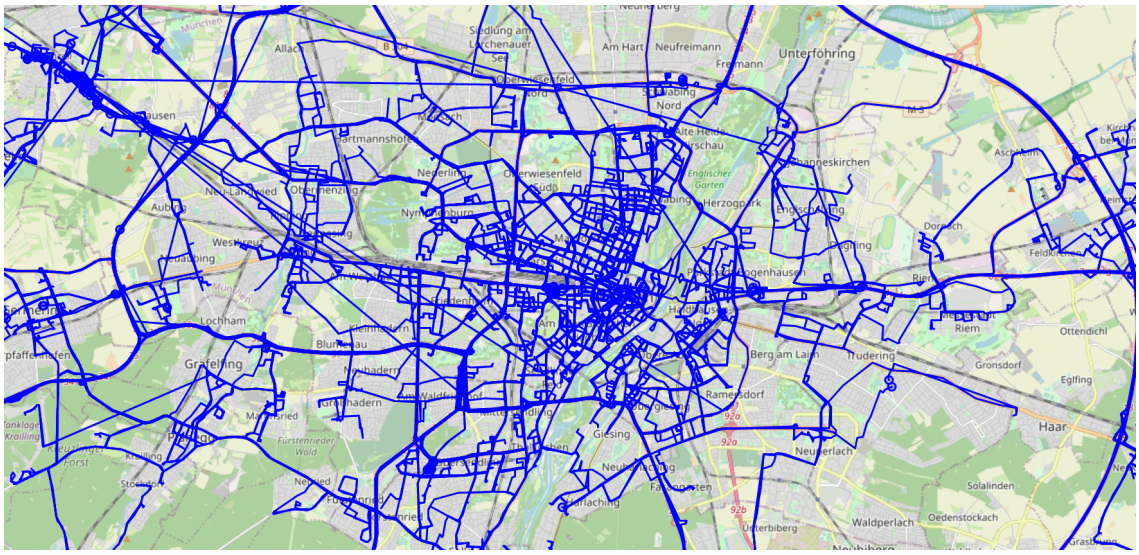
Entwicklung und Validierung eines Frameworks zur Verbrauchsberechnung von Mischflotten aus ICEV, HEV, PHEV, BEV und FCEV

Design and Validation of a Consumption Framework for Mixed Fleets Considering ICEV, HEV, PHEV, BEV and FCEV



Scientific work for obtaining the academic degree

Master of Science (M.Sc.)

at the Department of Mechanical Engineering of Technical University of Munich

Supervised by	Prof. Dr.-Ing. Markus Lienkamp Werner Schmid, M.Sc. Chair of Automotive Technology
Submitted by	Pol Masclans Abelló, B.Sc. Bonaplata 15 08034 Barcelona
Submitted on	April 16, 2021

Chair of Automotive Technology
Department of Mechanical Engineering
Technical University of Munich

Project description

Design and Validation of a Consumption Framework for Mixed Fleets Considering ICEV, HEV, PHEV, BEV and FCEV

The wide existing portfolio of powertrain technologies already in the market, which include different types of electrification and hybrid concepts, makes the choice of vehicles in a fleet management context an extraordinarily complex task. It is uncertain whether a technology is superior to the other in terms of total cost of ownership and greenhouse gas emissions, so only by defining the specific use case it is possible to optimize a fleet composition. This is particularly important for transporters, in which the payload may produce a substantial additional consumption that may not be neglected.

There are already many tools and methods for characterizing the mobility behavior of individuals, populations and vehicle fleets. In this regard, the large-scale fleet monitoring using satellite positioning such as GPS is a valuable technique able to capture the movement of individual vehicles with a high precision. How to use this data to optimize a vehicle fleet remains a fundamental challenge, as there is not a complete framework that allows to simulate the fuel consumption of a fleet using real driving data.

In this Master Thesis, a complete consumption framework for ICEV, BEV, HEV, PHEV and FCEV is developed and validated. In the literature research a suitable consumption model should be identified and then be instrumentalize to create a consumption framework for heterogeneous fleets. Using available real drive data from two existing vehicle fleets, the framework will be calibrated and proved, taking advantage of available real consumption data. The framework will be able to simulate two different types of situations: trips, that have been recorded with a GPS-tracking device including the corresponding speed profile, and “virtual trips”, for which no recodings and therefore no speed profiles are available.

The following points are to be worked on by Mr. Pol Masclans Abelló:

- Literature research on fuel consumption models and selection of a suitable simulation tool
- Development of the consumption simulation workflow.
- Development of a methodology for simulating the fuel consumption for “virtual trips”
- Assessing the impact of the payload on the overall fuel consumption and assuring the capability of the framework to account for it
- Validation of the framework using available real fuel consumption data
- Detailed analysis, evaluation, and documentation of the results

The elaboration should document the individual work steps in a clear form. The candidate undertakes to carry out the term paper independently and to indicate the scientific aids used.

Announcement date: November 16, 2020

Submission date: April 16, 2021

Prof. Dr.-Ing. Markus Lienkamp

Betreuer: Werner Schmid, M. Sc.

Chair of Automotive Technology
Department of Mechanical Engineering
Technical University of Munich

Geheimhaltungsverpflichtung

Herr: **Masclans Abelló, Pol**

Gegenstand der Geheimhaltungsverpflichtung sind alle mündlichen, schriftlichen und digitalen Informationen und Materialien die der Unterzeichner vom Lehrstuhl oder von Dritten im Rahmen seiner Tätigkeit am Lehrstuhl erhält. Dazu zählen vor allem Daten, Simulationswerkzeuge und Programmcode sowie Informationen zu Projekten, Prototypen und Produkten.

Der Unterzeichner verpflichtet sich, alle derartigen Informationen und Unterlagen, die ihm während seiner Tätigkeit am Lehrstuhl für Fahrzeugtechnik zugänglich werden, strikt vertraulich zu behandeln.

Er verpflichtet sich insbesondere:

- derartige Informationen betriebsintern zum Zwecke der Diskussion nur dann zu verwenden, wenn ein ihm erteilter Auftrag dies erfordert,
- keine derartigen Informationen ohne die vorherige schriftliche Zustimmung des Betreuers an Dritte weiterzuleiten,
- ohne Zustimmung eines Mitarbeiters keine Fotografien, Zeichnungen oder sonstige Darstellungen von Prototypen oder technischen Unterlagen hierzu anzufertigen,
- auf Anforderung des Lehrstuhls für Fahrzeugtechnik oder unaufgefordert spätestens bei seinem Ausscheiden aus dem Lehrstuhl für Fahrzeugtechnik alle Dokumente und Datenträger, die derartige Informationen enthalten, an den Lehrstuhl für Fahrzeugtechnik zurückzugeben.

Eine besondere Sorgfalt gilt im Umgang mit digitalen Daten:

- Für den Dateiaustausch dürfen keine Dienste verwendet werden, bei denen die Daten über einen Server im Ausland geleitet oder gespeichert werden (Es dürfen nur Dienste des LRZ genutzt werden (Lehrstuhllaufwerke, Sync&Share, GigaMove).
- Vertrauliche Informationen dürfen nur in verschlüsselter Form per E-Mail versendet werden.
- Nachrichten des geschäftlichen E-Mail Kontos, die vertrauliche Informationen enthalten, dürfen nicht an einen externen E-Mail Anbieter weitergeleitet werden.
- Die Kommunikation sollte nach Möglichkeit über die (my)TUM-Mailadresse erfolgen.

Die Verpflichtung zur Geheimhaltung endet nicht mit dem Ausscheiden aus dem Lehrstuhl für Fahrzeugtechnik, sondern bleibt 5 Jahre nach dem Zeitpunkt des Ausscheidens in vollem Umfang bestehen. Die eingereichte schriftliche Ausarbeitung darf der Unterzeichner nach Bekanntgabe der Note frei veröffentlichen.

Der Unterzeichner willigt ein, dass die Inhalte seiner Studienarbeit in darauf aufbauenden Studienarbeiten und Dissertationen mit der nötigen Kennzeichnung verwendet werden dürfen.

Datum: 16. April 2021

Unterschrift: _____

Chair of Automotive Technology
Department of Mechanical Engineering
Technical University of Munich

Erklärung

Ich versichere hiermit, dass ich die von mir eingereichte Abschlussarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Garching, den 16. April 2021

Pol Masclans Abelló, B.Sc.

Chair of Automotive Technology
Department of Mechanical Engineering
Technical University of Munich

Declaration of Consent, Open Source

Hereby I, Masclans Abelló, Pol, born on January 13, 1995, make the software I developed during my Master Thesis available to the Institute of Automotive Technology under the terms of the license below.

Garching, April 16, 2021

Pol Masclans Abelló, B.Sc.

Acknowledgments

I would like to express my gratitude to all the people who contributed in some way to the work described in this thesis.

First and foremost, I would like to thank my supervisor, Msc. Werner Schmid, who accepted me into his group and gave me the opportunity to work with him as a research assistant. I greatly appreciate the freedom you have given me to find my own path and the guidance and support you offered when needed.

I would like to extend my gratitude to Msc. Lorenzo Nicoletti and Msc. Benedikt Danquah. Their friendly guidance and expert advice have been invaluable at the first stages of the work.

I would like to acknowledge the Institute of Automotive Engineering at Technische Universität München (TUM). Both my graduate experience and this master thesis benefited greatly from the courses I took and the high-quality seminars that the department organized. I also would like to acknowledge my home university, UPC - School of Industrial Engineering of Barcelona (ETSEIB), for enabling me to participate in this double master's degree program.

I am extremely grateful for the funding sources (Erasmus + EU) that allowed me to pursue my master in this university.

Most of all I would like to thank my parents for their support during the whole master's degree and particularly during the elaboration of the Thesis. Last but not least, this work would have not been possible without the outstanding support, of my girlfriend Nuria, who had tirelessly helped, listen, advised and accompanied me during this two years.

Abstract

In this work a consumption simulation framework for ICEV, HEV, PHEV, BEV and FCEV is developed, tested and validated with real consumption data. The framework is based on FASTSim, a simulation tool created by the NREL to efficiently and accurately compare different vehicle powertrains. Two different versions are developed, one of them is oriented to a simulation using GPS recorded speed profiles, while the second one allows to simulate the fuel consumption of a trip just by introducing its origin and destination coordinates. For the testing and validation of the framework 30 different real vehicles from two vehicle fleets are simulated, with a total driving distance of 191 680 km. The version for estimating the consumption of recorded trips is validated using real consumption data of 103 trips driven by one of the simulated vehicles, for which its real consumption data was recorded using the Mercedes Pro platform. Additionally, a validation is carried out for each of the simulated vehicles using the available consumption rates of Spritmonitor. This version accounts for both road grade and additional masses. A comparative study of the influence of the payload masses of transporters in the overall fuel consumption is included in the work, showing that a 600 kg payload increases the fuel consumption by an average of 10 %. The version for simulating non-recorded trips is based in a consumption curve derived using real drive data. A series of speed bins are created and simulated individually for each vehicle, so the approach accounts for individual vehicle characteristics and even extra mass from payloads. A routing engine is used to determine the optimal route between the origin and destination and provide the model with the necessary data to estimate the consumption rate of the trip. This approach has similar or even better accurateness compared to the alternatives found in the literature and it is particularly well suited for performing large number of simulations.

Contents

List of Abbreviations

ADAS-RT	Advanced Driver Assistance Systems Research Platform
ADVISOR	Advanced Vehicle Simulator
BEV	Battery Electric Vehicle
CD	Charge Depleting
CMEM	Comprehensive Modal Emissions Model
CNG	Compressed Natural Gas
CS	Charge Sustaining
EEA	European Environment Agency
EMS	Energy Management System
EU	European Union
FASTSim	Future Automotive Systems Technology Simulator
FCEV	Fuel Cell Electric Vehicle
GIS	Geographic Information System
HBEFA	Handbook Emission Factors for Road Transport
HEV	Hybrid Electric Vehicle
ICEV	Internal Combustion Engine Vehicle
MAE	Mean Absolute Error
ME	Mean Error
MOVES	MOtor Vehicle Emissions System
NEDC	New European Driving cycle
NME	Normalized Mean Error
NREL	National Renewable Energy Laboratory
NTAE	Normalized Total Absolute Error
PHEM	Passenger Car and Heavy duty Emissions Model
PHEV	Plug-in Hybrid Electric Vehicle
SOC	State of Charge
US	United States
USEPA	United States Environmental Protection Agency
VSP	Vehicle Specific Power
VT-CPEM	Virginia Tech Comprehensive Power-based EV Energy Consumption Model
VT-CPFM	Virginia Tech Comprehensive Power-based Fuel Consumption Model
WLTC	Worldwide harmonized Light vehicles Test Cycle

Formula Symbols

Formula Symbols	Unit	Description
A	m^2	Frontal area
C_d	-	Drag coefficient
C_r	-	Rolling resistance coefficient
E_kin	J	Kinetic Energy
E_pot	J	Potential Energy
F_drag	N	Aerodynamical force
F_roll	N	Rolling resistance force
F_w	N	Force at wheel
a	$\mathrm{m/s}^2$	Acceleration
ϵ_i	kg	Translational mass of rotating components
g	$\mathrm{m/s}^2$	Acceleration of gravity
m	kg	Mass of the vehicle
v	$\mathrm{m/s}$	Speed
v_w	$\mathrm{m/s}$	Speed of wind
α	-	Road grade
ρ_air	$\mathrm{kg/m}^3$	Air density

1 Introduction

1.1 Motivation

The transportation sector accounted for 24 % of worldwide direct emissions from fuel consumption in 2019[1]. Particularly, road transport is responsible for three quarters of the entire sector's emissions [1]. In Europe, transport is the only sector whose CO_2 emissions are above 1990 levels. According to Transport and Environment [2], it is crucial to take immediate action in order to achieve carbon neutrality by 2050 and ensure that the increment of global temperature is kept well below 2 °C with respect to the preindustrial situation, as targeted on the Paris Agreement [3]. As a response to this issue, new powertrain technologies, such as Hybrid Electric Vehicle (HEV), Plug-in Hybrid Electric Vehicle (PHEV), Battery Electric Vehicle (BEV) and Fuel Cell Electric Vehicle (FCEV) have arisen, thus, placing the spotlight on the concept of electromobility. electromobility is defined as a road transport system based on vehicles that are propelled by electricity, which can be produced both inside or outside the vehicle [4]. This new paradigm opposes to the conventional Internal Combustion Engine Vehicle (ICEV), which has been the leading technology for more than a century.

The rise of these new technologies may pose a dilemma to the fleet managers: Which technology or mix of technologies is the most adequate for a vehicle fleet? To address such question, a model to compute an optimize a fleet composition, considering the mobility behavior of a fleet and the existing technologies in the market is needed. The target values to minimize are CO_2, energy consumption and total cost of ownership. This way, the mentioned objectives of greenhouse gases emission reduction are aligned with economical competitiveness, which is key to accelerate the path to a sustainable economy.

One of the fundamental modules of the model is the energy and fuel consumption estimation framework. Vehicle consumption estimation is a recurrent topic in the literature due to its many applications. Governments, researchers, manufacturers and businesses are all interested in assessing the energy consumption of its fleets both for economic and environmental reasons. It is, however, also a complex issue due to the many parameters that influence energy rates. Particularly, Ahn [5] classifies these parameters into six categories, relating them to travel, weather, vehicle, roadway, traffic or driver factors. Thus, an energy and fuel consumption estimation can strongly vary in its form and accuracy depending on the specific objectives of a project and the available resources. Therefore, there is not a one-fits-all solution, but an extensive set of approaches that tackle the majority of necessities.

This thesis provides a framework to estimate the energy and fuel consumption produced by each of the vehicles of a real fleet. By using an existing consumption model in the literature, the framework estimates the rates of the specific technologies and vehicles. It employs GPS recordings of individual trips as the main input data source, however, it also provides an alternative procedure for calculating the energy consumption for non-recorded trips. Consequently, this framework is prepared to be used in a fleet composition optimization algorithm.

1.2 Objectives

The objective of the project is to create a consumption simulation framework compatible with ICEV, HEV, PHEV, BEV and FCEV. Its input data will be the speed profiles derived from GPS data and otherwise, an origin and destination point. To this end, the following sub-objectives must be met:

1. Create a comprehensive and extensive virtual fleet that represents an existing one and includes the above-mentioned powertrain technologies. Also validate this fleet using the homologation data.
2. Create and validate the required input data for the estimation using the available database, so the mobility behavior of each vehicle can be fully characterized. This includes generating accurate speed profiles out of the GPS recorded trips and the alternative procedure for the virtual trips.
3. Validate the results of the consumption estimates calculated with GPS data by using real consumption results. These are the real consumption data from a subset of one of the studied fleets, from the Mercedes Pro platform, and publicly available consumption data rates from the website Spritmonitor.
4. Validate the consumption estimation framework for trips in which only the origin and destination coordinates are available.
5. Examine the impact of the extra load on transporters.

1.3 Structure of the thesis

The present work is divided into seven chapters, including the introductory chapter. The structure follows the standard research paper approach, with a state of the art chapter, followed by the methodology, the implementation, the results, the discussion and finally the conclusions.

The state of the art is based on an extensive literature review that covers five major topics, each of them introduced in a separated section. The first section describes the classification of consumption models and gives an overview of the main use cases for each category. The second section refers to the calculation of the vehicle specific power, a parameter frequently used for estimating fuel consumption and emissions of road vehicles. The third section presents the most relevant fuel consumption models in the literature and separates them accordingly to a scale criteria. The fourth section of the literature review presents some relevant use cases of synthetic speed profiles generation and fuel estimations without using previously recorded trip data. Finally, in the fifth section some relevant characteristics of the simulation of PHEV are introduced.

The methodology is an abstract description of the tools and methods used in this thesis for completing the objectives. It is divided in two sections: the first one addresses the methodology for estimating the fuel consumption of the GPS-recorded trips. It includes the discussion of which model to use, the description of the real vehicle fleets that are simulated, an explanation of the approaches used for solving some experimental data related issues and the validation of the results. The second section describes the methodology to estimate the fuel consumption for trips

using only an origin and a destination point, which includes the tools and the approach used to validate the results.

The implementation is an in-depth description of the elaboration and usage of the simulation framework. It has three sections, which correspond to the three main aspects of the present work that required the most effort. The first section explains the process of creating a representative virtual fleet using the available resources. The second section describes the algorithms for querying the trips and creating usable speed profiles. The third section describes the implementation of the methodology for estimating fuel consumption in non-recorded trips. In each section, additionally, the alternative implementations that are finally not part of the consumption framework are also introduced to provide a full understanding of the development process.

In the results chapter all the relevant findings are presented, divided in 5 different sections. The first section presents the results of the virtual fleet, which includes a calibration using the official homologation data. The second section shows an overview of the obtained trips using the GPS recorded trips. Since there are thousands of trips, this section only presents some relevant examples. The third section presents the results of the consumption simulation framework using the recorded trips. This includes the validation using the Mercedes Pro data set, the validation using the Spritmonitor data and the assessment of the influence of load mass in transporters. The next section is the presentation and validation of the consumption results for non-recorded trips. The results of alternative implementations are also presented for documentation purposes in each of the above-mentioned sections. Finally, the last section of the results chapter introduces the error chain, a concept used to describe and quantify the different sources of error present in the estimation procedure.

The discussion chapter contextualizes the methodology and the results with the state of the art and discusses the results and findings of the thesis. It is divided in 6 blocks covering the creation of trips using GPS data, the virtual fleet, the validation using the Mercedes Pro data, the validation using the Spritmonitor data, the error chain and the consumption estimation using non-recorded trips.

Finally, in the last chapter the conclusions and outlook of this thesis are discussed.

The document also includes two appendixes, the first contain additional tables and plots necessary for the documentation of the work. Particularly, it includes additional virtual vehicles that were created during the elaboration of the thesis, but did not take part of the results, the parameters of all the virtual vehicles, some maps of the fleet tests and the speed profiles used for the consumption curve creation, which is described later in the thesis.

The second appendix contains a short description of the documents found in the SD card that is included in the printed version of this thesis.

2 State of the Art

In this chapter, the theoretical basis for fuel consumption estimation is introduced. The findings from the literature review are presented in five different sections, starting with an overview of what is meant by a consumption model, which are its uses and how are the different models categorized. The second section defines the vehicle specific power, its physical meaning, and why its a relevant parameter for many consumption models. This is a necessary definition for understanding some of the differences between the consumption models presented in the third section. This part is divided in two subsections according to the type of model and present the most common models found in the literature. In the fourth section the topic switches to the existing approaches for estimating energy consumption without a defined speed profile. Finally, in the fifth section some relevant aspects to be considered during the simulation of hybrid vehicles are introduced.

2.1 Classification of Consumption Models

The proper estimation of vehicle fuel consumption and tailpipe emissions has been discussed for many years in the literature, since it has applications in many different fields. Fuel consumption and emission models are used, among many, for the calculation of emissions by the authorities, the simulation of mobility scenarios, the assessment of infrastructure projects or the analysis of emerging vehicle technologies [6, 7].

The calculation of emissions by the authorities are mostly reported annually as the so-called emission inventory reports and are the basis for many national or even international policy agreements regarding air quality and global primary energy consumption. In many cases, such as in the European Union or the United States, their respective environmental protection agencies are in charge of both the development of the models for creating the inventories and also the generation the periodical reports. The simulation of mobility scenarios, which is closely related to the assessment of new infrastructure projects and existing infrastructure management, includes often a consumption estimation. An example of this is the HDM-4, a widely used decision support tool for highway development and management created by the World Bank [8].

There are many different classification criteria for the consumption models, being the scale one of the most widespread used. In this regard, authors distinguish between Macroscale and Microscale models and, in some cases, even Mesoscale models as a category. According to Faris [9], a model is said to be Macroscopic when it uses aggregated fleet and network-wide parameters to estimate fuel consumption and emission rates. These kind of models are usually conceived of creating the above-mentioned emission inventories at a national or regional level. Microscopic models, on the other hand, use speed profiles and acceleration data to estimate the fuel consumption and pollutant emissions of a single vehicle [10]. The concept of a Mesoscale model is more diffuse. Yue [10] defines Mesoscopic models as those whose scale lie between Macro (a full network

simulation providing fleet-wise results) and Micro (a single vehicle performing a particular drive cycle). However, Chen [11] provides a more exhaustive definition. In particular, Mesoscopic models are said to combine the advantages of the previous two by applying real driving data to a microscopic model. They create a statistical distribution of fuel consumption and emission as a function of average speed, that can be used to fulfill similar needs of a purely Macroscopic model, namely, inventories or transportation policy assessments. This way, an evaluation of fuel consumption in a particular network is done without conducting a traffic simulation. In practice, all models can be in some way considered to be a Mesoscopic model depending on its usage, so for the sake of simplicity this paper will not consider Mesoscale as a category. This does not mean that the Mesoscopic models are ignored in this thesis, but simply that instead of labeling each of the models as Microscopic and Mesoscopic, they will be considered simply as Microscopic as their usage in a Mesoscopic scale can always be assumed as possible.

A complementary classification criteria is the model transparency. Authors usually distinguish between white-box models, black-box models and grey-box models [12]. The white-box models are based on either the engine physical process, or its chemical process. Its mathematical framework is highly deterministic and contains an extensive set of modules that require a full understanding of the system. All the relevant existing subsystems in a vehicle are parameterized and simulated, so they are usually difficult to calibrate and require an extensive set of parameters, some of them are difficult to obtain. On the other hand, in a black-box model the entire vehicle (including its engine) is modeled as one without describing each of the sub-processes individually. Finally, the grey-box model is a hybrid version of the former two, in which the power unit is modeled as a black-box, but the vehicle longitudinal dynamics related subsystems are fully described and parameterized [12].

For microscopic models, a further classification exists, which distinguishes between data-based and physics-based models [13]. The data-driven methods can be divided between regression models and the so-called lookup tables. The regression models are created by adjusting the output parameters such as the consumption rates to the input variables, which could be instant speed or wheel power demand. This category includes neural networks and similar artificial intelligence approaches, that are becoming more common. The look-up table approach consists on deriving a consumption or emission matrix for a vehicle by using chassis dynamometer test data. This way, a table filled with emissions factors expressed as a function of a parameter is created. Common parameters are the average speed or he Vehicle Specific Power (VSP). Normally, a look-up table is created for each vehicle category, and in some cases for a same vehicle category different tables are created considering cold or hot emissions, driving style etc.

Regarding the physical-based models, a distinction between deterministic and stochastic models is made. The first rely purely in a vehicle dynamics simulation and are based on the engine power demand that is introduced in the following chapter, while the probabilistic methods, besides relying on the dynamics also consider stochastic components. More precisely, the probabilistic methods are based on random velocity disturbances that intend to replicate the variability in a speed profile [13]. They differ to the data-driven approach as the data used in a probabilistic model is based in assumptions and a random distribution, while the data-driven models use data that is gathered in large scale experimentation.

In addition to the academical classification, a key factor to differentiate the numerous existing models is its scope. There are two aspects to consider in this regard. First, some of the models, specially the microscopic, are only intended for a specific powertrain technology. This means that they are exclusively designed to capture the behavior for instance, of a ICEV, ignoring

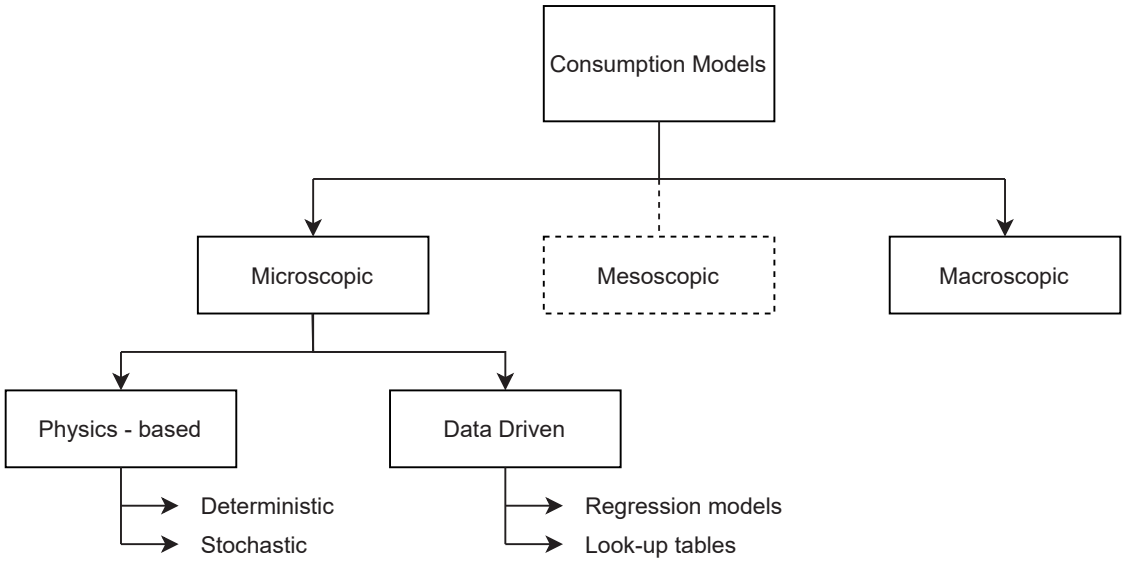


Figure 2.1: Classification of consumption models.

other existing technologies. This same happens with the vehicle typology, as some models are exclusively intended for passenger cars while others focus on high duty vehicles.

Secondly, the output variables must also be considered. Each model may simulate a broad spectrum of tailpipe emissions, such as CO_2, NO_x, CO, PM as well as energy or fuel consumption, or just one of those magnitudes. As a rule, the Macroscopic models, that are usually intended for the emission inventories, cover the majority of emissions, whereas the Microscopic models are likely to focus on just a few. In this regard, at the beginning of the section, the term fuel consumption and tailpipe emission model is mentioned, however the scope of the project is limited to the fuel and energy consumption estimation. The reason for that, a part of providing a wider overview of the topic, is that the CO_2 emission is proportional to the fuel consumption, and thus, a solely tailpipe emissions estimation model, if includes CO_2, is also valid for fuel consumption estimation.

Before going into detail to the consumption models in the literature, a closer look to the physics consumption estimation model is needed. This explanation is provided in the following section.

2.2 Vehicle Specific Power

The complete characterization of a trip for energy estimation purposes usually requires of speed profiles at 1 Hz frequency. This means, a time series containing both the elapsing time and the instantaneous speed in a second-by-second fashion. This information is enough to derive the acceleration profile and the traveled distance. In this regard, it is important to note the difference between a driving cycle and a trip. A driving cycle is a time series that represents driving patterns that describe the workloads imposed on the vehicles to assess its performance and its environmental impact [14]. In the context of the present work, a trip is simply a speed profile describing the movement of a vehicle from an origin to a destination. However, the literature also refers to the trips as driving patterns, speed profiles, real world drive cycles or even drive cycles. For the sake of clarity, in this work the concept of drive cycle will specifically refer to the definition by Huertas [14].

A complementary parameter for trip characterization found in the literature, is the VSP. This parameter was first used by Jimenez-Palacios [15] for modeling vehicle emissions and is also found to be representative for estimating fuel consumption [16]. In fact, many of the models that are later discussed use VSP distributions as the main descriptive variable of a trip. The physics behind this parameter are well-known to the scientific and engineering community as it is just a reinterpretation of the equation commonly used to calculate the wheel power demand of a vehicle. However, the novelty and relevancy of VSP lies in the fact that it is a measurable, one dimensional and physically meaningful parameter, that allows to establish a comparison and generalization between different vehicles. Here, by one dimensional it is meant that a VSP approach depends only on the power, while the also common alternative of using traditional engine maps requires of a combination of the torque and the rotational speed. Later on, in Figure 2.2 an example of an engine map and the advantages of a vehicle power approach over it are illustrated.

The wheel force in a vehicle is described as the necessary force so a vehicle can overcome all the resistances against him and move. These resistances are the rolling resistance, the air resistance and the road-grade resistance. Sciarretta [17] formulates the wheel force as in Equation (2.1):

$$F(v) = \frac{1}{2}\rho_{air} C_d A v^2 + C_r m g + m(g \sin\alpha) \tag{2.1}$$

Where F is the resistance force at the wheels, v the vehicle speed, C_r the rolling resistance coefficient, m the vehicle mass, g the gravitational acceleration, ρ_air the air density, A the frontal area of the vehicle, C_d the drag coefficient and α the slope of the road. Additionally, one can calculate the wheel power by simply applying Equation (2.2).

$$P = F v \tag{2.2}$$

Coming back to the VSP, in Jimenez-Palacions, [15] the approach for determining the wheel power is slightly different. Instead of considering the acceleration and road grade resistance as forces, they are defined as energies. Then by dividing by the mass, they obtain the specific power, as shown in Equation (2.3).

$$\begin{aligned} VSP &= \frac{\frac{d}{dt}(E_{kin} + E_{pot}) + F_{roll} v + F_{drag} v}{m} \\ &= v(a(1+\epsilon_i) + g(cos\alpha + C_r)) + \frac{1}{2}\rho \frac{AC_d}{m}(v - v_w)^2 v \end{aligned} \tag{2.3}$$

Here ϵ_i is the equivalent translational mass of the rotation components and v_w is the headwind into vehicle, that is usually neglected. Following this introduction to the basic physics principles in consumption simulation, the overview of the existing models in the literature is provided, in which the term VSP is recurrently used.

2.3 State-of-the-art Consumption Models

Several models are currently being used to perform a broad range of tasks. As defined in previous sections, models can be either Macro or Micro scale and particularly for Micro models, many different sub-types exist. The benchmark presented in Subsection 2.3.1 and 2.3.2 weights the relevancy of each model considering its suitability for the project and its relevance in the

literature. The classification distinguishes between Macroscopic and Microscopic models in different subsections to allow for a fair comparison. Most of the models are open source, but there are also some cases in which the software distributions are subject to periodical fees or subscriptions. This literature review only considers a small subset of licensed models that has been repetitively found in the literature.

2.3.1 Macroscopic Models

In this subsection 4four different macroscopic models are presented, namely COPERT, EMFAC, MOVES and HBEFA. The models are summarized in Table 2.1.

COPERT is the european tool to calculate transport related emissions for creating national inventories and assess the impact of policy measures. It was developed at the Aristotle University of Thessaloniki and funded by the European Environment Agency (EEA). Originally, it was published as an emission inventory guidebook called CORINAIR and COPERT is its official software implementation. It provides emission factors for all the existing vehicle technologies currently in the market. The values are periodically updated using chassis dynamometer testing. Its main limitations are, on the one hand, the use of constant speed as the input variable and on the other hand, the large confidence intervals of its results, originated from the existing variability in cars within a segment or typology [18]. Additionally to the standard distribution, a so-called COPERT Micro distribution was introduced, that allows to estimate emissions at a road-link level [19]. Nevertheless, this distribution doesn't match our definition for microscopic models, as it uses aggregated vehicle data and network parameters to perform calculations. This means that it is not able to capture the consumption behavior of an individual vehicle.

A similar approach to COPERT is found in the EMissions FACtor (EMFAC) tool developed by the California Air Resources Board. It is designed to be the official tool for creating emission inventories in the state of California and also uses the average speed as a main input. It provides actualized emission rates for the existing vehicle fleets in different aggregation levels, from a state wide level to a county level [20]. Its main drawback is that it is conceived to be used in California, so adapting the model to anywhere else would result in losing its singularities, particularly the ability of the model to provide emissions factors that consider the real composition of a regional fleet. Moreover, such attempt would require a significant effort.

The United States Environmental Protection Agency (USEPA), the American counterpart to the EEA, developed in 2003 the MOtor Vehicle Emissions System (MOVES) to provide a versatile tool for performing emission estimation at various levels. The model was designed as the evolution of the MOBILE series, the previous American reference model that was superseded by MOVES on 2004. It incorporates the novelty of considering VSP distributions as the descriptive parameter for a given drive cycle instead of using the average speed, which was the standard parameter at that moment. Its latest version is MOVES3 and it is regularly updated with the current vehicles powertrain technologies and the actualized rates. [21]. The model can have a speed profile as an input so it is said to be suitable at any scale, however it is as a Macroscopic model when it highlights the most as it is the first model of this kind that adopted the VSP as descriptive variable instead of using the average speed. The estimation in MOVES are based in 14 VSP bins, which are calculated for each of the several different sources. Specifically, for each vehicle type, the bins are calculated considering the fuel type, the engine technology, the model year (grouped according to the existing standards), the loaded weight, the engine size and the emissions technology. Oposite to the COPERT Micro, the use of VSP as input parameter makes it indeed a Micromodel, as it

is able to simulate the consumption of individual vehicles. MOVES is still nowadays used as a reference when developing an energy consumption and emission estimation model [22, 23].

Finally, some authors have also used the Handbook Emission Factors for Road Transport (HBEFA) for simulating vehicle emission and fuel consumption [24]. This model consist of an extensive database of emission factors for all existing vehicle segments and typologies and was developed by a joint effort of the environmental protection agencies of Germany, Switzerland and Austria later also supported by the European Union (EU). It works as a look up table of emission rates able to distinguish between different traffic situations and it is regularly updated [25]. The complete version, however, is under license.

Table 2.1: Overview of the described Macro models.

Model	Release Year (Last update)	Simulation type	Developer	Main Usage	Available data
COPERT	2005 (2020)	Lookup table: constant speed bins	EEA	EU standard vehicle emission model	Yeary actualization of energy rates for all vehicle types
EMFAC	2000 (2017)	Lookup table: constant speed bins	CARB	Emission inventories in California	Periodic actualization of energy rates and fleet data for existing technologies
MOVES	2003 (2020)	Lookup table: VSP and speed bins	USEPA	Energy and emission inventories for authorities	Yeary actualization of energy rates for all vehicle types
HBEFA	1995 (2019)	Lookup table: constant speed bins in different driving situations	AFRAS	Energy and emission inventories for authorities	Yeary actualization of energy rates for all vehicle types (Licensed)

2.3.2 Microscopic Models

In this section, seven microscopic models are introduced, ADVISOR, CMEM, the VT-models, the University of Cork model, FASTsim, PHEM and MOVESTAR. The models are summarized in Table 2.2.

In 1994 the National Renewable Energy Laboratory (NREL) from the United States (US) developed a system analysis tool named Advanced Vehicle Simulator (ADVISOR) to quantify fuel economy, vehicle performance and tailpipe emissions using alternative powertrain technologies. The tool was later released to the public to be used in a Matlab/Simulink environment and has been maintained by its developers up to 2003 and later by its own users. ADVISOR is a white-box model based in engine maps that combines backward and forward simulation to deliver very accurate results in many different fields. Their main usage was the comparison and assessment of different powertrain technologies. It can simulate Gasoline and Diesel ICEV, FCEV and HEV in various architectures. Its high level of detail is also its major disadvantage, as it is difficult to calibrate and requires an extensive parametrization. Moreover, the complexity of the model leads to a poor time performance even when simulating short trips. Despite its lack of actualized data, it is still one of the most used microscopic models in the literature nowadays [26].

The University of California at Riverside developed the Comprehensive Modal Emissions Model (CMEM) to fulfill the need for microscopic emissions modeling in the late 1990's. This model is based on a power demand approach and uses a physical foundation to characterize the relevant

existing processes involved in the fuel consumption and the tailpipe emissions of vehicles. Its latest version dates from 2005, includes 28 light-duty ICEV categories [27, 28]. According to Dowling [29, p.94] it was said to be the most detailed model for estimating exhaust emissions. However, some authors claimed recently that, despite the ability of CMEM to accurately simulate emissions and fuel consumption, it is too difficult to calibrate it due to the complexity of acquiring the necessary parameters. Therefore, its ability to perform large estimations is disputed [30].

The Virginia Tech Comprehensive Power-based Fuel Consumption Model (VT-CPFM) is a microscopic vehicle fuel consumption and emission estimation tool for ICEV. It is a black-box and power based model introduced in 2011, that can be calibrated by using publicly available data. It has two variants, the VT-CPFM-1 and the VT-CPFM-2. The first is just a simplified version of VT-CPFM-2 that does not require any engine specific data. Most of the parameters required in the second version are usually publicly available [31]. However, more complex values such as drivetrain efficiency or wheel slippage are calibrated using commonly accepted values in the literature and therefore are not vehicle-specific. The model validation dating between 2001 and 2011 indicates a satisfactory estimation accuracy that makes the model suitable to be integrated to traffic management tools, to perform energy assessments or be used in similar applications [32]. The same department in charge of the VT-CPFM development introduced in 2015 a new variant for EV which was able to capture regenerative breaking. This new model is called Virginia Tech Comprehensive Power-based EV Energy Consumption Model (VT-CPEM). It is also a black-box model using a quasi-steady backward approach specifically meant for microscopic simulations [33].

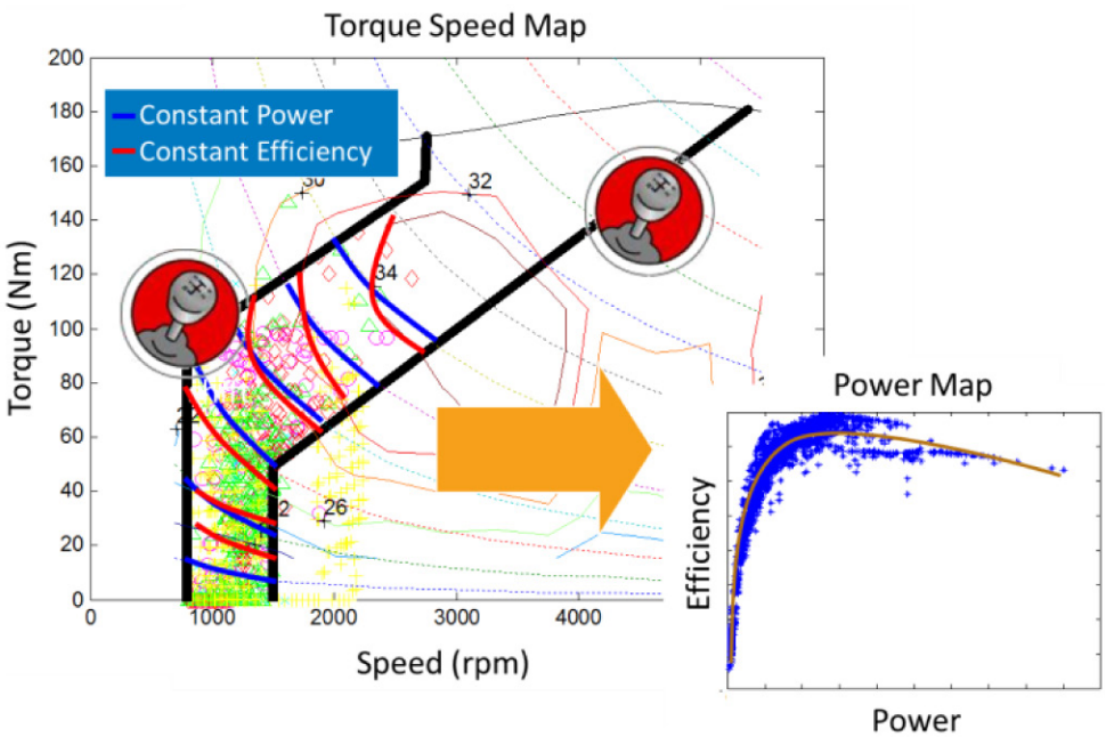


Figure 2.2: The single efficiency curve approach in fastsim as alternative to traditional engine maps [34].

A slightly different approach was used by the researches from the university of Cork to develop its EV consumption model for specific vehicles. In their work they used a simplified white-box model and assumed standard values for efficiency, inertia and some other engine related

parameters. Instead of using power, the input parameter of the model is the required vehicle speed. They presented the full methodology with a Nissan Leaf and the Tesla Roaster and achieved excellent correlations between the experimental results and the model prediction [35]. In 2014 an improved version of the model was published, which uses the coast-down parameters provided by USEPA to enhance the model performance. Using this parameters, the vehicle load force can be calculated as a function of the vehicle speed and account for additional speed related vehicle losses that can not be modeled using solely the aerodynamic drag coefficient and the rolling resistance [36].

In 2015 the NREL, the same department that developed ADVISOR, introduced the Future Automotive Systems Technology Simulator (FASTSim) to provide an efficient, accurate and robust tool for comparing vehicle powertrains. It includes conventional ICEV, HEV, PHEV, BEV, FCEV and even Compressed Natural Gas (CNG) vehicles, using high level parameters to model the relevant subsystems. The authors showed that a single efficiency curve for each engine type, scaled according to its power, can represent most engines with a 5 % error. This is visualized in Figure 2.2, where a typical engine map, also described as torque-speed map, relates the rotational speed (horizontal axis) and torque (vertical axis) with the efficiency contour of an engine. The figure shows that the gearing strategy forces the engine to stay in a narrow working area, in which the constant power and constant efficiency isolines are almost parallel. Thus, when plotting the engine working points as a function of the power and efficiency, as in the right hand side smaller plot, they follow a curve. This way, the complete powertrain can be effectively described using one curve, in which not only the engine power, but also the gearing is being indirectly simulated. The validation for a complete vehicle shows that FASTSim consumption estimates are most of times below 5 % and in almost all cases within 10 % when using the official and publicly available homologation data and without the need to calibrate the vehicles. The model was conceived to efficiently perform several different simulation types, including batches of drive cycles and a Python distribution was recently published, making this a very versatile tool [34].

Recently, the University of California at Riverside developed an open-source model named MOVESTAR which is based in the existing MOVES model. It is conceived as a user friendlier version of the already presented MOVES, which is usually difficult to operate for inexperienced users. It focuses on MOVES' ability to perform microsimulations of different vehicle typologies using the already mentioned VSP bins. As input data for a simulation the speed profiles must be provided at a 1 Hz frequency, however it doesn't account for the road grade, which is assumed as 0. The user can chose between 13 vehicle models and for each vehicle, between six fuel types [23].

Finally, it is worth mentioning the Passenger Car and Heavy duty Emissions Model (PHEM) developed by the Graz University of Technology in cooperation with FVT GmbH, a vehicle longitudinal dynamics simulation tool that uses a backward approach for deriving vehicle emissions and fuel consumption. It uses a white box approach able to fully describe the engine emission behavior using engine maps and even taking into account exhaust aftertreatment elements. The model has a time resolution of 1 Hz and includes traditional ICEV as well as electrified powertrains and other new technologies. It has a variant particularly meant for Micro simulations of single vehicles, the PHEMlight, which also has a reduced complexity. It is however a licensed software [37, 38].

Table 2.2: Overview of the described Micro models.

Model (Developer)	Year (Last update)	Simulation type	Main Usage	Available data	Powertrain technologies
ADVISOR (NREL)	1992	White-box, Engine Map based, backward-forward	Fuel estimation and vehicle performance comparison	Vehicle specific data (Outdated)	All existing technologies
CMEM (University of California)	1999 (2005)	Grey-box, VSP-based	Instantaneous fuel consumption and emission estimation	Only calibrated parameters, vehicle specific parameters not provided	ICEV
VT-CPFM (Virginia tech)	2011 (2013)	Black-box model, Power-based	Instantaneous fuel consumption and emission estimation	Calibration with Homologation data	ICEV
VT-CPEM (Virginia tech)	2016				BEV
EV Model (University of Cork)	2011 (2014)	White-box model with power-efficiency curve	BEV range estimation	User data	BEV
FASTSim (NREL)	2015 (2020)	White-box model with power-efficiency curve	High level powertrain analysis	Default vehicles, expandable with public available data	All existing technologies
PHEM (TU Graz)	2003 (2016)	White-box, Engine Map based	Instantaneous fuel consumption and emission estimation	Extensive set of vehicle, HBEFA data (licensed)	All existing technologies
MOVESTAR (University of California)	2020	VSP and speed bins, based in MOVES	Development of advanced vehicle technology	13 vehicle types	All existing technologies

2.4 Generation of synthetic speed profiles for Energy Consumption

The prediction of speed profiles, also referenced in the literature as synthetic speed profiles have been used to optimize the energy management of hybrid vehicles, to predict the vehicle energy consumption and also for eco-routing purposes [39]. The predictions can either be short-term, this is, within the scope of the driver's vision and sensor perception, or long-term, which includes a full trip prediction. Moreover, the predictions can be either Mathematical based or data-driven [39, p. 48]. In this work, the focus will be on long-term predictions, since only those are able to describe a complete drive cycle from start to finish. Following some relevant use cases will be introduced, the first and second of them are mathematical based and the remaining three are data driven.

A methodology to calculate emissions with simulated Traffic Conditions was introduced by Hirschmann [40]. Hirschmann coupled the traffic simulation model VISSIM with the instantaneous emission model PHEM to provide assessments of the environmental impact of traffic networks, improve management strategies and assist technology implementation. Their case study was limited to an urban arterial of the center of Graz and its surroundings, including 12 intersections. The traffic simulation model was enriched with road features such as stop lines, pedestrian and cyclist facilities and bus stops. Also GPS recordings where used for calibration purposes. Their implementation showed a fuel consumption reduction potential of up to 12 % with an optimized traffic lights management.

Karbowski presented a research work done at the Argonne National Laboratory, in which a process able to predict the energy consumption of a vehicle following a realistic user defined trip

was introduced [41]. For doing so, the Advanced Driver Assistance Systems Research Platform (ADAS-RT) was used to create the trips and import relevant road features from a digital map and the simulation tool Autonomie to run and analyze the simulation. They included features such as the average traffic speed, the speed limits, the stops related to intersections and traffic lights as well as realistic road grade profiles to enrich the drive cycle.

The disadvantage of these works is that both the traffic modeling tools used, VISSIM and ADAS-RT, as well as the consumption models, PHEM and Autonomie, are licensed software which make it hard to reproduce the presented methodology at a reasonable cost.

Boriboomsim [42] developed an eco-routing navigation system able to determine the most eco-friendly route between a trip origin and destination. Their approach combined a microscopic emissions model, CMEM, with a large vehicle activity database. They estimated the consumption of this extensive set of drive cycles using CMEM and they obtained the average fuel consumption as a function of the average speed throughout a link. The results were assigned to each road link and were passed to the routing engine. For their model they used congestion data from a real-time traffic information database and several probe vehicles, which allowed them to achieve fuel consumption estimations with an error of around −15 %. Note that in this case a negative value means that there was a fuel consumption underestimation.

Grubwinkler, from the TUM's Institute of Automotive Engineering presented its own system to predict EVs energy consumption using crowd-sourced speed profiles. In their work, data from 14.000 tracks with a total length over 200 000 km was used to create and enhanced digital map of Munich for routing. The energy consumption is calculated using vehicle-characteristic power consumption maps and a probability distribution function of the vehicle speed and accelerations at each road type. The results were applied to the map in links of around 18 km long [43]. Their more accurate predictions had an error of 7 % [44].

Finally, Holden [45] presented a data driven energy estimation model for trips before they are driven. They gathered nearly 100 million real drive spatial points acquired at a 1 Hz frequency and enhanced them with road grade data and other relevant road features, such as the speed limit, the road type or even the type of turn. Using the resulting cycles as input for the FASTSim model, they determined its second-by-second fuel consumption and created the so-called passes, a link level aggregation of the fuel consumption. The passes were then labeled and sorted according to the above mentioned features, namely the road type, the speed limit of the road and the type of turn (left, right, no turn). Thus, fuel consumption results were associated to its corresponding 10 mph bin, resulting in a lookup table which was the basis for the energy prediction. By using only speed limit and road type, they achieved a 17 % Normalized Total Absolute Error (NTAE). By incorporating the road grade, results improved to 13 % and the link orientation (type of turn) increase the accuracy even more, up to 12.5 % NTAE.

2.5 Hybrid electric vehicle simulation

Since the hybrid vehicles, both HEV and PHEV, combine an electric motor with an internal combustion engine, there is a need for defining a control strategy that specifies where to retrieve the power from at each instant. This is strategy is referred in the literature as Energy Management System (EMS). The control strategies are usually divided in two operating modes, the Charge Depleting (CD) mode and the Charge Sustaining (CS) mode. In CD, the vehicle operates entirely or partially using the electric drive in a way that there is a net energy loss from the batteries,

resulting on a reduction of the State of Charge (SOC). In CS, the major part of the energy is provided by the engine and the electric drive is used only as momentary storage to improve fuel economy [46, 47].

The EMS is not trivial and may lead to huge differences in the overall fuel consumption, so when simulating hybrid electric vehicles, the control strategy algorithm and the energy management strategy should be always considered. For achieving an optimized energy consumption in HEV and PHEV, a complete deployment of the battery stored energy is obviously necessary. However, simply operating on a CD mode, consisting of using as much electric power as possible until emptying the batteries and then running purely on the engine power is not optimal [39, p. 90]. There are two classes of EMS for hybrid vehicles, the reactive strategy and the route-based strategy. The reactive EMS can only generate approximate optimal solutions, since they only use the current driving conditions. The route-based EMS, on the other hand, use the full trip information to optimize the performance of the controller, which can drastically improve the fuel economy specially in PHEV [48]. In an extensive review of existing EMS, Martínez determined that PHEV may only have an optimal fuel economy when a detailed information about the future route of the vehicle is fully available [49].

This is due to the fact that the benefits in terms of fuel economy are highly related to the load profile of the trips. Fontaras showed that HEV can obtain consumption reductions in urban driving up to a 60 %, whereas the reductions at higher speed are much lower, so the fuel economy of HEV was found to be equivalent to those of conventional cars when driving above 95 km/h [50]. Moreover, the existence of a limited size battery makes the trip distance together with the battery recharging as key factors when determining the overall consumption of PHEV. Silva showed that for a daily trip of 20 km, the total CO_2 could more than double throughout the vehicle life cycle when the driver didn't charged his vehicle frequently [51].

3 Methodology

In this chapter an abstract description of the methodology is presented. The methodology defined in this chapter refers to the definitive version of the framework, disregarding the elaboration process and the studied alternatives, which are also presented in the results, but were found to be not optimal. This alternative methods are defined in the implementation and later also presented in the results. The methodology is divided in two sections. First the procedure for estimating fuel consumption out from the GPS-derived speed profiles is defined, consisting in the simulation tool, the identification and configuration of vehicle fleets, the generation of speed profiles and the validation of the results. Secondly, a consumption estimation methodology for trips that do not have a recorded speed profile is detailed. The complete Framework is summarized in Figure 3.1.

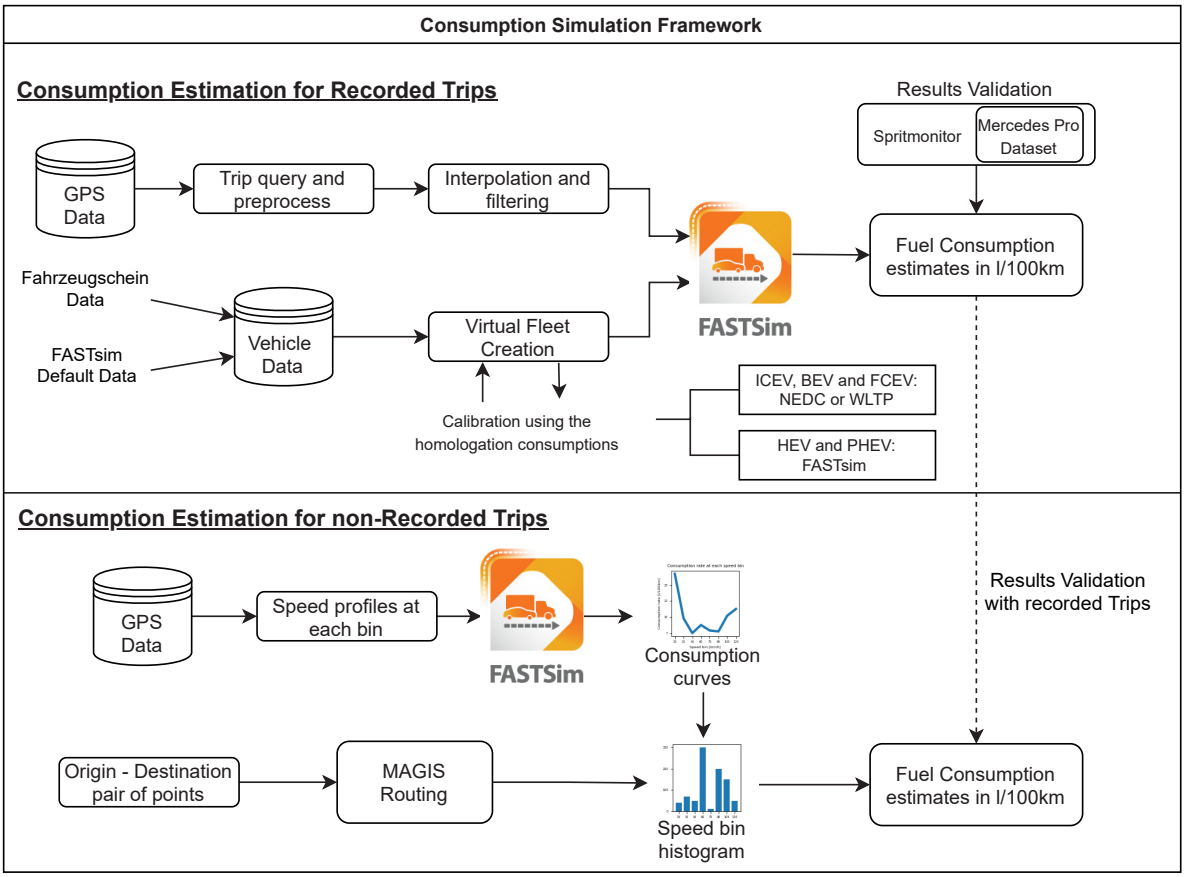


Figure 3.1: Diagram of the complete simulation framework.

3.1 Consumption Simulation for Recorded Trips

The consumption simulation for recorded trips covers 5 fundamental aspects. In the first place, the choice of a consumption model, secondly, the identification and creation of virtual fleets, third, the data query for generating trips, fourth the trip filtering and interpolation of speed profiles and finally, the validation of the results.

Simulation model
The decision on which simulation model to use is discussed in the following paragraph, based on the literature review of Section 2.1 and 2.3. The first step correspond to the simulation scale. In this regard, it is clear that only a Microscopic model can fulfill the requirements and objectives of this work. This is because we need to simulate individual vehicles instead of averaged fleets, that also follow specific trips instead of driving in an average network. Thus, COPERT, EMFAC and HBEFA can not be included in the thesis. Further reasons for not considering these are, that COPERT only uses constant speed bins, so the advantages of having GPS trip data would be lost, that using EMFAC outside California is not cost-effective and finally, the fact that HBEFA is licensed, which also makes it unsuitable for the work. Regarding MOVES, which as mentioned in Subsection 2.3.1, it can be used at any scale, so in principle it could be a suitable model for our work. However, it has two problems, on the first hand, despite it ability to simulate individual vehicles, MOVES only provides averaged vehicle types. This means, that it can differentiate between hatchbacks and sport cars, or Gasoline and a Diesel vehicles, but not between specific models. Therefore by using this model the overall maximum achievable accuracy would diminish. On the other hand, MOVES is said to be relatively difficult to use, and probably impossible to customize without having to code the software from scratch, since its distribution is not open code. This is in fact the reason why MOVESTAR was created. For all these reasons, MOVES is also not considered.

The obvious alternative, which is in fact the standard approach used in the literature, is to consider a Microscopic model. The first option would be ADVISOR, since it is able to deal with all existing powertrain technologies and its Matlab/Simulink structure makes it fully adaptable. Its disadvantage, however, is its lack of actualized data. Using ADVISOR to simulate present vehicles would require and outstanding calibration effort. A similar situation arises with CMEM, since the majority of its required parameters are user-defined, while the model only provides some default parameters. Moreover, CMEM only accounts for ICEV, so a complementary model would be necessary for the other technologies. The BEV counterpart of CMEM could be the University of Cork's EV model, which has a similar structure, including the necessity of many parameters for its calibration.

More promising are the Virgina Tech models, VT-CPFM and VT-CPEM, since they can be entirely calibrated using the homologation data. Its black-box structure makes it also possible, to integrate them in a custom made framework, however they don't provide a model for neither HEV nor FCEV. This leaves only three possibilities, FASTSim, PHEM and MOVESTAR (which was mentioned early as a variant of MOVES, but not ruled out as an option). The three of them can handle all vehicle technologies and provide enough data to perform simulations, either directly in its distributions, or by accessing public available databases. As already mentioned, however, PHEM is under license which means that using it would require a considerable economic effort and the software distribution wouldn't be easy to adapt (if even possible) to specific use cases. MOVESTAR, as said, can not simulate individual vehicles, but only averaged ones, which limits its accuracy. Therefore the resulting model is FASTSim.

Moreover, FASTSim also have further advantages, as it comes with an open code Python distribution as well as an EXCEL file that supports the process of importing new vehicles to the model. Thus, FASTSim fit all the requirements and is the model in which the simulation framework will be based.

Identification of Fleets
In this work, two vehicle fleets will be used to perform the calculations. Fleet 1 consist of 20 vehicles and is a combination of passenger cars and light commercial vehicles. The majority of the vehicles are transporters that are visiting several different costumers every day, this means on average 30 different stop locations every working day. In some cases there can have up to 90 assignments in a single day. In each stop they unload a payload and load back to the vehicle a different but equivalent payload, so the mass of the vehicle can be assumed constant during a daily trip. Their destinations cover most of the state of Bavaria, and have a mean daily distance of 130 km, with some trips up to 400 km long.

The Fleet 2 consist of 15 light commercial vehicles, most of them equipped with a refrigerated trunk. This is because it belongs to a company that deals with groceries. In this case, the assumption of a constant load is less accurate, as they only load goods in their stops. The range of action of is more reduced, with a mean daily distance of 79 km and is centered in the city of Munich. They make an average of 17 stops per day.

A digital twin (also named virtual fleet) of Fleet 1 and 2 is created using the FASTSim build-in vehicle instance. The process includes a combination of default data with the homologation data of the real vehicles, that can be found in the Vehicle Registration Certificate. The vehicles of both fleets are grouped in models and the variants of a specific model, if any, are only distinguished when there is a relevant difference, such as a different powertrain technology or consumption rate. Once the vehicles have been grouped, a so-called virtual vehicle is created to represent each of the groups. The accuracy of the created virtual vehicles is validated using the homologation consumption data. When it differs from the expected value, some parameters are optimized to enhance its improved performance. Since FASTSim is meant for the American market, its validation module is not directly applicable to european vehicles. The original module compares the real homologation consumption data of each vehicle with the results achieved by running the corresponding homologation cycle, however the US and the EU use different drive cycles for this. To circumvent this issue in ICEV, BEV and FCEV, a custom validation is made with the corresponding European drive cycles, namely the New European Driving cycle (NEDC) and the Worldwide harmonized Light vehicles Test Cycle (WLTC). This circumvention is not possible with the hybrid vehicles, as the derivation of its official consumption values is more complex. Thus, both HEV and PHEV are calibrated and optimized using the fastsim built-in module and therefore require the US homologation consumption figures.

GPS data query for trip generation
Once the vehicle instance has been created, the next step is to generate the GPS data-based trips. FASTSim uses speed profiles at a 1 Hz frequency for its simulations. Additionally, the user can also specify a second-wise road grade in terms of altitude gain divided by the longitudinal distance traveled, and even a road charging infrastructure, a feature which is not used. In this project, the trips are created using real-world GPS recordings of fleets, which allow to create precise and realistic speed profiles. The raw data is gathered using a custom device capable of registering vehicle tracks at a 10 Hz frequency. The design and manipulation of the tracking device, the fleet monitoring and the database maintenance are not part of the project, so they can be assumed to be created beforehand. A comprehensive description of this process, applied to previous fleet monitoring projects is provided by Wittmann [52].

in short, the first step for the creation of trips is querying such raw data from the database and creating a continuous speed profile at the desired frequency rate. Then, the road grade throughout the trip is calculated and added to the speed profile as time series.

Trip filtering and interpolation

Since we are working with experimental data, some inaccuracies should be expected, mainly because of low signal quality in urban areas or even spatial-temporal gaps due to missing data, for example in tunnels. Moreover, the created speed profile may still contain idle noise, which consists of very low intermittent speeds for a relatively long period. The noise originates in traffic lights and other stops as a consequence of the GPS inaccuracies. The trip filtering and interpolation process addresses such issues.

Focusing on the data gaps, a common approach for addressing them is using dead reckoning. Dead reckoning is a positioning technique based on the integration of an estimated or measured displacement vector. It generally uses multiple sensors such as gyroscopes, odometers and accelerometers. A perfect integration requires a perfect knowledge of the velocity at all times, which in practice means that dead-reckoning systems are subject to accumulated positioning errors that increase with time [53]. Thus, its application is best suited to short periods. In Zhao [54] they presented an extended Kalman filter algorithm able to provide improved vehicle speed information, that enables continuous positioning even in urban areas and tunnels. Their results showed that the filter was effective most of the time when suffering a GPS signal mask.

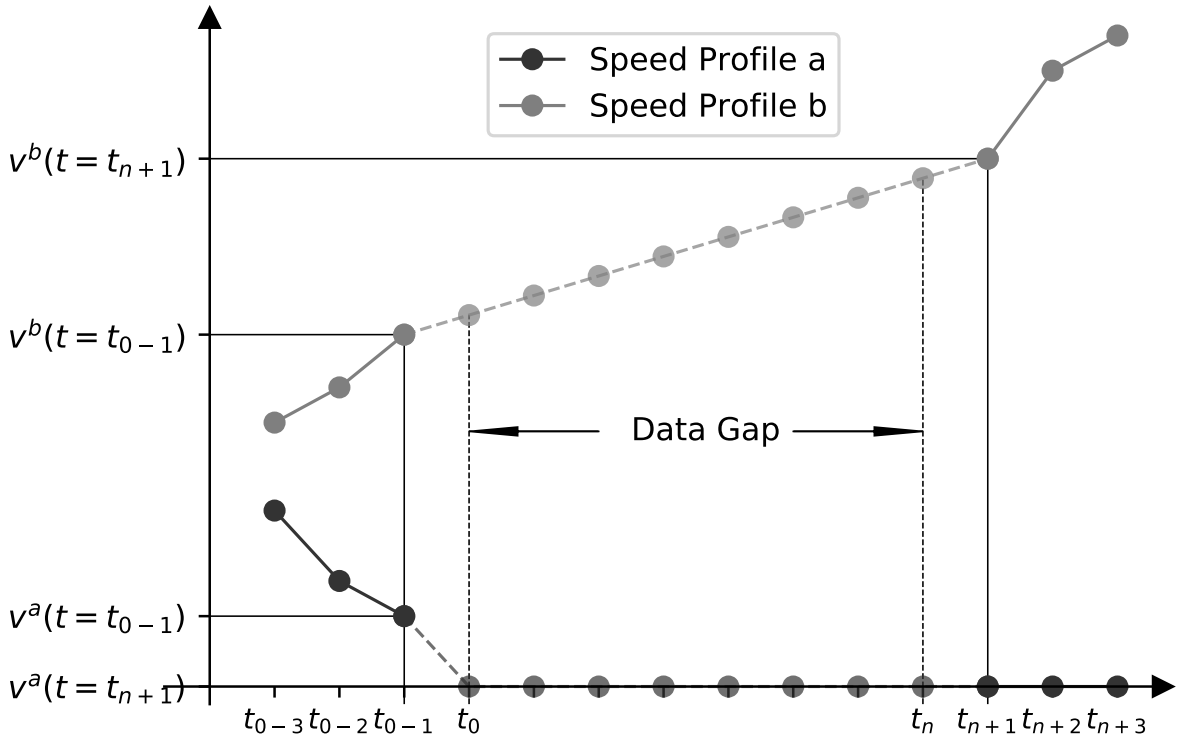


Figure 3.2: Two sample speed profiles showing the two possible interpolation behaviors.

Given the difficulty of building and implementing a Kalman filter, and provided that we have the speed values just before and after the data gap, in this work, the dead reckoning problem is solved by simple linear interpolation. Figure 3.2 shows and example. For each speed profile, an algorithm identifies the missing data gaps, these are sequences in which there is no speed data available. Then it determines whether the instant before and after the gaps have a positive speed. If both instants do, the algorithm will interpolate the speed between those points and the gap

will be filled with the resulting speed values, as in the speed profile b. If on the contrary, one (or both) of the pre and post-gap instants have zero speed, the algorithm assumes that the vehicle was not moving for the entire data gap and thus, will be filled with 0s. This is the situation in the speed profile a. Once this process has been done for all the existing gaps, the noisy idle speeds are set to 0 and the excessively long idle times are removed. The speed profile resulting from this process will be ready to use for consumption simulations.

Validation of the results
The last step in the consumption simulation is to perform batch calculations of the complete fleet and validate the results. For doing so, two different strategies are developed. One strategy consists of calculating the average consumption rate of each vehicle during a large period of time and comparing the results with the publicly available data of Spritmonitor [55]. This website provides crowd-sourced consumption values from nearly a million vehicles adding up to 18 billion km. The second approach is only limited to a subset of vehicles of Fleet 1 (which is presented in Section 5.1), monitored by the Mercedes Pro platform. This platform provides the real consumption rates of each of the monitored vehicles on a daily basis. These daily rates are compared with the rates achieved using the simulation framework, so validation of the consumption rates is made using real-world data.

3.2 Consumption Simulation for non-Recorded Trips

In this section, the methodology to create virtual trips and simulate their consumption is presented. First, the methodologies for creating synthetic speed profiles, presented in Section 2.4, are discussed. Secondly, we explain the tools used for creating trips with a routing API. Then the consumption curve calculation procedure is explained and finally the approach for estimating individual trip consumption and assessing the results accuracy is introduced.

Creation of virtual trips
From the literature review it becomes obvious that despite the interest in the topic, there is still a big margin for improvements in this field. The mathematical-based approaches used a traffic simulation model such as VISSIM or ADAS-RT to create the speed profiles. Its implementation, however, it's not straightforward and requires considerable effort. This is because such models need a comprehensive characterization of the road network in order to predict the traffic flows, which means that can only be used in a limited area. To implement this approach in the present work would require characterization of at least half of the Free state of Bavaria, which would be, to the best of our knowledge, almost impossible. Moreover, the mentioned software is licensed, limiting, even more, its applicability for the present project.

The alternative is using a data-driven approach, which is very well suited to the present work given the vast amount of real drive data that we have. In this regard, the three use cases presented follow a similar approach: they employ a large database containing thousands of speed profiles to derive an average consumption per road link or average speed bin. Then, one can calculate the consumption rate of a trip by concatenating the corresponding road links, or as the sum of the averaged consumption rates. This methodology assumes that the traffic, weather and driver related factors can be neglected and dilutes the importance of accelerations. Therefore, the validation errors are around 10 % to 20 %. To enhance its accuracy, authors include some extra features such as the type of turn (right, left, straight), road type or road grade.

Leaving aside the particularities of each approach, they all have in common the derivation of a consumption rate curve as a function of the link speed, with the possibility to add the already mentioned additional features. Thus, the approach of this project is founded in this same principle, namely the derivation of consumption rate vs. link speed curves for estimating the fuel consumption of trips before they are driven, as in [42, 45]. These curves will be defined in speed bins, this is, ranges of 10, 15 or $20km/h$.

Routing
The virtual Trips are created out of an origin and destination pair of coordinate points. Before going into detail on the consumption estimation, a fundamental task is to determine the route between those points. For doing so, we employ a Geographic Information System (GIS) developed in the TUM's Institute of Automotive Engineering called MAGIS [56]. This GIS provides an easy-to-use API with a specific service for routing purposes, so creating the routes is as straightforward as simply passing the origin-destination pairs to the API and obtaining the routes between those points.

The complete approach for estimating fuel consumption requires from independent variable produced by the routing engine that is representative of the possible consumption rate. This variable is the link average speed, this is, the average speed of the vehicles driving in a specific road link and is in fact, the information that the routing algorithm uses to determine the best route between two points. Therefore, the above-mentioned speed links will be based on this speed.

Consumption curve
In order to create the consumption curve, a large amount of speed profiles is gathered and assigned to its correspondent speed bin. This way, each bin contains an extensive speed profile coming from several individual trips and thus representative of the average. The consumption curve will be created by simulating each of those large speed profiles using FASTSim and will have a consumption rate in l/100km as an output. An advantage of this approach is that allows for creating vehicle-specific curves, more precisely, as many curves as virtual vehicles in FASTSim.

Consumption estimation and validation
Once the consumption curve has been derived, the fuel estimation is straightforward. First, a route is created using an origin-destination pair of points and the MAGIS routing algorithm. Secondly, a cumulative distribution of the distance traveled in each speed bin is calculated. Finally, the fuel consumption of the trip is derived by multiplying the cumulative distribution with the corresponding value of the consumption curve.

The validation procedure is then relatively simple and consists of comparing the rates achieved with the GPS-derived speed profiles, with the results of the virtual trip approach when following the same route. The tracks used for the validation are different from the ones used for creating the consumption curve. Also, the data quality of each tracks will considered, so only a representative subset of data with the highest possible accuracy is used.

4 Implementation

In this chapter the implementation of the methodology is described. The first section explains the process of implementing a virtual fleet with a detailed step-by-step description. The second section describes the data querying and filtering process including the implemented algorithms and the third section examines how the methodology is implemented for non-recorded trips consumption estimation.

4.1 Virtual fleet implementation

The EXCEL distribution of FASTSim includes a vehicle import wizard that in principle could speed up the process of creating virtual vehicles. However, it is incomplete and meant for the American market, therefore, it is not always applicable. This import tool allows to easily create virtual light-duty vehicles by selecting specific models from a list and importing their parameters from a database. Most of the data is also found in the USEPA internet page which is periodically updated [57], so the database can be regularly updated by downloading the actualized database, filtering it using a script and adding the resulting data to the EXCEL file. The main problem of this approach is that at each actualization the spreadsheet increments notably its size decreasing drastically its time performance. It is thus an inefficient way to import a limited fleet of vehicles. Moreover, the scripting efforts to filter the actualized data, which usually contains changes from one actualization to the other, do not pay off when compared to the gains of using the import wizard.

Therefore, the import of new vehicle parameters is done manually using the also available manual import tool. Figure 4.1 shows in the left hand the manual import tool with its 9 parameters categories, while in the right hand, there is the button to access the import wizard for an automatic data import. As seen, the vehicle parameters are distributed in 9 different categories, some of them may be left blank (such as Motor or Traction Battery) depending on the vehicle type. At each category, several parameters are always left with the default values since they are model-dependent, while others are to be filled with the vehicle information. Neither the Energy management nor the Miscellaneous parameters are modified.

The parameters of the vehicle are filled with the Vehicle Registration Certificate ("Zulassungsbescheinigung" in german) data, which is available to us for each of the real vehicles on both fleets. The official template of the Vehicle Registration Certificate is found in Figure 4.2, while Table 4.1 shows the relevant parameters to import from it. Note that the specifications of hybrid vehicles, such as the electric motor power or battery capacity are detailed in row 22 of the certificate.

Table 4.2 shows the specific parameters to import to FASTSim. The vehicle type defines the model framework and architecture of the powertrain. It can be a conventional architecture (for Diesel

Figure 4.1: The vehicle manual import tool from Fastsim.

Figure 4.2: The official template of the Vehicle Registration certificate in Germany [58].

Table 4.1: Relevant parameters in the Vehicle Registration Certificate.

Field name	Code	Unit
Emissions class and Homologation type	14	-
Tyres	15.1	-
Width	19	mm
Height	20	mm
Mass of the vehicle	G	kg
Type of fuel or power source	P.3	-
Rated Power	P-4	kW
CO_2	V.7	gCO_2/km

or Gasoline ICEV), a HEV architecture, a PHEV architecture or a BEV architecture. The term fuel converter is broadly used in the vehicle simulation field and refers to the type of engine. In this case, the user can choose between five options that will define the efficiency map of the vehicle, which are a SI engine (Gasoline), an Atkison SI engine (Gasoline), a CI engine (Diesel), a heavy-duty CI engine (Diesel) and a Fuel Cell. The engine power is introduced in kW an can be left blank in BEV. The vehicle mass is introduced in American pounds (lbs) to the excel file and transformed to the glider mass, and inner parameter of the model. The wheel radius is calculated by using the Tyre information, while the wheel rolling coefficient is normally left as a default value as it is not an easily available parameter. Finally, the Frontal area of the vehicle is derived from the vehicle's width and height using Equation (4.1). This calculation follows the SAE J1263 proposed in FASTSim, where A is the frontal area in m^2, W and H are respectively the vehicle's width and height found in the vehicle certificate and 0.8 is the profile factor.

$$A = 0.8WH \tag{4.1}$$

Since the vehicle registration certificate provides the consumption rate in gCO_2/km, it must be converted back to $l/100\,\text{km}$. This is done using the conversion of Equation (4.2) for petrol vehicles and Equation (4.3) for the diesel vehicles [59].

$$l/100\,\text{km} = 0.043552 gCO_2/\text{km} \tag{4.2}$$

$$l/100\,\text{km} = 0.038388 gCO_2/\text{km} \tag{4.3}$$

Finally, two additional parameters are specified for electric vehicles, the electric motor power and the battery capacity, both of them are also available in the registration certificate.

Table 4.2: The main parameters to import to FASTSim.

Parameter	Description
Vehicle Type	Conventional, HEV, PHEV, BEV
Fuel Converter Type	SI engine, Atkison, SI engine, Diesel engine, Fuel Cell or Heavy-duty Diesel engine
Engine Power	in kW, 0 if BEV
Vehicle Mass	in lbs
Frontal Area	Using Equation (4.1)
Wheel Radius	Using Tyre information
Drag coefficient	Real value if available, if not, Default
Rolling resistance R_f	Real value if available, if not, Default
Electric Motor power	in kW
Battery storage Capacity	in kWh

The FASTSim documentation states that the model is able to simulate most vehicles within a 5 % of error and almost every vehicle within a 10 %. However, this only applies when using the default data. To further improve the performance of the model, calibration is made to fine-tune some of the default parameters to minimize the difference between the actual homologation value and the achieved one. The parameters that may change during the calibration are the drag coefficient, the rolling resistance and the engine efficiency improvement. This last parameter accounts for a possible additional efficiency improvement, derived from an overall better construction of the engine or enhanced engine performance. Table 4.3 shows the parameters to optimize and its corresponding range of possible values.

Table 4.3: The parameters that can be fine-tuned during the calibration.

Parameter	Range	Description
Drag coefficient	(0.27 - 0.4)	When no specific data is provided
Rolling resistance R_f	(0.0068 - 0.012)	When no specific data is provided
Engine efficiency improvement	(0 % - 2 %)	Accounts for additional efficiency improvements of a vehicle

4.2 Data query and filtering

As seen in the methodology, the creation of trips out of recorded GPS-data consists in two steps. The first one correspond to the trip query and pre process, that includes the calculation of the road grade. The second step is the trip filtering and interpolation.

Trip query and pre-process

This part is completely done at the database using SQL programming and deals with the resampling of the signal and the addition of road grade information. For each trip, both the position data and the speed data are queried as a time series at the raw acquisition frequency of 10 Hz. The position data is used to calculate the road grade by first obtaining the altitude profile from a cartographic information dataset and then dividing it by the traveled distance between each pair of points. For doing so, the time series is down-sampled to 33.3 mHz. The reason for such a long time step is to avoid unreal fluctuations and peaks of the road grade that would be originated with higher frequency data. Moreover, since the raw data was not map matched, it would not be possible to keep constant length segments for the road grade.

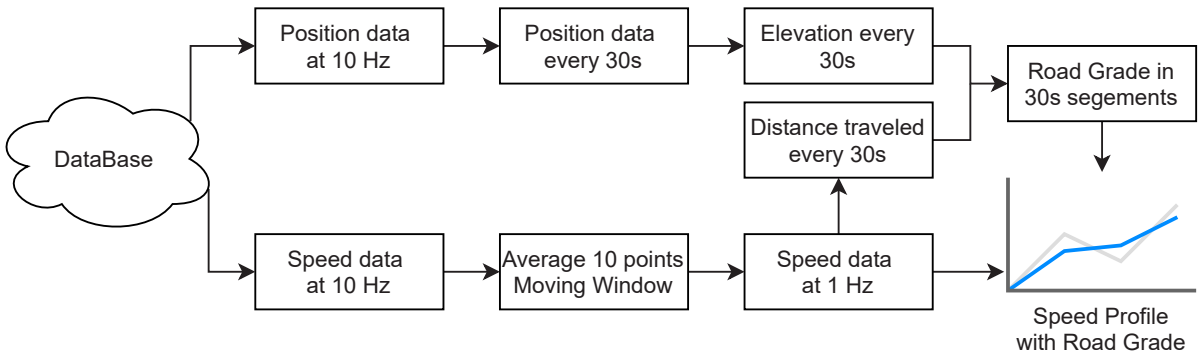


Figure 4.3: Diagram of the data processing.

Many different alternatives were tried for deriving an accurate road grade profile throughout the trip. Using the GPS-derived altitude instead of the digital map alternative, was not considered

from the very beginning due to the well-known inconsistent altitude data when using a GPS tracking device. Despite the tracker provided good results most of the time, there were occasions in which erroneous altitudes were registered. In those cases the results were so inconsistent that a filter would be useless, therefore the alternative of deriving the altitude from a digital map was chosen. The problem with this alternative was the lack of map matching, so the only possible matching between the road grade profile and the speed profile was the time stamp. Using distance instead would create a spatial shift between both time series, which would be worsened at each speed gap.

The challenge then lied on fine-tuning the sampling frequency of the altitude profile, which is required for creating the road grade profile. As opposed to the speed profile generation, in which a high sampling frequency is a necessary condition for a realistic result, the disadvantages of using high frequency rate when calculating the road grade outweigh its advantages. This is due to the fact that the road grade is derived as the gradient of the altitude throughout the distance. Therefore, when driving at low speeds, a small altitude gain may produce unrealistic peaks. When filtering those peaks, the best one can obtain is a smothering of the signal, but since the peak was an artifact, the signal will still be flawed. The solution for this is to ensure that the longitudinal distance is sufficiently large so these peaks are never produced. The signal losses resolution by doing so, but the impact of this loss in the overall fuel consumption, which after all is the desired output, is negligible. On the other hand, keeping the artificial peaks could indeed impact the overall fuel consumption in short trips and even create a malfunction in the simulation.

As said, the implemented frequency of the altitude signal is 33.3 mHz, which translates in an altitude time step of 30 s. Additionally, when calculating the road grade there is a minimum distance threshold of 100 m that ensures that the longitudinal distance is indeed long enough to avoid the above-mentioned peaks. In the unlikely event of a trip segment not fulfilling this condition, its road grade is set as null. Note that the distance condition must be proportional to the time step in order to control the amount of out-filtered segments. Otherwise, most of the trip would have a null road grade due to insufficient longitudinal distance. The selected distance threshold implies an average speed of 12 km/h or more. In preliminary implementations both shorter and larger time steps were tried. As a quality assessment, the difference between the real altitude profile (from the digital map) and the altitude profile derived from the created road grade time series is compared.

In parallel to the road grade profile calculation, the speed data is averaged using a 10 point moving window to create a precise speed profile at 1 Hz. It is with this same speed data that the elapsed distance is calculated. Finally, both time series, namely the speed profile and the road grade are joined using its common time attribute and thus creating the pre-filtered trip. Since the road grade is defined in 30 s segments, when joined to the speed data is kept constant within each segment. The complete process is described in Figure 4.3.

Alternative implementations considered other speed profile import methodologies rather than the use of a 10 point moving average window. The most simple one consisted of importing the original raw data as is, so at 10 Hz, and letting a FASTSim built-in function resample the speed profile at the desired frequency. In a different implementation, instead of using a 10 point moving window to average and resample the speed, the time stamp of the speed profile was truncated to seconds and all the speeds falling in each of the resulting timestamps were averaged to create the second-by-second speed profile.

Interpolation and filtering

The second step of the trip generation corresponds to the interpolation and filtering. As described in Section 3.1, the raw data comes with some irregularities and signal losses that the querying process does not correct. The selected approach for treating the trip data consists of creating an algorithm that finds the signal losses and either interpolates the speed when both at the beginning and end of the gap the vehicle was moving, or leaves the speed at 0 when not. Once all the gaps have been filled, the algorithm corrects the idle noise data and cuts the idle times longer than 100 s. The algorithm is shown in Figure 4.4.

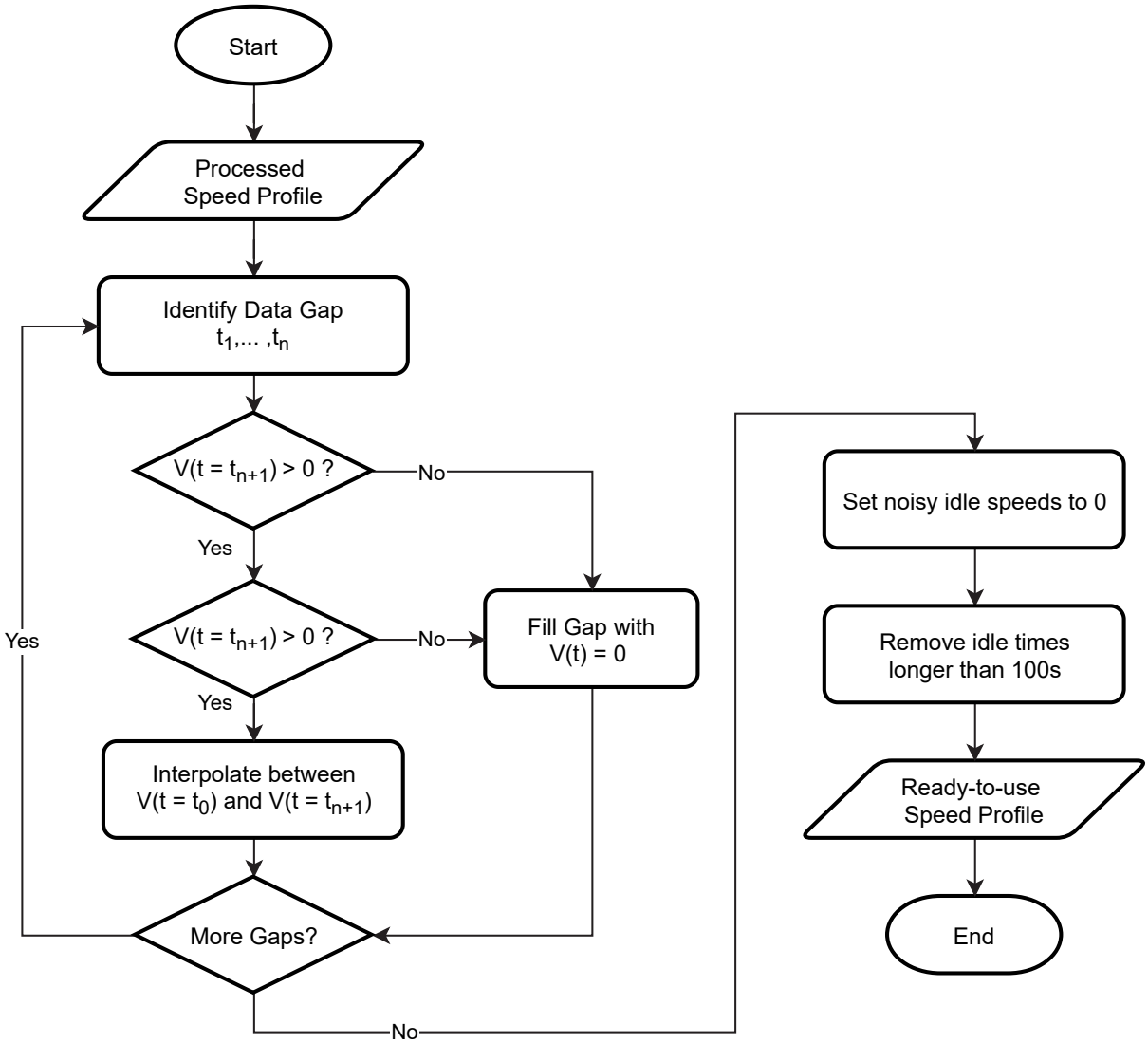


Figure 4.4: Algorithm used for filtering the trip data.

This algorithm is the result of a series of trials in which many alternatives were considered. It should be kept in mind that, since the algorithm is trying to compensate the data losses, there is not an unambiguous way to determine whether one option is effectively filling the gaps correctly, as the necessary data for doing so is precisely the one that is missing. There is, a way to indirectly determine whether the resulting speed profile is likely to be correct by using the trip distance. The trip distance can be calculated by creating a line out of all spatial points and measuring its distance, or by integrating the speed profile. In the trips without any relevant data gap, this two distances are equal, while in the incomplete trips the distance originated from integrating

the speed profile fall shorter. With this information, different interpolation strategies can be compared.

The interpolation between two instants in which the vehicle is moving was relatively straightforward once discarded the laborious and complex option of creating a Kalman filter, however, what to do when the vehicle was idle at one of the boundaries of the gap was uncertain. Simply interpolating would only work in extremely short gaps and would create an unrealistic ramp in the longer ones. It was unclear whether just leaving the speed as a null would create inconsistencies due to substantial magnitude leaps. For this reason, several implementations tried to create short artificial ramps at the end of the gap that would account for the acceleration. At this point, it is worth mentioning that FASTSim only follows the speed profile trace as long as the simulated vehicle has enough power for doing so, otherwise, it calculates the maximal speed it can reach with the available power. The idea of the ramps is to avoid the unlikely accelerations at full power that simulate aggressive driving hardly seen on the streets.

4.3 Non-recorded trips implementation

The simulation of non-recorded trips has two different phases. The first phase can be named as the training phase analogously to the machine learning terminology and consist of creating the consumption curve using the real drive recorded data from the database as the speed profile input to FASTSim. The second phase is the actual consumption estimation of non-recorded trips, which includes the validation of the results. Additionally, at the end of the section, there is a description of a radically different approach for simulating the fuel consumption of non-recorded trips, which has also been tried.

Training phase
The challenge of routing and track identification comes from the trip preparation, which is necessary for deriving the consumption rate curves. Particularly, we need to create a speed profile at each of the mentioned speed bins using an extensive set of tracks. Here, by tracks it is meant segments of the speed profile of a certain trip that has a stable attribute, in this case, the link speed. As said, this attribute does not come from the GPS recorded data, but from the routing algorithm. Thus, the tracks originate from matching the recorded GPS trips to the routed trips, this is, the route created by MAGIS that have the exact origin and destination points that the GPS-derived trip. Since there are multiple possible routes for each origin and destination pair, we must ensure that we are dealing with the same one. Moreover, the GPS derived speed profile has to be matched with the corresponding road links to be certain that the created tracks correspond indeed to the desired speed bin. This is done by transforming the speed profile from a speed vs time to a speed vs distance perspective.

Table 4.4: The speed bins used for creating the consumption curve.

	Bin 1	**Bin 2**	**Bin 3**	**Bin 4**	**Bin 5**	**Bin 6**	**Bin 7**	**Bin 8**
Speeds in km/h	0 to 15	15 to 30	30 to 45	45 to 60	60 to 75	75 to 90	90 to 105	>105

To ensure that the recorded trip matches the routed trip and therefore its tracks do indeed represent an existing road link with a particular link speed, only those trips in which the routed and GPS distance match over a 99.9 % are used. This means that only a small subset of the complete database can be used as training data, a limitation that comes from the lack of map-

matched trips. For this reason, the curves are created using the trip data from all of the vehicles from Fleet 1 and 2, so there is sufficient data available.

All the recorded tracks that are long enough in terms of distance and time are assigned to their corresponding speed bin, which are shown in Table 4.4. Then the tracks at each bin are concatenated creating a continuous speed profile and passed to FASTSim. At this point, any of the vehicles of the virtual trip can be selected as the modeled vehicle and additionally a payload can be defined to simulate loaded transporters. Finally, the consumption curve is completed by repeating this process at each of the eight speed bins. It is important to note that the consumption curves of a HEV must be created assuming a SOC of 40 %. This way, FASTSim recognizes that the vehicles is driving in CS mode, which is the most accurate assumption for this typology of vehicle in which the batteries can not be externally recharged when the vehicle is stopped.

There were several alternative implementations of this part covering the different aspects of the training phase. Apart from using different quality thresholds to optimize the trade-off between the amount of data and the quality of the data, also different subsets of vehicles were tried with the objective of emulating different driving styles. The subsets tested include all the transporters, which pretended to infer a light commercial vehicle driver, all the Passenger Cars, that pretended to infer an average auto driver and all the Sprinters, in which the particularities of driving a Sprinter were isolated. The final implementation, did not account for this differentiation, however the author would like to encourage future works in which different driving behaviors are captured by the consumption curves.

An additional parameter to optimize was the speed bin configuration. The choice of the bin length, finally set at 15 km/h is not obvious and has many implications. The first thing to keep in mind is the first bin width. The number of tracks falling in the first bin is very limited due to the fact that there are very few available tracks at such a low speed and the existing tracks are generally very short. On the other hand, the second speed bin is much less scarce and captures a driving situation that may have a heavy impact on the overall fuel consumption of a trip. Moreover, despite the vast amount of data available, its distribution may not be uniform if the speed bins are to narrow and setting a short bin length might lead to overfitting.

Table 4.5: Configuration A with speed bins using powers of 10, speeds in km/h.

Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9	Bin 10	Bin 11	Bin 12
<20	<30	<40	<50	<60	<70	<80	<90	<100	<110	<120	>120

Three additional speed bin configurations are worth mentioning as plausible alternatives. The first, named configuration A, has a narrower speed bin length of $10km/h$, aiming to capture a more extensive range of speeds and trying to be proportional to the speed limits, that are powers of ten. In order to have enough data in the first speed bin, it has a double size. The bins are show in Table 4.5.

Table 4.6: Configuration B with speed bins using powers of 10, speeds in km/h.

Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9	Bin 10	Bin 11	Bin 12
<15	<25	<35	<45	<55	<65	<75	<85	<95	<105	<115	>115

The second variant, named configuration B, has the same bin length of 10 km/h as the former, but instead of having a double-sized first speed bin, it has a length of 15 km/h, the same that the definitive implementation. It is presented in Table 4.6.

Table 4.7: Configuration C with speed bins using powers of 20.

	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7
Speeds in km/h	0 to 20	20 to 40	40 to 60	60 to 80	80 to 100	100 to 120	>120

In contracts, the third additional configuration, named C, is a simplification of the curve in which the speed bins are powers of 20 km/h. It is displayed in Table 4.7.

Simulation and validation

The simulation of non-recorded trips consists of creating a route out of an origin-destination and use it to estimate the fuel consumption utilizing a consumption curve. The route is created by the routing engine of MAGIS via API, which generated a json file as output that includes the length and average speed of each link throughout the trip. With this data, a histogram of instants at each speed bin is built, representing the trip speed distribution. Finally, the number of observations is multiplied by its correspondent speed bin consumption in the consumption curve and the aggregated value is divided by the lengths of the trip in seconds to achieve a consumption rate in l/100km.

The validation of the results is done using a large number of trips, excluding the ones used in the training phase. Particularly, the ones that have at least a 95 % of matching distance. The trips are grouped according to the vehicle they represent, so for each vehicle and trip, the procedure for estimating fuel consumption using both recorded and non-recorded data is performed. The results are then compared using the Normalized Mean Error (NME), Mean Absolute Error (MAE), the Mean Error (ME) and the standard deviation.

Alternative approach for simulating non-recorded trips

Besides the final implementation, the consumption estimation of non-recorded trips was also implemented without the need for the consumption curves. For doing so, the json file containing link speeds and lengths created by the routing engine was used to create a Speed Profile. Then, this speed profile was properly treated and passed to FASTSim to perform simulations as if it was a normally recorded trip. This approach was named the virtual cycle approach, for differentiating it from the consumption curve approach. The created speed profiles, however, looked substantially different two the ones created with real data, as the speed was practically constant throughout the links. Therefore, they didn't capture a realistic driving behavior that well. Still, the results achieved with the virtual cycle approach are also presented in this work.

5 Results

In this chapter, the results of the project are presented. The chapter is divided into 5 sections covering different aspects of the work. The first section presents the two virtual fleets created, including their composition and validation with the corresponding homologation cycle. The second section addresses the GPS-derived trip acquisition, the results of the data processing and the overall data quality. The third section presents the consumption results of the simulations done with the recorded trips, including the Mercedes Pro and the Spritmonitor validation. There is also a comparative analysis of the influence of a payload in the consumption rate of a transporter. The fourth section presents the validation of the consumption results using virtual speed profiles. In the last section, the error chain is presented, this is, a complete display of the existing sources of error in the framework and its measure.

5.1 Virtual fleet

Two virtual fleets are created and validated in this section, Fleet 1 corresponding to a delivery company and Fleet 2, from a grocery distribution company.

Fleet 1
Fleet 1 is the most extensive and is composed of 20 vehicles including 8 different models. The majority of the vehicles are Mercedes Sprinters, some of them are relatively new vehicles, while others are rather old. For this reason, the Sprinters have been split into in two classes according to their consumption. There is 1 Gasoline vehicle, 17 Diesel vehicles, 1 HEV and 1 PHEV, as detailed in Table 5.1.

Table 5.1: The vehicles in Fleet 1.

Vehicle Name	Number of vehicles	Powertrain type	Homologation Cycle	Error
Mercedes-Benz Vito Pro	1	Diesel	WLTC	−0.5 %
Hyundai i30	1	Gasoline	NEDC	−2.6 %
Opel Astra	2	Diesel	WLTC	0.7 %
Volkswagen Caddy	1	Diesel	NEDC	−0.4 %
Toyota Prius HEV	1	Gasoline HEV	FASTSim	0.9 %
Mitsubishi Outlander PHEV	1	Gasoline PHEV	FASTSim	6.3 %
Mercedes Sprinter I	9	Diesel	NEDC	1.8 %
Mercedes Sprinter II	4	Diesel	NEDC	1.2 %

The hybrid vehicles are directly validated by FASTSim while the rest are validated using the corresponding drive cycle by our own process. The validation setting includes additional 136 kg of mass, which stand for the mass of the driver and fuel. This additional mass value is the default value by FASTSim and determines the minimum extra mass of the vehicle that is used in all simulations, leaving open the possibility of a higher value that could account, for instance, for

additional passengers or a payload. The results are shown in Figure 5.1. Thanks to the calibration process, all vehicles excepting the Mitsubishi Outlander have a validation error below 5 % and in total, the weighted average absolute error of the fleet is 1.6 %. This is the average of the absolute value of the error, weighted by the number of vehicles in the fleet.

In fact, the real fleet includes two additional passenger cars that are not displayed on the table and in the statistics, since there was no available tracking data for them and had any relevant role in the design and validation of the framework. However, in the event of a possible trip recording of these vehicles during the elaboration of the present work, a virtual vehicle of them was also created. It involves a Skoda Fabia and a Hyundai Tucson, which achieved an error of −0.20 % and 12.8 % respectively. For documentation purposes, further details of these vehicles are found in Section A.1.

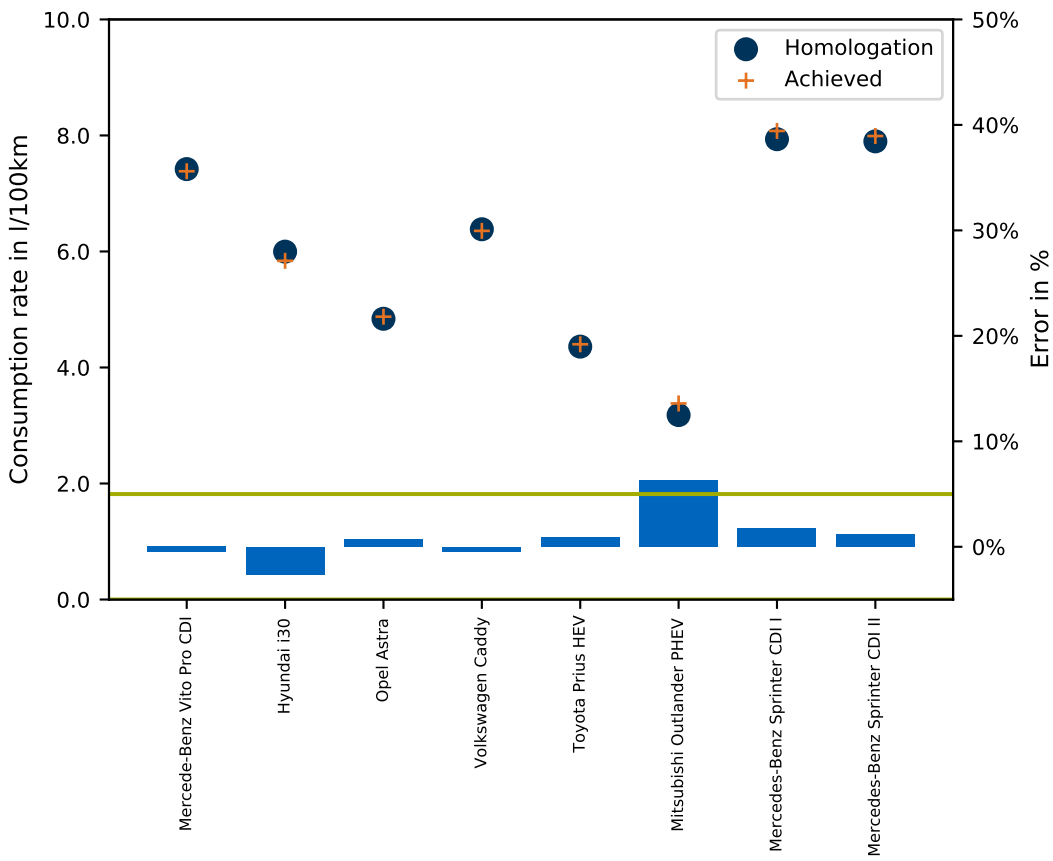


Figure 5.1: The calibration results of Fleet 1.

Fleet 2

Fleet 2 is composed of 13 vehicles, all of them Diesel Transporters, which are divided in 6 different models. They are much older compared to the vehicles of Fleet 1 and are in general much bigger and customized. This hinders both the calibration process and the validation, since the real data is much more scarce and difficult to find. Moreover, as light commercial vehicles, the information found in their vehicle registration certificate is related to the base variant and some key parameters, such of the frontal area, are likely to be inaccurate. Table 5.2 shows the fleet composition.

Table 5.2: The vehicles in Fleet 2.

Vehicle Name	Number of vehicles	Powertrain type	Homologation Cycle	Error
Opel Movano	5	Diesel	NEDC	3.9 %
Iveco 70C15	1	Diesel	-	-
Iveco 50C15	2	Diesel	-	-
Ford Transit	3	Diesel	NEDC	7.7 %
Peugeot Boxer	2	Diesel	NEDC	2.1 %
MAN TGE	1	Diesel	NEDC	4.8 %

The particularity of Fleet 2 is that it has no homologation data for the Iveco vehicles, since it is not mandatory for manufacturers to provide a consumption value for transporters above a certain mass threshold. Therefore they can not be validated with their declared fuel consumption. In this case, the validation shows slightly higher error results, but still all of them are well below the 10 % threshold and only the Ford Transit approaches it with a 7.7 % error. Thus, the weighted average absolute error of Fleet 2, excluding the Iveco vehicles, is 4.7 %. Considering that both Iveco results fall to the upper limit of 10 % error, the weighted average absolute error of Fleet 2 would increase up to a 5.84 %. The results are shown in Figure 5.2.

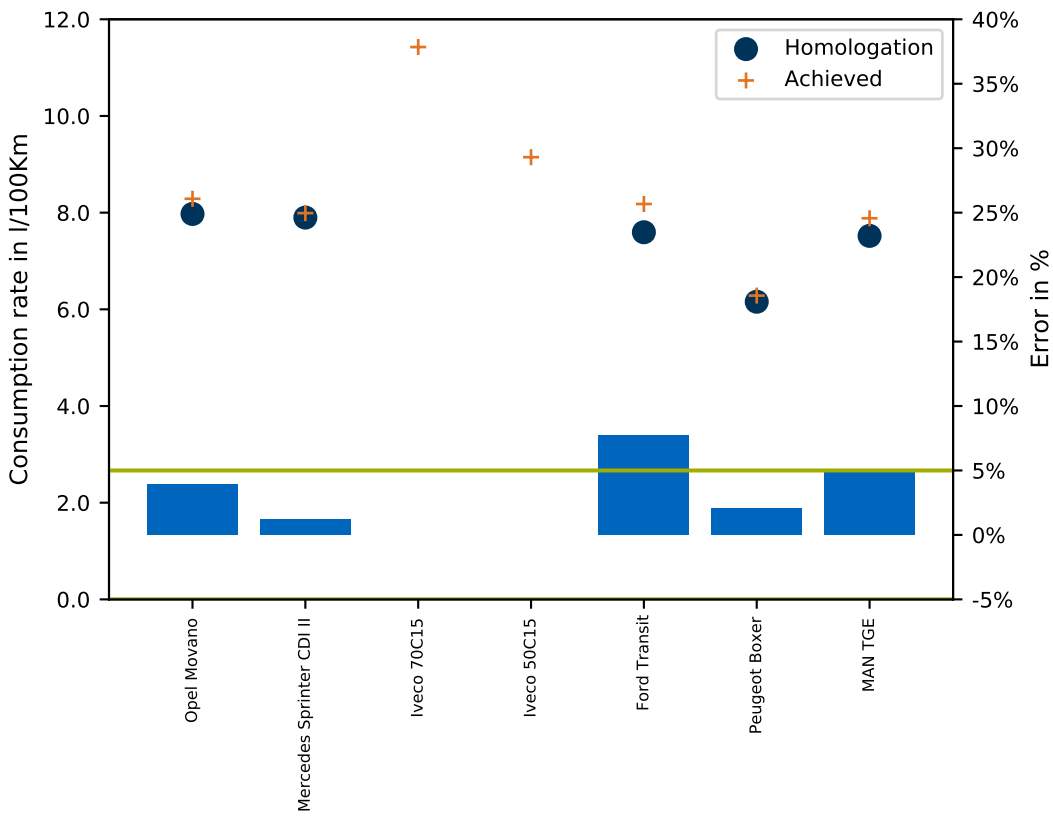


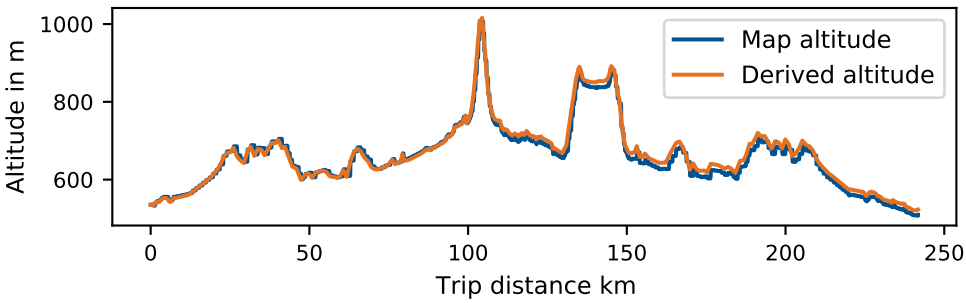
Figure 5.2: The calibration results of Fleet 2.

In this case all the vehicles of the real fleet have been used to some extend in the present project in exception of a Mercedes Sprinter, that did not have real drive data. Since Fleet 1 already have recordings of thousands of kilometers for other similar Sprinters, the lack of data for this particular vehicle in Fleet 2 is unimportant. It is worth mentioning that there is much less real drive data from Fleet 2 compared to Fleet 1. Therefore, there will be much more relevant results from the Fleet 1.

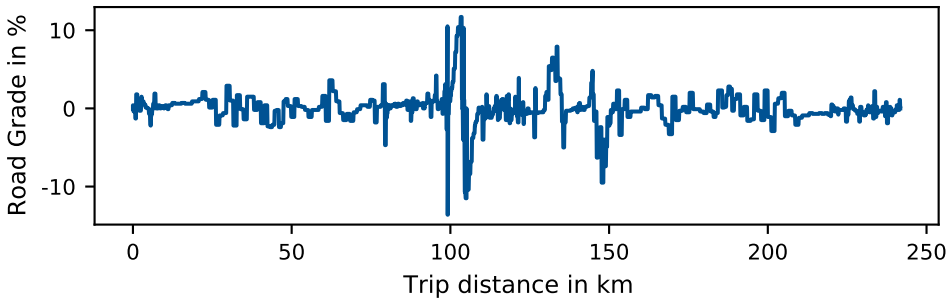
5.2 GPS-Derived Trips

In this section an overview of the creation of the GPS-derived trips is presented. First, we show the data processing in a particular example, with special detail to the road grade calculation. Then, the interpolation procedure for solving low signal issues is presented. Finally, the data quality assessment, which determines whether the created trips are accurate or not, is presented.

Data Processing and road grade calculation
The GPS recorded trips are queried from the database and then treated in two phases. The first phase is the data processing, which includes the derivation of the road grade. The algorithm for doing so is presented in the methodology and displayed in Figure 4.3. In this regard, the most relevant and visible part is the derivation of the road grade, which is done in segments of 30s. In the majority of the trips, the road grade is almost negligible and varies between 0 and 1 %, however, there are also other trips headed to hilly areas such as the Alps. In this case the slope is no longer insignificant and must be taken into account.



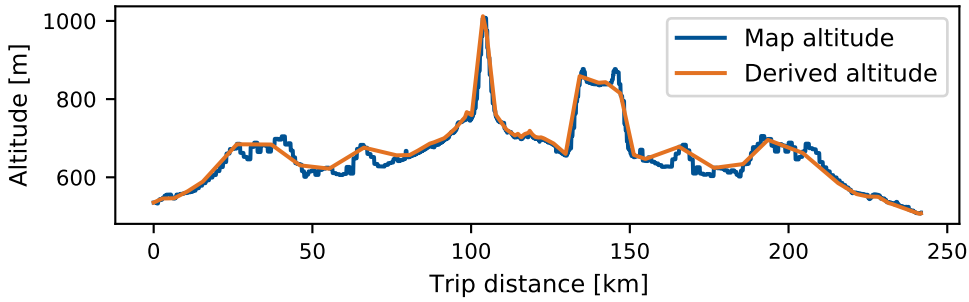
(a) Altitude profile.



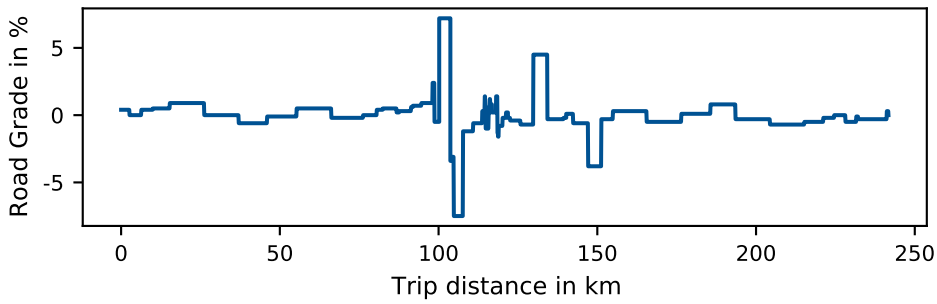
(b) Road grade.

Figure 5.3: Altitude profile and road grade of a round trip to the Alps.

This is the case of the trip shown in Figure 5.3, where the vehicle drove to the german Alps. To ensure that the algorithm is effective in these situations, several similar trips are manually identified and analyzed. In Figure 5.3a the altitude at each point using the official elevation map of Bavaria is compared with the one achieved with the calculated road grade throughout the trip, shown in Figure 5.3b. For comparison purposes both signals start at the same height. The signals are perfectly correlated throughout the trip and after almost 250 km they only diverge 2 %, which is imperceptible for the simulation model. Moreover, the extreme points of the road grade, which have 10 % slope, are checked manually in the map to verify whether they exist in reality or if they are simply noise. The scan shows that they indeed exist and also the peaks found in the trip are happening in the reality.



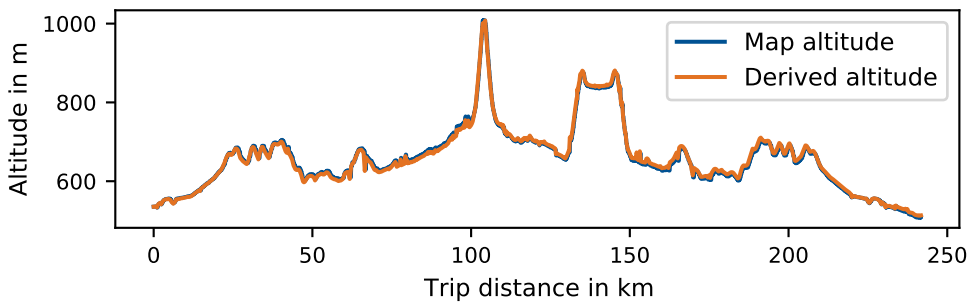
(a) Altitude profile.



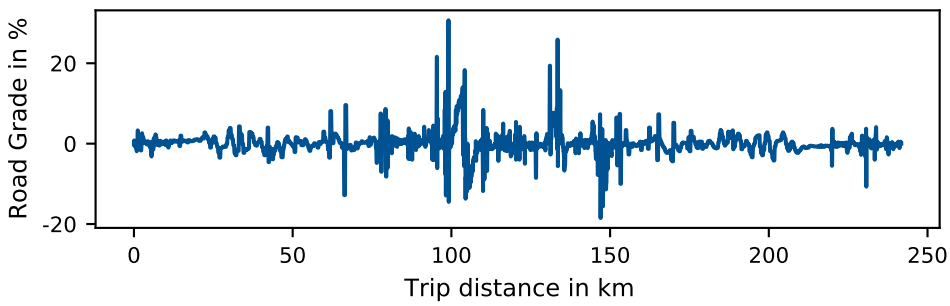
(b) Road grade.

Figure 5.4: Altitude profile and road grade of a round trip to the Alps using a longer segment.

This results are achieved using the definitive implementation, however it is also interesting to see what would it look like when changing some parameters. In Figure 5.4 the same trip is presented with a variation of the parameters. In this case the segments are incremented to 5 minutes, keeping the distance threshold at 100 m.



(a) Altitude profile.



(b) Road grade.

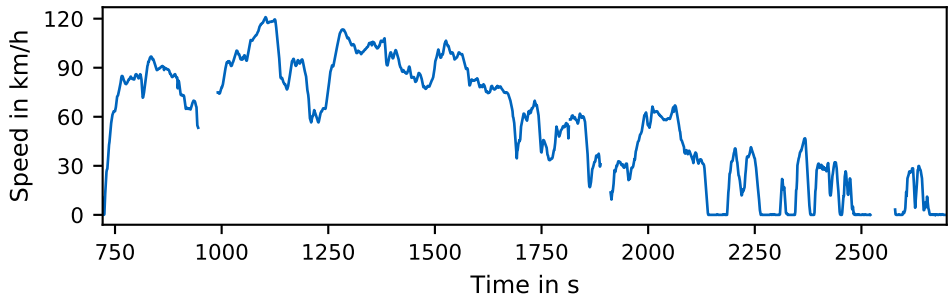
Figure 5.5: Altitude profile and road grade of a round trip to the Alps using a shorter segment.

The difference between both altitudes after 250 km is much narrower and is only a 0.05 %. However, the effect of the longer segments becomes very evident, since the derived altitude in Figure 5.4a is not capturing many of the features and the smaller hills.

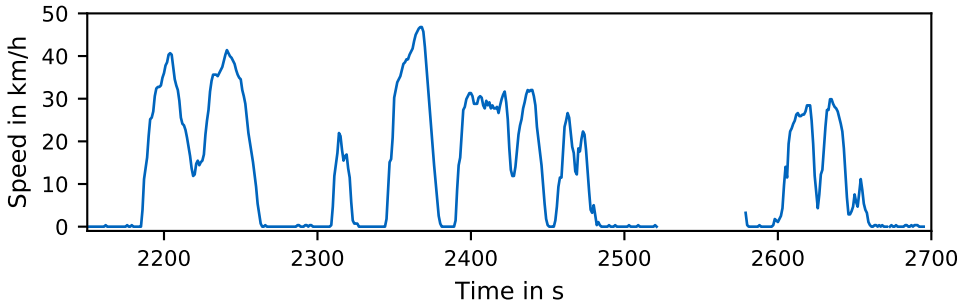
Finally, a possible alternative is to reduce the segment to only 10 s and also reduce the minimum distance accordingly to 30 m. This way the resolution of the altitude profile could be enhanced, but the undesired peaks would still be limited. The resulting altitude profile is found in Figure 5.5. In this case, the altitude traces are perfectly correlated and there is only a 1 % difference at the end of the trip. The problem, however, is in the road grade signal presented in Figure 5.5b, in which there are many peaks over 20 % slope up and downhill, which are not found in the reality. The signal is much noisier and clearly doesn't capture a realistic road grade profile. Additionally, there is a time performance problem when obtaining an altitude profile from a digital map at a high-frequency rate. Using the LRZ server, the final implementation requires 6.8 s, which is already a relatively long time, while the variant with a shorter segment time requires twice times as much. The majority of the computational effort is directly assignable to the altitude profile querying from the digital map.

Trip filtering and interpolation procedure

The processing algorithm is able to create a usable speed profile out from the raw data in the database, however, it doesn't account for possible data inaccuracies and more especially, for signal gaps. Using such trip data without further filtering, would result in inaccurate results that would fail to reproduce the recorded trips.



(a) Full speed profile.



(b) Zoomed-in segment.

Figure 5.6: The sample speed profile.

The procedure for trip filtering and interpolation is presented in a particular trip to show its functionality. In Figure 5.6a the speed profile of a trip without filtering is shown as an example. In the plot there are four data gaps, found around the 1000s, 1800s, 1900s and 2600s timestamp. Moreover, there is abundant idle speed noise, which is best perceived by zooming in the last part of the segment as shown in Figure 5.6b.

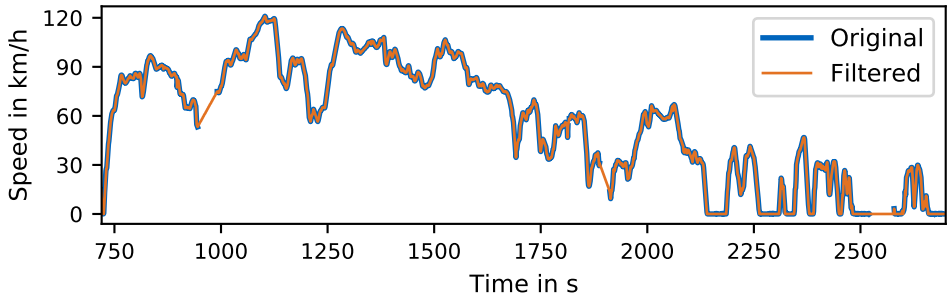
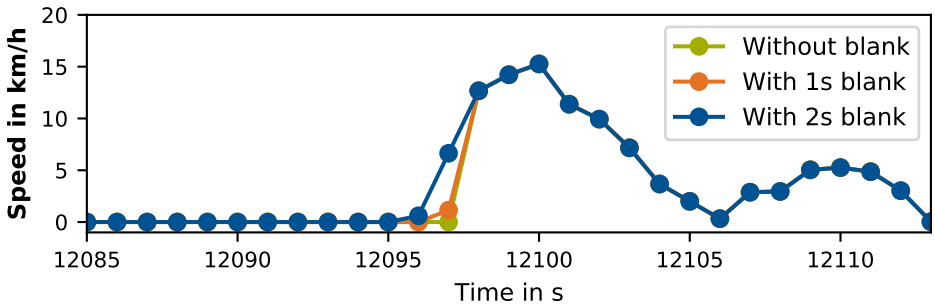
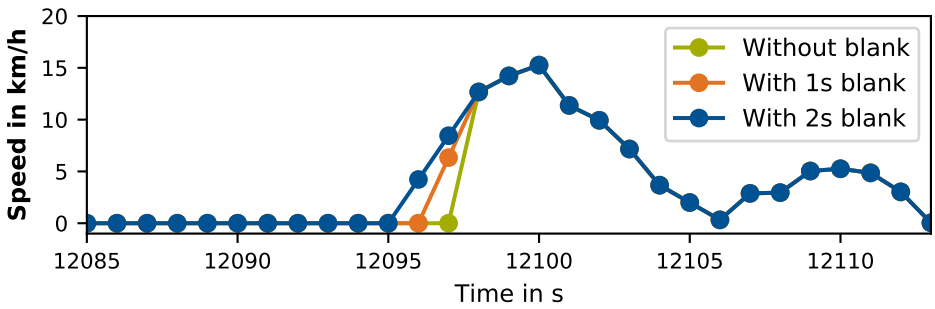


Figure 5.7: Comparison between the original speed profile and the one resulting from the filtering procedure.

The filtering algorithm will interpolate the speed of the three first gaps, as they are between non-zero speeds, however, it will fill the remaining gap with zeros as it starts from idle. This way, we avoid creating unrealistic ramps that could be extremely long (in fact, as long as the length of the data gap). Still, this solution leads obviously to an inaccurate speed profile, since the missing speed data were most like not 0. The need for a unique solution and the difficulty to predict the speed profile in a specific segment leaves no better alternative than this. Finally, the existing idle noise best shown in Figure 5.6b is effectively removed by the algorithm. The comparison between the original speed profile and the filtered (and definitive) one is presented in Figure 5.7.



(a) Using data at 10 Hz.



(b) Using data at 1 Hz.

Figure 5.8: Interpolation starting from idle using different sampling frequencies and leaving zero, one or two instants blank for emulating a realistic acceleration.

Alternative implementations of the interpolation considered smoothing the acceleration curve when starting from idle as a consequence of the algorithm introduced above. For doing so, the algorithm would fill the previous instants with a blank (no data) instead of filling it with zeros.

By doing so, the mentioned unrealistic ramps could be controlled and instead of being as long as the data gap, they could have the desired length. Then, a more realistic acceleration ramp would be created using simple interpolation, thus achieving a smoothed acceleration signal. To get an even more realistic acceleration profile, it was also considered using the original 10 Hz data for the interpolation and resample the signal at the end of the filtering and data treatment process. This process was very dependent of the particular situation, since the length of the acceleration should be a function of the achieved speed. In other words, the number of instants left blank should be greater when accelerating from 0 to 100 km/h, compared to when accelerating from 0 to 20 km/h.

An overview of this variant is presented in Figure 5.8, where Figure 5.8a presents an example of the above-mentioned interpolation directly from the raw frequency and then resampled at 1 Hz and Figure 5.8b the interpolation at the desired final frequency. In both cases, there are three speed profiles, one without any blank, this is the base situation and the actual final implementation. Additionally, there are other two speed profiles in which this interpolation for smoothing accelerations from idle has been implemented, either leaving one second blank, which correspond a one point interpolation, or leaving two blanks, corresponding to a two point interpolation. This results in a straight line when using the data at 1 Hz or in an exponential when using the 10 Hz data. Many more alternatives with longer acceleration ramps were also tried, however, they would not help in making a clear and clean plot and therefore are not displayed.

This alternative implementation had substantial performance issues, especially the one using the 10 Hz raw data, that forces a complete change of the query and trip generation pipeline and undermines the time performance of the process. The difference in the overall consumption on the presented trip were minimal despite the trip contained many gaps in which the interpolation was applied. In particular, there was just a 0.02 % difference in the overall consumption between the two more distinct strategies. Moreover, there is not a consistent way to assess whether the resulting speed profiles were indeed more accurate with the interpolation or not, so this procedure was not implemented in the definitive version.

Data Quality Assessment

As in any project involving experimental data, it is crucial to assess the overall quality of the gathered data in order to correctly assess and understand the results. The obtainment of the GPS-recorded tracks and development of the database structure in which this work is based is out of the scope of the work, so the interested reader should refer to Wittmann [52]. However understanding the reliability of the trips and especially the speed profiles are a fundamental part of the simulation and the evaluation of the results, so this is included in the present work.

The key parameter for defining the data quality of the trip is the trip distance quality, which is the difference between the geometrical distance of the trip, using the spatial coordinates of each recorded point, and the integrated distance, using the speed profile of the trip. For this work, data from 33 different vehicles, adding up to 1646 daily trips with a total length of 191 680 km have been used. The vast majority of the data, a 96 %, have a distance error between −1 % and 10 %, which can also be shown in the histogram at Figure 5.9. These are the daily trips that contain none to relatively few gaps, making the geometrical distance higher than the integrated distance. There are a few outliers with a negative error, which are related to circular trips with huge data gaps or a too optimistic interpolation of the speed profile, which overestimates the traveled distance. There are also a few trips with an error above 10 %.

The recorded trips, however, are not equally distributed among the 14 distinct vehicle types simulated. Nor is the data quality of such trips as it varies substantially from one vehicle to

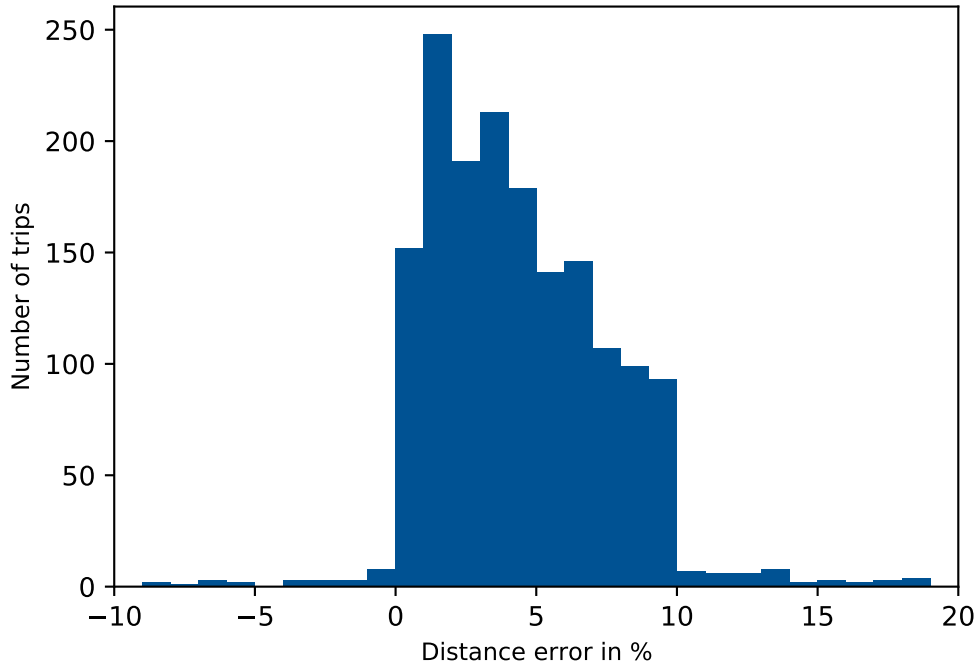


Figure 5.9: Histogram of the data quality of the daily trips. Each column represents a 1 % bin.

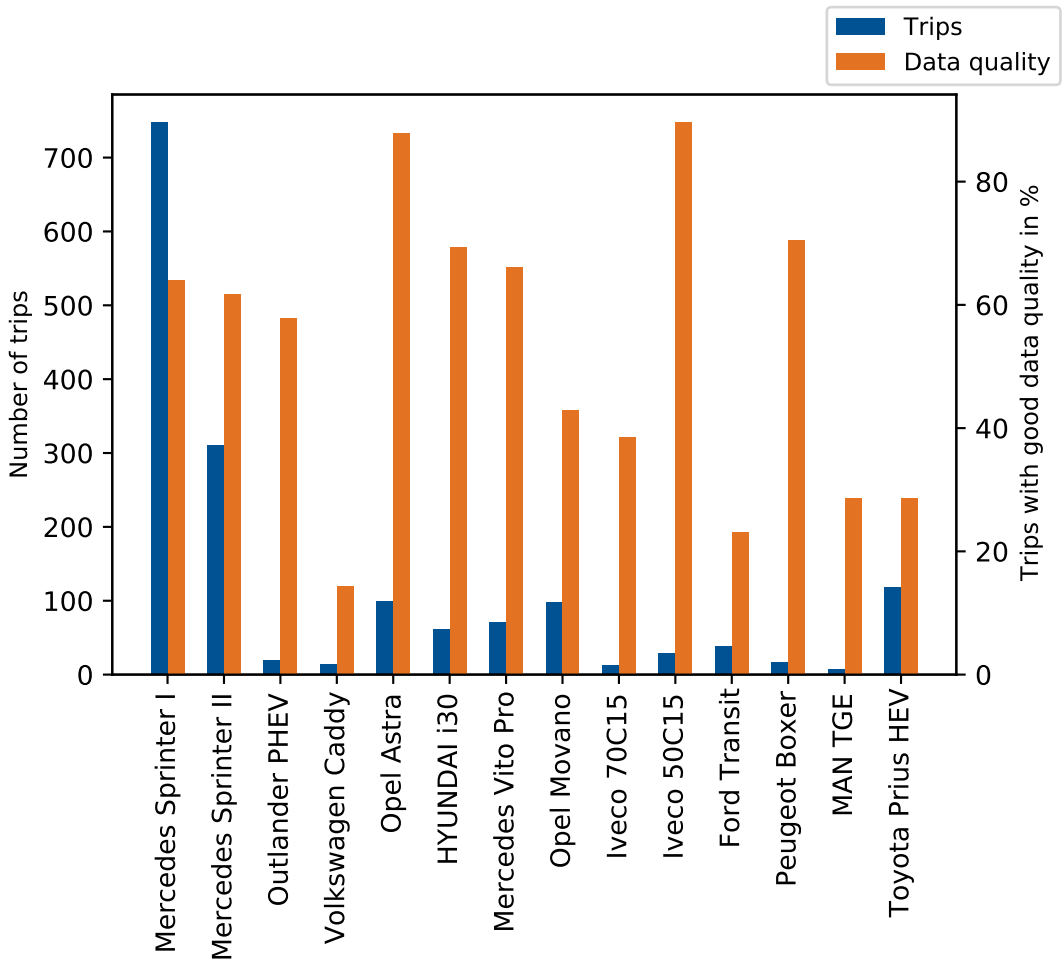


Figure 5.10: Number of trips per vehicle and percentage of trips with a satisfactory data quality

another. Two-thirds of the trips belong to a Mercedes Sprinter, either type I or II. Even considering that 13 out of the 30 studied vehicles are Sprinters, this data shows that they are over-represented. On the contrary, for the MAN TGE there are only 7 daily trips available. This can be worsened when, additionally to a low number of trips, their distance quality is also poor. In Figure 5.10 the number of trips of each vehicle and the percentage of trips with satisfactory data quality are shown. A satisfactory data quality is defined in this plot as a distance error in the range between −1 % and 5 %. Vehicles such as the Volkswagen Caddy, the Ford Transit and the MAN TGE show both a very limited number of and trips and poor data quality on the majority of them.

5.3 Simulation Results using Recorded Trips

In this section, the consumption results obtained by using the simulation framework for GPS-derived trips are presented. The section is divided into several parts. First, the results are validated using the Mercedes Pro data. Secondly, the validation using the Spritmonitor data is presented and finally, the influence of the payload in the overall fuel consumption rate is examined.

Mercedes Pro Validation

This first validation is only possible with a subset of Fleet 1, specifically, the 9 Mercedes Sprinter I and. The existing real-world consumption data is provided for each of those vehicles through the Mercedes Pro fleet monitoring Platform, and the rates are available individually for each vehicle and on a daily basis for a limited time range. This consumption data is referred to as the Mercedes Pro rates. For those same vehicles, the consumption framework is applied using the GPS-derived trips to get the daily consumption rates, which are referred to as the Fastsim rates. Finally, to put into perspective the overall accuracy results, the Mercedes Pro and Fastsim rates are compared with the homologation consumption rate of the Mercedes Sprinter, as this would be the only available consumption rate, should the consumption framework didn't exist.

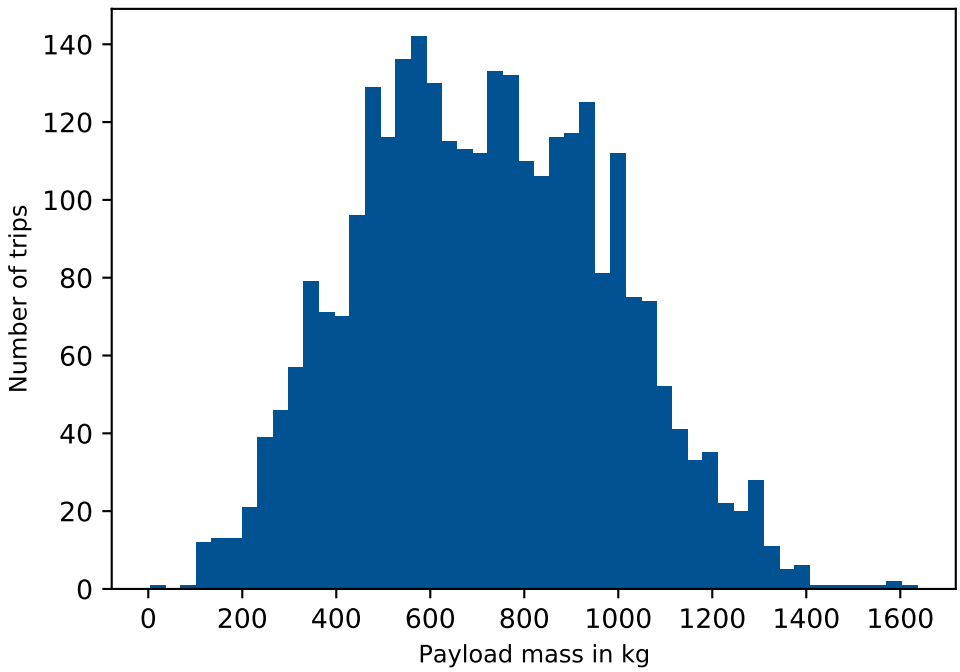


Figure 5.11: Histogram of the payload mass of the Fleet 1 transporters

In order to achieve the highest possible accuracy, only those daily trips with a distance error under 2 % are used, so the total number of trips is 103. Note that in this case, by daily trips is meant the concatenation of all the individual trips of a specific vehicle during a day. This is obligated given that the Mercedes Pro rates are provided on a daily rate. Before performing the simulations, it is necessary to have some general knowledge about the task that these transporters are doing. The daily mission of each vehicle consists of visiting several destinations that are relatively close to each other and at each stop, exchange some goods. It can be assumed that the mass of the outgoing goods is equal to the mass of the incoming, therefore, the vehicle extra mass can be considered constant. We do have historical data on the payload mass for many daily trips, however, the exact mass of the load for each vehicle and day is unknown. That's why the distribution of the payload masses, in Figure 5.11 is useful. In this regard, we have done three different simulations considering the most representative extra loads of 500 kg, 750 kg and 1000 kg respectively.

For each simulation, 4 key parameters are calculated. The NME, which is calculated as in Equation (5.1), with $FC_{\text{Framework}}$ as the sum of the fuel consumption of each day using the framework and $FC_{\text{MercedesPro}}$ as the sum of the fuel consumption recorded in the Mercedes Pro Platform.

$$NME = \frac{FC_{Framework} - FC_{MercedesPro}}{FC_{MercedesPro}} 100 \tag{5.1}$$

The MAE and ME are calculated as in Equation (5.1) and (5.3), where n is the number of trips, and $Rate^{\text{i}}_{\text{Framework}}$ and $Rate^{\text{i}}_{\text{MercedesPro}}$ are the consumption rates of each of the trips, using either the framework or the Mercedes Pro respectively.

$$MAE = \frac{\sum_{i=0}^{n} \frac{\left|Rate^{i}_{Framework} - Rate^{i}_{MercedesPro}\right|}{Rate^{i}_{MercedesPro}}}{n} 100 \tag{5.2}$$

$$ME = \frac{\sum_{i=0}^{n} \frac{Rate^{i}_{Framework} - Rate^{i}_{MercedesPro}}{Rate^{i}_{MercedesPro}}}{n} 100 \tag{5.3}$$

The standard deviation of the results, which is calculated analogue to Equation (5.3). The above presented formulas are also used analogously when calculating the results with homologation consumption rate and the consumption rates are expressed in $l/100km$.

The results are presented in Table 5.3, which shows that the MAE is around 6 % for all configurations and that the simulation with a 750 kg payload, which is the closest to the mean payload, has almost no mean error and just a 3 % NME. In contrast, the estimations using only homologation data present errors well above 20 %.

Table 5.3: Results of the Mercedes Pro Validation.

	Payload	NME	MAE	ME	Std. dev.
Framework	500 kg	−7.00 %	6.48 %	4.51 %	9.32 %
	750 kg	−3.45 %	5.91 %	0.83 %	9.53 %
	1000 kg	0.18 %	7.14 %	−3.01 %	9.86 %
Homologation	-	−26.5 %	23.7 %	−23.7 %	7.46 %

Spritmonitor validation

Since the validation using the Mercedes Pro platform data is limited to one vehicle type, we need a further validation procedure that includes the complete fleet. This is done using the publicly available Spritmonitor data, a crowd-sourcing website where the users upload the real consumption rates of their own vehicles. For each of the vehicles in our fleet, we use the available filters to find a comparable subset in Spritmonitor that matches as much as possible our virtual fleet. Therefore we select the vehicles with the same powertrain, a similar fabrication year, same power and if possible, the exact same variant. This filtering returns a extremely variable amount of data depending on the model. The most popular passenger cars have hundreds of users uploading their consumption rates and allow precise filtering, while the more exotic ones (normally transporters) have much less data, which makes it impossible to choose the exact variant. This fact should be taken into account when discussing the results.

Table 5.4: Results of the Spritmonitor validation, the consumption rates are in l/100km.

Vehicle	Sprimonitor				Fastsim		Difference
	Users	Min	Mean	Max	Trips	Rate	
Fleet 1							
Mercedes Vito	55	6.04	8.35	10.3	64	8.41	−0.71 %
Hyundai i30	398	4.13	7.32	10.6	35	6.48	13.9 %
Opel Astra	561	4.01	5.90	8.83	87	5.35	10.3 %
Volkswagen Caddy	128	5.04	6.50	9.27	14	7.33	−11.3 %
Toyota Prius	51	4.12	5.52	6.60	32	3.98	38.7 %
Mitsubishi Outlander PHEV	231	0.23	4.28	8.94	15	4.31	−0.70 %
Mercedes Sprinter I	97	8.06	10.5	16.4	734	10.5	−0.10 %
Mercedes Sprinter II	176	8.06	11.3	17.7	303	10.6	6.92 %
Fleet 2							
Opel Movano	6	7.31	9.65	10.5	80	9.93	−2.82 %
Iveco 70C15	13	9.12	14.3	17.1	11	12.3	15.6 %
Iveco 50C15	16	9.28	12.7	15.6	25	8.99	40.8 %
Ford Transit	40	7.58	9.28	12.1	32	10.6	−12.0 %
Peugeot Boxer	10	8.97	10.4	11.3	13	8.66	19.6 %
MAN TGE	6	9.67	11.9	13.0	20	9.28	28.7 %

Spritmonitor provides the minimum, mean and maximum consumption rate of the selected vehicles. In this comparison, the Spritmonitor mean rate will be compared to the mean consumption rate of each vehicle simulated with the framework. The passenger cars will be simulated with the minimum extra mass of 136 kg and for the transporters, a partially loaded vehicle will be considered, so the extra mass will be 750 kg. Only daily trips with at least 20km and a maximum of 10 % of distance error are considered. The results are shown in Table 5.4.

Influence of extra load in Transporters

When estimating the fuel consumption of a fleet, it is interesting to identify the parameters that influence the most the overall consumption rate. In this regard, the extra load in the light commercial vehicles is certainly suitable to have a notable impact, since the mass difference between an empty and a fully loaded vehicle may reach the 1000 kg. To quantify this influence, the transporters from both Fleet 1 and 2 have been simulated considering an empty vehicle with only a driver, with a total extra mass of 136 kg and a loaded vehicle, with a total extra mass including the driver of 750 kg, this is, with a 614 kg payload.

For the simulation, only the daily trips longer than 20 km with a distance error smaller of 10 % are considered. In Table 5.5 the results are presented individually for each type of transporters and aggregated for all the studied trips. Note that most of the trips correspond to a Mercedes

Table 5.5: Results of the extra load comparison in the transporters.

Vehicle	Trips	Consumption rate [l/100km] Empty	Loaded	Difference
Fleet 1				
Mercedes Vito	64	7.41	8.41	13.4 %
Volkswagen Caddy	14	6.20	7.33	18.2 %
Mercedes Sprinter I	734	9.47	10.5	10.6 %
Mercedes Sprinter II	303	9.45	10.6	11.7 %
Fleet 2				
Opel Movano	80	8.91	9.93	11.5 %
Iveco 70C15	11	11.3	12.3	9.27 %
Iveco 50C15	25	8.24	8.99	9.10 %
Ford Transit	32	9.4	10.6	12.2 %
Peugeot Boxer	13	7.58	8.66	14.3 %
MAN TGE	20	8.23	9.28	12.7 %
Total consumption difference	1296	-	-	10.2 %

Sprinter, since this is the most used model. The results show that driving a partially loaded transporter increases its consumption around 10 % on average. This increment is not equal for all vehicles, so for instance the Caddy registers up to 18 % additional consumption due to the payload, while the increment is limited to a 9 % for both Ivecos.

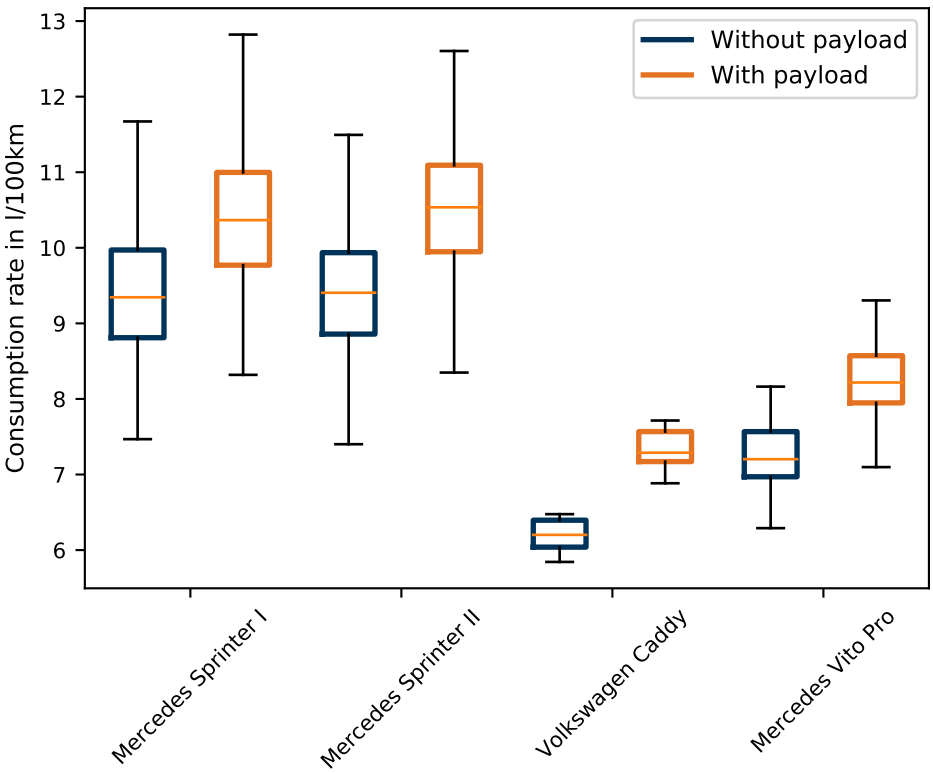


Figure 5.12: Mass comparison of Fleet 1.

It is also interesting to observe how the consumption rates are distributed since there is a non negligible deviation between the results. This can be easily represented using box plots, as in Figure 5.12 and 5.13. Two main ideas arise from these plots. First, there is a high deviation between the fuel consumption within a trip, and this is specially true for the vehicles with a high

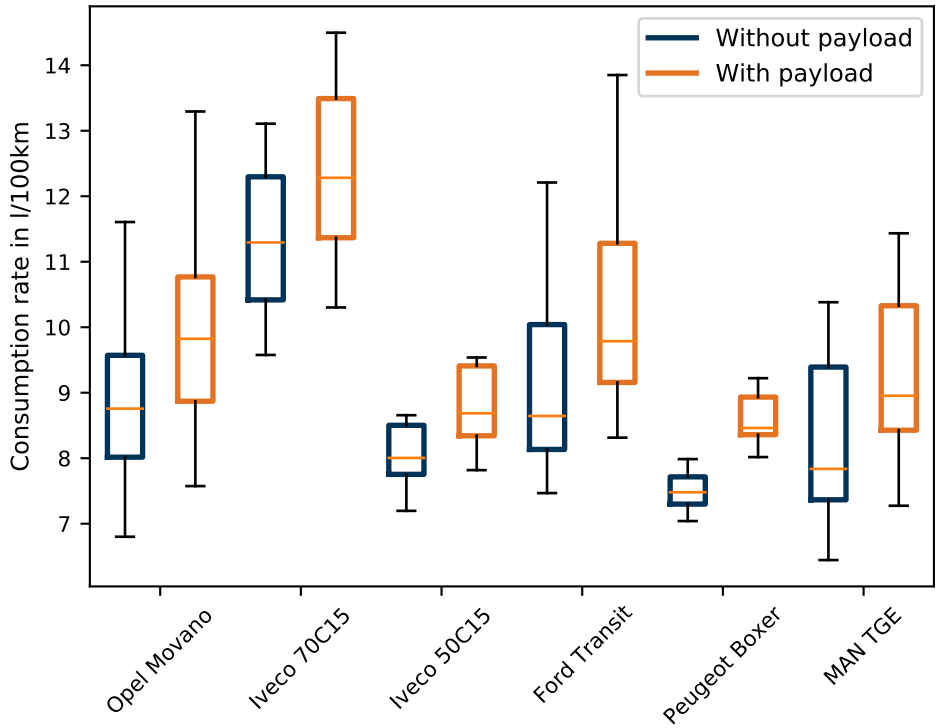


Figure 5.13: The validation results of Fleet 2.

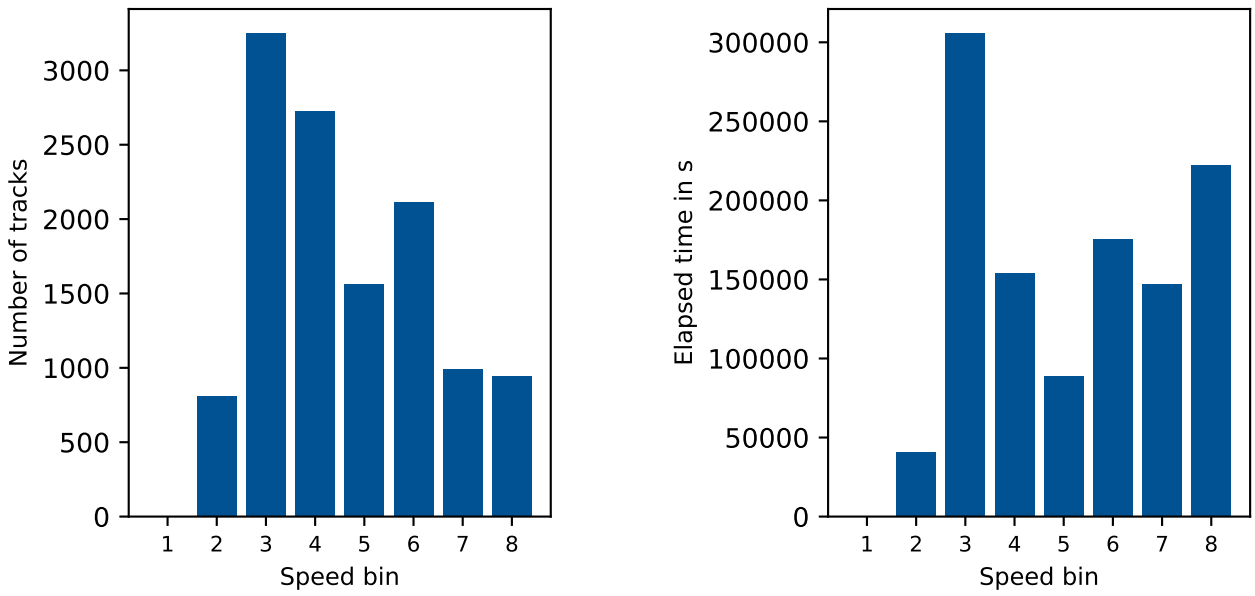
number of recorded trips. Secondly, the influence of the payload is lineal and proportional for all trips, so the deviation on the consumption rates is kept constant and the difference between the boxplots with and without payload is an upwards shift indicating the consumption increase.

5.4 Simulation Results using non-Recorded Trips

This section presents the results of the non-recorded trips and its validation. This simulation procedure is based in the already mentioned speed bins of 15 km/h, however, besides this speed bin configuration, which have turned out to be the most effective, there has been experimentation with some other configurations. The results achieved with this alternatives are also presented in this chapter.

The first relevant results in this section correspond to the training tracks, this is, the tracks used for creating the necessary speed profiles for deriving the consumption curves. There are two aspects to consider in this regard, first, the number of tracks in each speed bin, shown in Figure 5.14a. The bar plot shows that there are hundreds of available tracks in all speed bins except for the first, which is the one corresponding to lower speeds. In fact, the lack of tracks for the first bin is responsible for most issues and difficulties during the process. On the other hand, it is also remarkable the high number of tracks in the speed bin 3 and 4, these are the speeds between 30 and 60 km/h, suggesting that this is the average speed at which the vehicle drive the most of the time.

A second key indicator is the length of the derived speed profiles at each bin, that originate from the concatenation of the above-mentioned tracks. This is shown in Figure 5.14b. Again most speed bins have a large speed profile, while the second and especially the first one lack of that



(a) Number of tracks.

(b) Number of tracks.

Figure 5.14: Bar plot of the training data distribution.

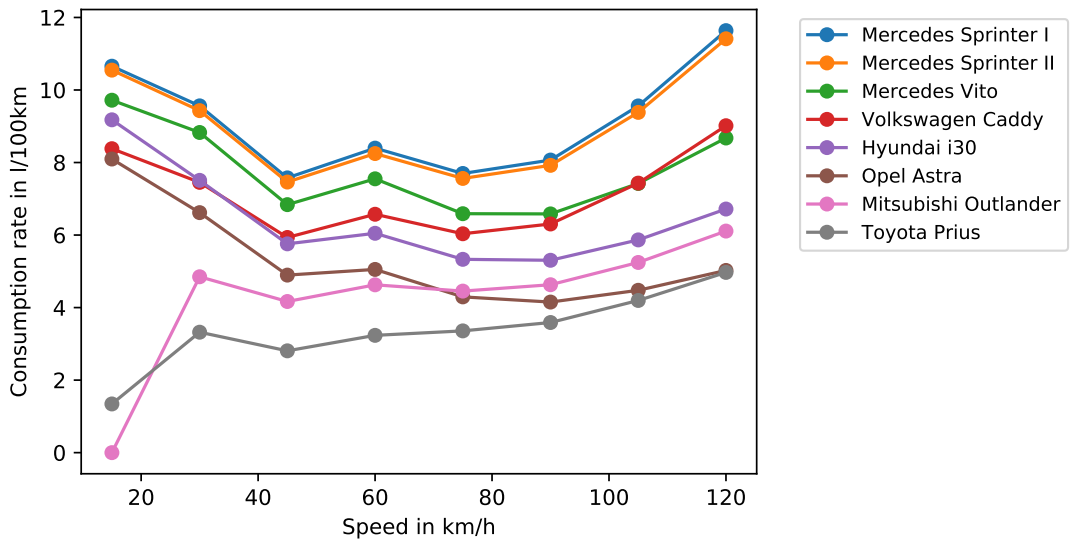


Figure 5.15: Consumption curve from fleet 1.

much data. This is a direct consequence of having a small number of tracks. The high-speed bins, such as 7 and 8 present a much longer speed profile than it could be expected, had all the tracks the same average length. It is however reasonable to think that the tracks recorded in roads and highways are considerably longer than the ones in urban areas where the speed distribution is much less constant. The speed profiles at each speed bin of all the versions presented in this section are shown in the appendix, in Section A.4.

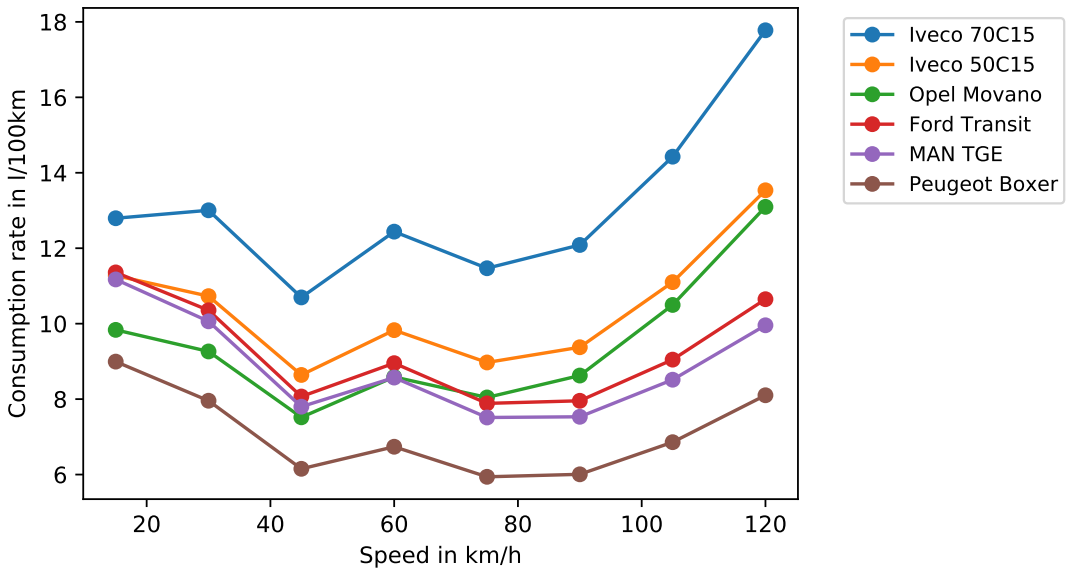


Figure 5.16: Consumption curve from Fleet 2.

The above-mentioned speed profiles are used for deriving vehicle-specific consumption curves. They are said to be vehicle specific because FASTSim calculates the rates at each bin using the corresponding virtual vehicle. However, it is important to remark that the input speed profile at each bin is exactly the same for all vehicles. This curves are shown in Figure 5.15 and Figure 5.16.

Since Fleet 1 is the most diverse in terms of vehicle typologies, their consumption curves have many interesting trends to highlight. The two Mercedes Sprinter variants have perfectly parallel curves with a small shift that sets variant I as the more pollutant. This is produced by separating the same vehicle into two very similar variants. The rest of the curves, in contrast, are not parallel and shifted, but present various shapes, which indicates different consumption behaviors. Other transporters such as the Mercedes Vito and the Volkswagen Caddy have however a similar shape to the Sprinter, characterized by a high consumption increase at high speeds due to the aerodynamic drag. The passenger cars such as the Astra or the i30 are much more aerodynamic and thus, their consumption doesn't grow that much, however, when driving at low speeds it is considerably higher. A completely different tendency is highlighted by the hybrid vehicles. Both the Prius as a HEV and the Outlander, a PHEV have very low consumption at the first bin. In fact, the Outlander has zero consumption, meaning that it is able to rely purely on its electric motors al low speeds. In the following speed bins the consumption increases gradually, but it is kept well below the other vehicles.

Fleet 2 is much more homogeneous in terms of shape, due to the fact that it is purely integrated by Transporters. It is worth mentioning, the extremely high consumption that the Iveco 70C15 records at the high speed bins. Also, the Opel Movanos curve cuts the Ford Transit and MAN TGE curves, which indicates a different curve shape, due to a higher than normal consumption rate at

high speeds. Finally, there is also a double minimum affecting all curves, specifically, the first minimum is located in the third bin and the second minimum is at the fifth bin.

Table 5.6: Validation results of the non-recorded trips simulation framework.

Vehicle	Trips	Length [km]	NME	MAE	ME	Std. dev.
Mercedes Sprinter I	3079	26276	−1.64 %	11.8 %	−0.17 %	14.5 %
Mercedes Sprinter II	1134	7400	−4.76 %	12.7 %	−4.70 %	14.6 %
Opel Movano	11	58	17.2 %	11.1 %	10.4 %	10.5 %
Mercedes Vito Pro	161	3432	0.36 %	10.4 %	2.32 %	12.8 %
Opel Astra	291	4059	−3.36 %	12.6 %	−2.97 %	15.2 %
Toyota Prius	165	659	6.70 %	9.14 %	6.23 %	9.26 %
Aggregate	4841	41886	2.16 %	11.9 %	−1.07 %	14.61 %

The validation of the results is done using a large data set of trips and excluding the ones used in the training phase. For each track, an estimation for both recorded and non-recorded trips is conducted and the results are compared in terms of l/100km. To ensure proper comparability, this data set is filtered to discard the trips with low data quality, particularly, with a distance error higher than 5 %. This is a much less strict threshold than the one used for the training phase since in this case, the studied variable is only the trip consumption rate, which is an aggregate value and therefore less sensitive to small deviations. Still, this threshold is too high for most of the vehicles, since its data quality is low. Therefore, only 7 of the 14 vehicles have at least one suitable track for comparison. The aggregated results of the validation include 4841 tracks with a total distance of 41 886 km.

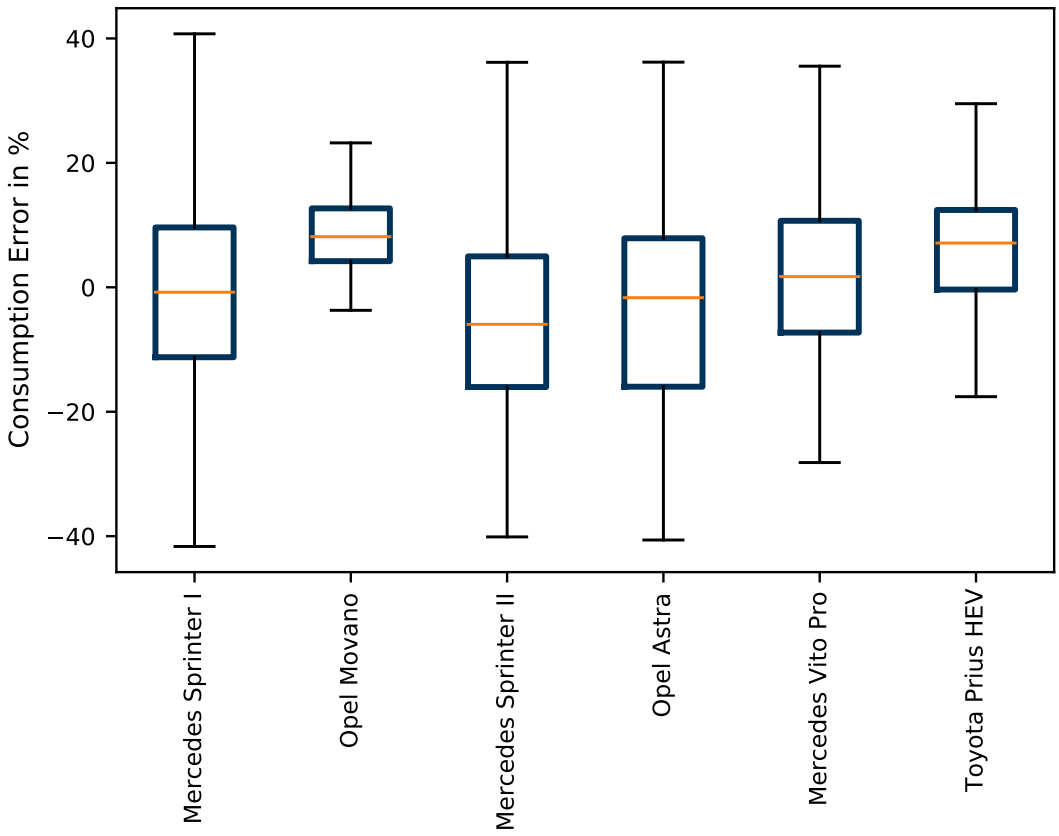


Figure 5.17: Boxplot of the consumption errors of each vehicle.

The results are shown in Table 5.6. Note that the Mitsubishi Outlander results are not displayed in this table, because those simulations produced inconsistent values. The rest of the vehicles, however, present very promising validation results. As said, the majority of tracks are Mercedes Sprinters and both variants account for over 80 % of the total trips. Both variants present similar MAE, around 12 % and standard deviations at 14.5 %, but the NME and ME are a bit more distant and the Sprinter II has a greater underestimation. The underestimation is denoted by the minus sign and refers to a minor rate using the simulation framework for non-recorded trips. The Opel Movano has only 11 trips, so this data doesn't allow for many conclusions. Both the NME and ME are clearly above-average indicating that there is an overall overestimation of the consumption and additionally, within a smaller range since the standard deviation is only a 10 %. The Mercedes Vito has much more accurate results and it almost matches the overall fuel consumption, as show its 0.36 % of NME. Similarly, the Opel Astra have a slightly underestimation with a NME and ME of around 3 %. Whereas the Toyota Prius, a closer look is required since it has a different powertrain technology as it is a HEV. Provided that its results are generated using 165 tracks, they are indeed statistically meaningful opposed to the case of the Movano. The validation shows an overestimation of the consumption rate of around 6 %, however, both the MAE and the standard deviation are much lower indicating a smaller range and thus, higher consistency of the estimations.

In aggregate, the validation of the consumption estimation for non-recorded trip present a 2.16 % of NME, and −1.09 % of ME. The MAE is 11.9 % and the standard deviation is as high as a 14.6 %, indicating a high variance in the results. This high variance is best shown in Figure 5.17 where the consumption errors for each of the vehicles are summarized in a boxplot. All vehicles present a large interquartile range and even a larger range between the minimum and maximum values. The median, however, is close to 0 indicating that the results are well centered, which together with the good average values presented, indicates a good accuracy, despite the low precision.

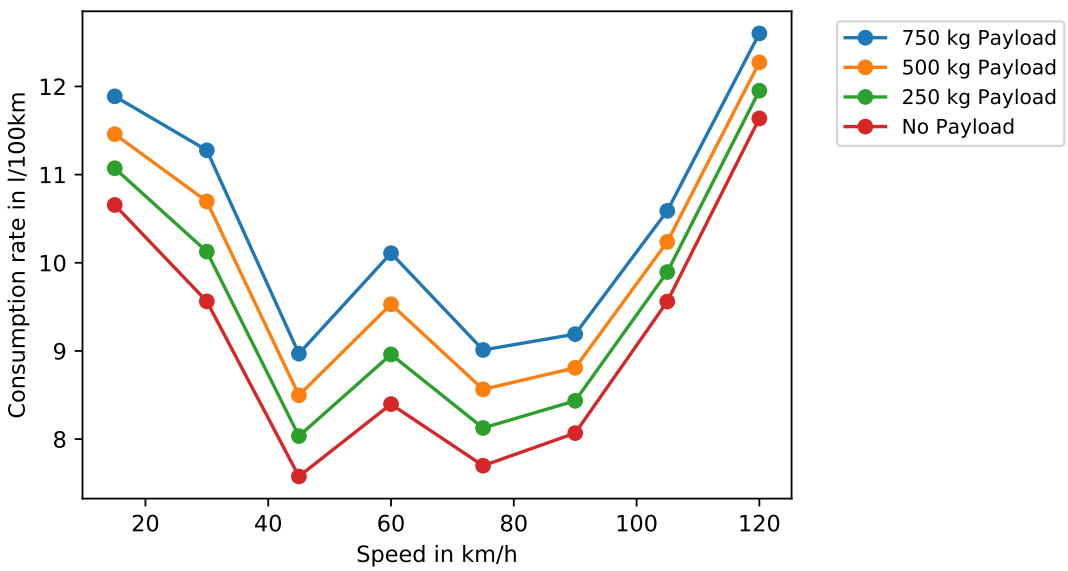


Figure 5.18: Consumption curves from the Mercedes Sprinter I considering different payloads

As seen in Section 5.3, the payload has a significant impact on the fuel consumption of a transporter. Therefore, it is necessary to check if the consumption curves can also handle it and produce accurate estimations that also account for the additional mass of the transporters produced by its payload.

For doing so, the consumption curves have to be rerun and include the corresponding payloads. An example of this is presented in Figure 5.18 and illustrates the resulting consumption curves of the Mercedes Sprinter I considering no payload and additional payloads of 250 kg, 500 kg and 750 kg. Logically, all the curves are parallel to each other and vertically shifted, so the greatest payload creates the highest consumption curve.

Table 5.7: Validation results of the non-recorded trips simulation framework for transporters considering a payload.

Vehicle	Trips	Length [km]	NME	MAE	ME	Std. dev.
Mercedes Sprinter I	3165	27785	0.07 %	12.4 %	1.24 %	15.3 %
Mercedes Sprinter II	1145	7847	−4.43 %	13.1 %	−3.80 %	15.4 %
Opel Movano	11	58	21.0 %	14.6 %	14.0 %	11.3 %
Mercedes Vito	184	4022	2.86 %	10.8 %	3.39 %	13.7 %
Aggregate	4505	39712	−0.58 %	12.54 %	0.80 %	15.4 %

To assess the effect of the payload in the estimations, the same validation process as before is performed. However, this time only the transporters are be considered, namely both Sprinter variants, the Movano and the Vito. All trips are simulated with a constant payload of 750 kg. The results of the validation are in Table 5.7. All relevant parameters have similar values to the former validation, however, the Opel Movano presents a huge NME and ME, which are even higher than the former and imply a substantial overestimation. On the other hand, the errors of the other vehicles are minor in this validation, and in aggregate, both the NME and the ME are extremely low, in both cases below 1 %.

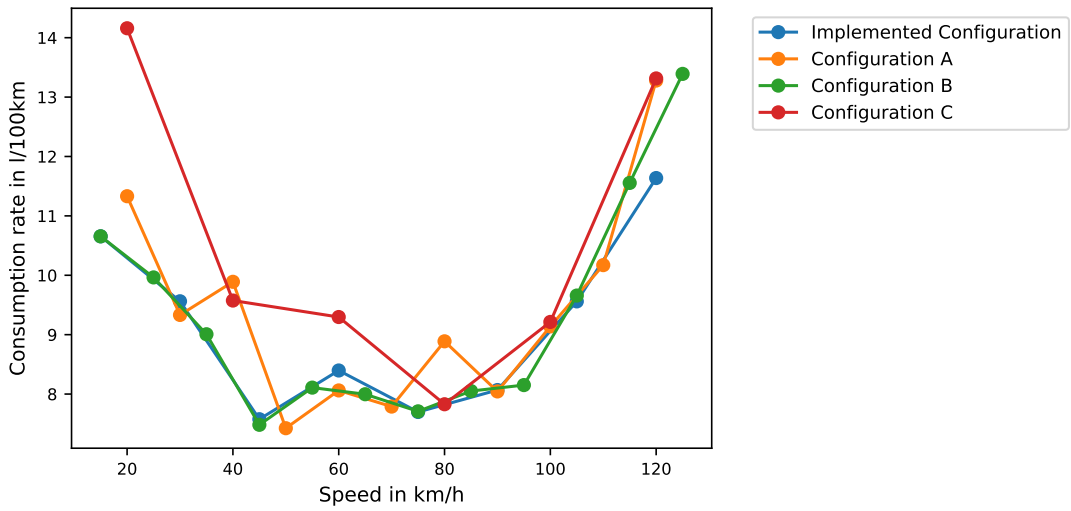


Figure 5.19: The consumption curves of alternative implementations compared to the implemented configuration. All curves are derived using a Mercedes Sprinter I.

As mentioned in the Implementation chapter, there has been experimentation with alternative speed bin configurations and vehicle subsets for creating the most reliable consumption curve possible. The most relevant attempts are the ones considering the alternative speed bin configurations A,B,C and the speed curves created using only tracks from transporters, instead of using all the available vehicles. Moreover, there is the virtual cycle approach also discussed in the implementation chapter, in which a speed profile is created using the average speed of the links of the routing engine. In all situations, the validation data used was the same, specifically the data from the Mercedes Sprinter I, which includes over 3000 trips with more than 25 000 km

driven, so the only difference lies in the approach attempt, the training data used and the speed bin configuration.

In Figure 5.19 the consumption curves using all the different speed bin configurations are shown. All the curves are simulated using the same vehicle, a Mercedes Sprinter I, and the same training data, with the exception of the configuration C in which a smaller subset of training data was used, do to the fact that it has fewer speed bins to calculate. The consumption curve of the configuration A shows an evident over-fitting when compared to the implemented configuration. This over-fitting is reduced in the configuration B, however, as with A, the consumption in the high-speed bin is much higher than in the implemented version. In contrast, the configuration C, which is the one registering the worst results among the four variants, has a very high consumption in all the urban-driving related bins, in this case, the bins 1, 2 and 3. As a reminder, the speed bin configurations are presented in the implementation chapter and summarized in Table 4.4 to Table 4.7.

Table 5.8: Validation results of alternative implementations for the non recorded trips simulations. All the results correspond to a Mercedes Sprinter I.

Speed bin configuration	Tracks used	NME	MAE	ME	Std. dev.
Configuration A	All vehicles	−5.37 %	14.18 %	−9.1 %	15.3 %
Configuration B	All vehicles	3.06 %	12.18 %	0.31 %	14.9 %
Configuration C	All vehicles	−9.94 %	18.1 %	15.4 %	16.4 %
Implemented configuration	Only Transporters	2.42 %	11.7 %	0.95 %	14.3 %
Virtual cycle approach		−21.2 %	21.0 %	−19.5 %	13.8 %
Benchmark		−1.64 %	11.8 %	−0.17 %	14.5 %

The achieved results, shown in Table 5.8, are relatively worse than the ones achieved with the final implementation, which are also presented in the Benchmark row of the table. Note that these values are exactly the same that the ones in the first row of Table 5.6. Still, the differences are not that big and actually the simulation using tracks only from Transporters achieves a slightly better MAE and smaller standard deviation. On the other hand, the virtual cycle approach registers a very poor accuracy and is clearly biased towards an underestimation of the consumption.

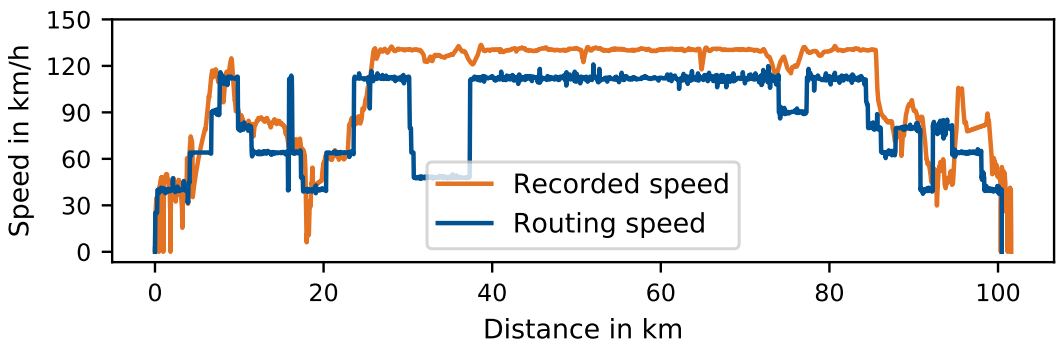


Figure 5.20: Comparison of a virtual speed profile, using the routed speed, and the correspondent recorded speed profile.

To understand this behavior, it is mandatory to look deeper into the simulated speed profile. In the example presented in Figure 5.20, the recorded speed profile using the GPS tracker is compared to the virtual cycle, derived from the routing. Both speed profiles correspond to the same trip and have the same length. Also, the average speed is more or less equivalent, however,

the virtual cycle lacks the stop-and-go behavior typically found in urban driving, which in this case occurs at the beginning and end of the cycle. Moreover, the real vehicle drove at a higher speed on the autobahn compared to the predicted speed of the virtual cycle. This is, in any case, just one example out of the thousand trips simulated that can be considered as relatively accurate. Still, the virtual cycle approach underestimated the fuel consumption of this trip by a 14.9 %.

5.5 The error chain

In this section, the error chain of the consumption framework is presented. The intention is to provide a detailed description and context on the existing sources of error of the framework and also to summarize the results presented in the previous sections. As in previous chapters, it is divided into two parts, the first part addresses the error chain when estimating fuel consumption with recorded trip data, and the second one describes the error when estimating fuel consumption without recorded trip data.

Error chain using recorded trips
There are four main sources of error when estimating the fuel consumption of recorded trips using the presented framework, as described in Figure 5.21. The first one, named distance error is a direct consequence of the vehicle tracking system inaccuracies, that have been already mentioned in former chapters. This source of error lies in imperfect raw data quality and the key indicator for assessing its quantity is the difference between the geometrical distance of the spatial points and the trip distance derived from integrating the speed profiles. Note that the vehicle tracking and database maintenance are not part of the present work, however, its influence on the distance error can't be neglected. The trip generation, extensively described in Section 4.2 and presented in Section 5.2 is indeed in the scope of the Thesis and has a non-negligible effect on the distance quality and thus, in the overall fuel consumption. The error in the road grade, on the other hand, is negligible in most of the trips and therefore is disregarded in the error chain.

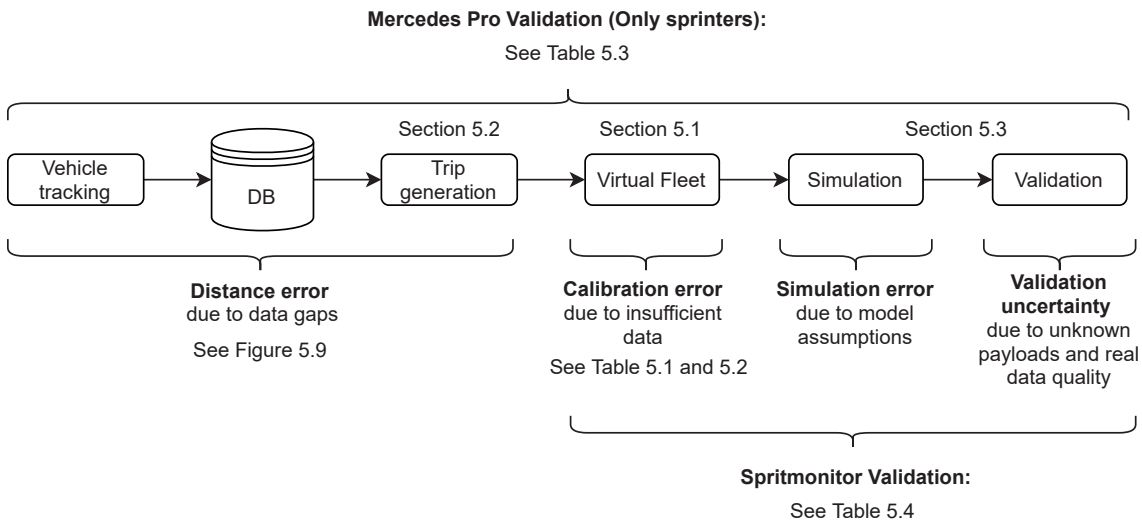


Figure 5.21: Diagram of the error chain for the recorded trips simulation.

The second source of error is found during the calibration of the virtual vehicles and is due to either insufficient other inaccurate vehicle data. This includes the homologation fuel consumption

used for validating the virtual vehicles. The results of this issue are presented in Section 5.1 and show a weighted average absolute error of 1.6 % and 5.84 % for Fleet 1 and 2 respectively.

A third source of error is due to the simulation model itself an lies in the assumptions in which the model is based. This error is not isolated from the others but it certainly exists and is discussed in the validation report of the model written by its own developers.

Finally, there is a certain level of uncertainty regarding the validation, since there are some unknowns regarding the data used. For the Mercedes Pro validation, the main unknown is the additional payload of each trip, as the exact information about the mass of payload that was being carried by the vehicles was not provided. Thus, for the validation different constant payloads were assumed. Whereas the Spritmonitor validation, the unknown is the actual comparability of the vehicles and their consumption rates, as the Spritmonitor data has some limitations specially in transporters. The reference rates are calculated by Spritmonitor after specifying a filter that allows to only consider specific models and variants, construction year, powertrain type and engine power. This filters yield a reduced amount of trips for the less popular vehicles, providing an unstable basis for comparison. Moreover, there is no information about the payloads that the real vehicles were carrying, so for the transporters simulation a total extra mass of 500 kg, including the driver, was assumed.

Two different validations of the framework have been presented. The Mercedes Pro validation can be used to validate the complete framework, from the vehicle tracking system to the simulation of the results since it is based on real trip-specific consumption rates that were gathered independently. Its results, presented in Table 5.3, correspond to Sprinters alone and show a ME that varies between −3.0 % and 4.5 %, depending on the assumed cargo mass. The Spritmonitor Validation, on the other hand, can only assess whether the simulated results are consistent with real drive consumption rates gathered by a crowd-sourcing platform. The distance error generated by the tracking device is not taken into account, since the trip-specific consumption rate is not examined.

Error chain using non-recorded trips

The error chain of the non-recorded trip consumption simulation is divided into two parts, the first corresponding to the training phase and the second to the simulation. The diagram in Figure 5.22 summarizes the results.

The training phase has two main sources of error. The creation of the speed profiles at each speed bin, in which the recorded speed profiles are matched with their equivalent routed trip produces a matching error since the trip data is not map-matched and thus, a track of a specific speed bin could be shifted and wrongly assigned to such bin, totally or partially. This would produce a completely flawed consumption curve, so an extraordinary high distance correspondence of 99.9 % is enforced, as explained in Section 4.3. Therefore, the maximum matching error could be assumed to be a 0.1 %.

The second source of error in the training phase is produced while calculating the consumption curves using FASTSim and the virtual fleets, so this phase is subject to the same calibration error and simulation error that the consumption framework for the recorded trips, which has already been introduced.

Whereas the simulation phase, there are three sources of error. First, there is a routing error, which can be compared to the distance error found in the framework for recorded trips. This error is due to the multiple possible routes that can be driven between and origin and destination. The routing algorithm chooses the fastest one according to its algorithm, but the actual driver

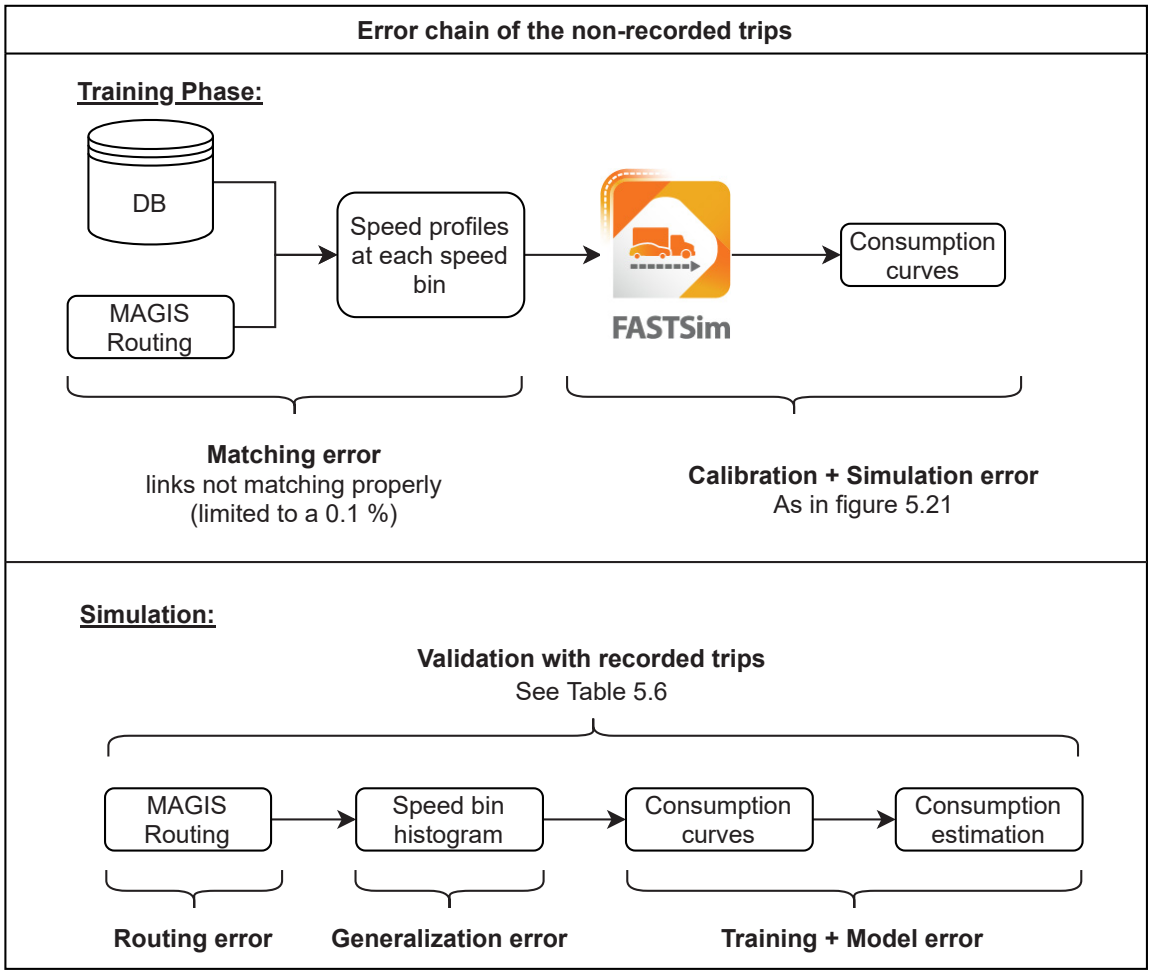


Figure 5.22: Diagram of the error chain for the non-recorded trips simulation.

may choose a completely different one, for many different reasons. In some cases, the difference is found in the approximation to the destination or in the first instants of a trip, which generates a minor error. In other situations, the chosen route is completely different from the start to the end of the trip. Contrary to the matching error, which must be limited to ensure the overall quality of estimation, this error must be considered as a part of the framework and incorporated into the overall simulation error.

The second source of error is the averaging of the load profile, which is done by creating a speed bin histogram out of the routing trip data. The raw data of the histogram is the average speed at each of the links of the routed trip, which was calculated by the developers of the routing algorithm by averaging the speeds of many vehicles driving in that same link. Thus, the created histogram is representative of the average trip but is not accounting for specific driving behavior. This averaging error, in fact, is also produced when creating the speed profiles for each speed bin and originates in the routing, however, the creation of the histogram is the materialization of the error and for the sake of simplicity, it is only attributed to this step.

Finally, in the consumption estimation using the consumption curves, two additional errors arise. First, the already mentioned training error, that influences the simulation at this point and secondly the model error. Again, the model error is assigned to the last step, in which it becomes evident despite it is actually present in all of the previous phases. In this case, the model error is based in the following assumptions:

- The average speed of the link is valid and sufficient as a descriptive variable for the fuel consumption.
- The created speed profiles at each speed bin are representative of the average driving behavior.
- The fuel consumption of a vehicle can be predicted as a function of its average speed.

These assumptions appear to be valid for all the vehicle typologies except the PHEV, in which the third assumption, that the fuel consumption is a function of the average speed, doesn't hold anymore. The reasons for this are found in the discussion.

The above-mentioned errors are computed in the validation using the recorded trips, summarized in Table 5.6. This validation was limited to a subset of the vehicles, due to a lack of data with the required quality, but still could include 4841 trips from 6 different vehicles and achieve satisfactory results, which are further commented in the discussion.

6 Discussion

In this chapter the methodology, the implementation and the results are discussed and referenced to the literature. The chapter is divided into different parts according to the relevant topics of the project. First, the creation of trips using the GPS data is discussed. Next the discussion shifts to the virtual fleet. Then the result of the Mercedes Pro and Spritmonitor validation are commented, which also includes the discussion on the error chain. Finally, the complete methodology for estimating the fuel consumption of non-recorded trips is discussed.

Creation of trips using GPS data

In the present work, thousands of trips that were recorded with a GPS tracking device are used for estimating fuel consumption. Since it is experimental data, these recordings have inaccuracies and errors that should be solved in order to produce reliable estimations. Even the most sophisticated data treatment may contain inaccuracies, however, the objective is not to create perfect trips but to limit some of the errors so they don't affect the overall estimation significantly. Moreover, the data treatment must ensure a way to detect the significant inaccuracies on the trip and discard the trip, where the errors unassimilable. This is done using the trip length rate, which compares the trip length derived from the speed profile with the trip length derived from creating a line using the GPS points. Since the line won't be affected by data gaps, they should be almost equivalent. The more divergent this lengths are, the more errors and inaccuracies the trip may contain.

The creation of trips is closely related to the simulation model used. FASTSim requires a speed profile defined at 1 Hz and a road grade time series. Thus, the data processing transforms the original raw data to the desired frequency using a moving window. Additionally, the road grade information is added using an elevation map. This process can be implemented in many different ways, but the evidence of comparing different implementation procedures showed a very low impact on the overall output. On the other hand, the trip interpolation and filtering does impact considerably the overall consumption, especially the gap-filling algorithm and the removal of long idle times. For the gap filtering, the approach was as simple as possible and only covers the gaps where the vehicle was obviously moving, leaving aside other situations of doubt. Thus, the majority of the distance error will originate from this filtering. There is not a realistic way to correct this, since the problem lies in an absence of data. The commonly used Kalman filters would require a very demanding calibration error and the overall gains would be unclear. Whereas the road grade calculation, it doesn't have a major impact on most of the trips, since the vehicles were recorded in an area that is mostly flat, however some cases have been shown in which the road grade was indeed non-negligible, so the procedure for calculating makes the overall framework much more robust.

Virtual fleet

During the virtual fleet creation, 14 different virtual vehicles representing their real-world counterparts were created. The process is supported by the excel distribution of FASTSim and the data used is mainly found in the Vehicle Registration Certificate of each vehicle. The most

relevant part of the process is the calibration using the homologation cycles. This process is straightforward for conventional vehicle architectures since the homologation rates are calculated by simply running the cycle. In hybrid vehicles, however, the process is more complex, so it could not be customized using the European homologation cycles. This limits the ability of the framework to model some hybrid vehicles that are not sold in the US, but in general, it is something rare.

The calibration results, which for most vehicles are very accurate, should be taken with caution. Just because we have been able to simulate the consumption throughout a drive cycle with around 1.6 % of error, it doesn't mean that the estimation error can be assumed to be the same in real conditions. However, it is a good indicator, since the drive cycles used to calibrating the selected parameters are meant to be representative of real driving situations.

A limitation of the present work is that the simulated fleets do not include neither a BEV nor a FCEV, despite the ability of the framework to consider this kind of vehicles. This limitation is directly related to the unavailability of such vehicles from the industry partners that provided the real drive and fleet data for the present work.

Mercedes Pro validation results

The results of the validation using the Mercedes Pro consumption rates are the most consistent assessment of the estimation error using the consumption framework, because they are the only validation with real-world data. The estimation error is around 5 %, which is in the same order of magnitude stated in the model documentation. If we look to the results with the most likely payload mass, the NME is only a −3.45 %, which is something remarkable. It should be kept in mind, that the scope of FASTSim is to find a balance between the simulation effort and the estimation accuracy, as stated in its validations report [60]. To achieve better results than that would require an extensive calibration effort that would make the framework unable to handle as many different vehicles and technologies as needed.

The validation of the results, however, is only possible for one of the virtual vehicles, the Mercedes Sprinter I, since is the only one that has the real rates data. Moreover, there is an extra source of uncertainty originated from the lack of precise payload mass. We do have a distribution of payload masses, but there is no way to match the payloads to the respective trips. The comparative analysis conducted in this work shows that a 614 kg payload increased the fuel consumption a 10.2 %. For the particular case of the Mercedes Sprinter I, for which there is the most data available, the increase was a 10.6 %. With these figures, the impact of the payload on the validation can not be neglected. By segregating the validation in three different payloads, this uncertainty is diluted, but should still be considered. Additionally, there is also some uncertainty regarding the speed profile accuracy, as the recording device is not completely precise. This last source of error, however, can be neglected after setting a data quality threshold. The possible inaccuracies of the resulting trips are very unlikely to affect significantly the consumption rates as the trip distances match a 98 %.

Spritmonitor results

The validation using the Spritmonitor rates produces mixed results that vary extremely from one vehicle to another, as seen in Table 5.4. Thus, The Mercedes Vito, both Sprinters, the Mitsubishi Outlander and the Opel Movano present a very similar mean consumption rate using the framework, to the one recorded by the Spritmonitor users. This can be considered as very good results. A second group presents rather good results, this is, an absolute difference below 15 %. This is the case of the Hyundai i30, the Opel Astra, the Volkswagen Caddy and the Iveco 70C15 (which is actually slightly out of the range). Finally, the rest of the vehicle present bad or

very bad results, especially for the Toyota Prius HEV and the Iveco 50C15. In order to assess the results there are many aspects to consider. First, comparing crowd-sourced consumption data with simulation data will inevitably lead to deviations, since it is not possible to isolate some relevant parameters such as the influence of the driving behavior, the type of trips or the influence of a variable payload. Moreover, some vehicle models present dozens of variants and the filter of Spritmonitor may not be able to deal with it properly. This is particulary true for the transporters and the Iveco fleet, in which it was only possible to select the 70C or 50C line, without further specification of the specific version. Finally, despite the immensity of the Spritmonitor database, it does not have an equally distributed amount of rates along with all existing models. As one could expect, it has much more information about the most popular vehicles, whereas the rarer vehicles are hardly represented. This is clearly represented by the column 'Users' of the Table 5.4. There are 561 users that have an Opel Astra similar to the one that we are simulating. Thus, we should conclude that the Spritmonitor rates are reliable. On the other hand, only 6 users posted rates from an Opel Movano (or a MAN TGE) having similar properties that the ones in our virtual fleet. Therefore their rates should be used with caution. On the simulated results side, the same logic applies. The results of the Sprinters, which are simulated using hundreds of daily trips, are much more representative than the results of the Volkswagen Caddy, in which there were only 14 daily trips available.

Keeping all the above mentioned in mind, the minimum and maximum recorded rates from Spritmonitor are a very useful piece of information, since it allows to distinguish whether a high percentage of difference comes from an inaccurate simulation or it may be just a consequence of the above-mentioned sources of error. None of the vehicles registered a mean consumption above the maximum threshold, but 4 of them do have lower mean consumption than the minimum Spritmonitor value. Those vehicles are the Toyota Prius, the Iveco 70C15, the Peugeot Boxer and the MAN TGE. While the three transporters have a relatively low amount of data to compare with,s o its results should be treated carefully, for the Toyota Prius there is enough data to conclude that the simulation is underestimating the real consumption of this vehicle. This can be due to the fact that it was calibrated using a drive cycle that wasn't representative of the real drive conditions, however, this would also apply to the other vehicles that do have a satisfactory estimation rate. More precisely, the Mitsubishi Outlander is also calibrated using the same FASTSim default procedure and shows an almost perfect mean consumption rate. In short, it is needed a deeper analysis with other HEV to understand the origin of such low performance of the Toyota Prius. All the other vehicles, apart from the 4 that have just been mentioned, are within the expected consumption range and therefore it can be assumed that they are able to satisfactorily estimate the fuel consumption in real conditions within an acceptable error range.

The error chain of the simulation framework using recorded trips

The error chain presented in Figure 5.21 summarizes the results achieved up to this point. The three sources of error are all assessed and despite the validation uncertainty, the framework has been proved to work and provide satisfactory results. The distance error is well assessed by controlling the data quality and it can be circumvented by assuming that it doesn't affect the consumption rate, but just the overall consumption value. Thus, this error can be reduced the majority of the time.

The calibration error is much more difficult to solve since the low availability of vehicle data is one of the hardest issues that the authors face when developing a consumption model. For this reason, some of the models presented in the literature review are entirely calibrated with publicly available data, such as VT-CPFM and VT-CPEM. FASTSim can also be calibrated with public available data for passengers cars sold in the US, however, the available data of the transporters is

harder to find and can be even be missing, as of for the Ivecos. In absences of the real parameters, FASTSim provides default values, but this of course results in an increased error.

The simulation error, closely related to the model assumptions, is both related to the necessary simplification that facilitates the model calibration and also to the necessity of acceptable time performances. The validation report of FASTSim explicitly announces that there is a special willingness of designing a model in the "sweet spot" between accuracy and complexity [60, p. 3], so a moderate simulation error has to be expected.

Whereas the validation uncertainty, it is meant to put in context the validation results given the limitations of this work. The Mercedes Pro validation, that despite including only a vehicle type is the most accurate comparison set available, could not be completely leverage due to the unknown payloads at each trip. This is solved by simulating three different payloads that are representative of the overall payload distribution that was provided. However, this is only a shortcut that partially undermines the potential of the validation. To put in context the magnitude of the problem, in this same work it has been shown that a 614 kg payload increases an average of 10.0 % the consumption rate of a transporter. This same problem occurs when using the Spritmonitor values since there is not a way to determine the average payload that lead to the crowd-sourced consumption rates. The payload mass analysis also showed that the influence of the extra mass is greater in the lightest transporters, as the Caddy, which has a 18.2 % increase, compared to the heaviest, such of both Iveco transporters.

Despite the mentioned problems, the consumption framework provides a valuable tool for estimating the fuel consumption of a wide range of vehicles in a large-scale simulation.

Consumption estimation of non-recorded trips

The consumption estimates using non-recorded trips have many interesting takeaways. As seen in the literature review of Section 2.4, a solid approach for creating synthetic speed profiles on a large scale has, to the best of our knowledge, not yet been created. The literature in this field focuses on the short-term predictions rather than predicting a complete speed profile of a trip. A reason for this is probably the computational complexity, the difficulty of creating accurate predictions and the availability of real data, which makes a hypothetical effort for creating a large-scale speed profile predictor unlikely to pay off. The existing methodologies for doing something similar require proprietary software and are limited to relatively small and limited areas. A shortcut for this could be using an Artificial Intelligence approach. The complexity of such procedure makes it nonviable for the present work.

The main alternative for simulating the consumption of trips without having its speed profile relies on a regression curve of the vehicle consumption enhanced with road and vehicle-related attributes. Note that in this case the speed profile is no longer relevant, thus, it doesn't make sense anymore to speak of synthetic speed profiles. These approaches, however, have indeed the capability of estimating the fuel consumption of trips that have not been recorded, which is exactly one of the objectives of the project.

The approach of the present work bases on the three use cases, defined as data-driven approaches, that involve a digital map from which to extract road relevant features and a consumption model from which to create a regression. This makes it relatively simple and flexible but has the disadvantage that it doesn't capture specific driving behavior, but just an average. This is a fact to keep in mind since it limits the capability of this approach to estimate individual trips with precision.

In this regard, the training data obtainment and the selected speed bins have a major impact. The speed bins of 15 km/h balance the capability of capturing several different consumption situations (characterized by the link mean speed) without over-fitting. The alternative speed bin combinations also presented in the results recorded poorer validation results. Still, the selected configuration has a lack of training data for the first bin. This is due to the fact that on seldom occasions the tracks have been recorded in such low speed links with the necessary data quality. Therefore, the consumption rate at such bin is not well averaged and it is unlikely to capture properly the driving behavior. However, since this bin is rarely found, this inaccuracy will have a very low impact on the overall trip consumption, especially on long distances.

The resulting consumption curves present different patterns depending on the vehicle type. In general, the transporters are high consumers at all bins and particularly at high speeds due to the aerodynamic drag. The passenger cars are less consuming and less affected by high speeds. In both cases, the consumption is also high at low speeds due to the accelerations and frequent stops typically found in urban areas, which is a highly inefficient driving behavior. The hybrid vehicles are able to avoid these inefficiencies thanks to their electric motors and present very low or even no consumption in these situations.

For the validation of the results a problem with an irregular distribution of the available data arises. There are thousands of available trips from Mercedes Sprinters, while for other vehicles there isn't even a trip with enough data quality for the validation. The main reason for this is the number of vehicles since 13 out of the 20 vehicles of Fleet 1 are Sprinters. However, there are other factors, such as the routes, which in case of being mostly urban may contain several gaps, a short period of experimentation and more. In aggregate, the validation is done using almost 5000 tracks with a total length of 50 000 km. Thus, the resulting parameters can indeed be assumed as representative.

The validation shows very promising indicators for most vehicles and also for the aggregated results. The 12.9 % of MAE is comparable with the results of Holden [45], where they achieved NTAE between 12.5 % and 17 %. Regarding the work of Boriboonsomsin [42], they expressed their results in terms of ME and achieved underestimations of around −15 %, worse results than the ones achieved in the present work. Only the Movanos, which is by far the worst vehicle regarding its validation results is close to these results. On the other hand, the TUM's approach for estimating BEV energy consumption by Grubwinkler achieved a MAE of 6.9 %, far better than in the present work. The reason for this is the fact that our approach doesn't account for the driver behavior, which produces the averaging error mentioned in the error chain. In contrast, the work from Grubwinkler did manage to account for the driving individual behavior.

The exception in the satisfactory results is the Mitsubishi Outlander, the PHEV for which the model produced inconsistent results. This was caused by the complexity of this technology, for which the EMS is crucial, as seen in Section 2.5. FASTSim determine the energy management mode at each instant depending on the SOC of the vehicle. When the SOC is sufficiently high, the vehicles operate under CD mode, while when the SOC diminishes to a certain value, the model shifts to CS mode. In regular HEV, this problem is solved by forcing the vehicle to operate under CS mode. It is a rough assumption but works well as the batteries are relatively small and can not be recharged other than using the energy recuperation system. However, since PHEV have a significant purely electric range and can be charged using the electric grid, their consumption rate is no more function of the average speed but of the SOC of its batteries and the length of the trip. Thus, the approach for estimating the consumption without using GPS data is not valid.

Additionally, the ability to capture the extra mass on a transporter vehicle has been proved, so the model can simulate variable payloads keeping the error constant or even slightly improving it. This is a relevant feature provided the non-negligible impact that payloads have on the fuel consumption of a transporter, which has already been discussed.

Overall, the estimation of non-recorded trips offers good results that are comparable with the literature and far better than just considering the homologation rate, which would produce a MAE around 26 % and could not account for the extra payload. The results, however, are far better when performing batch simulations of a fleet or a relatively long period of time, than when simulating individual trips under particular circumstances. Since the objective of the work is to create a consumption framework to be used in a fleet optimization model considering its mobility behavior, the attributes of the presented framework are perfectly aligned to this needs.

7 Summary and Outlook

This thesis has fulfilled the objective of developing and validating a consumption simulation framework adapted to a mixed fleet. The framework provides a complete and functional solution for simulating ICEV, HEV, PHEV, BEV and FCEV using a speed profile as input data. The trip generation using GPS recordings of a large scale fleet test is also included in the framework and even considers the road grade. The results have been validated using two sources, the real consumption rates gathered by the Mercedes Pro platform, that where available for a subset of trips and vehicles, and the publicly available Spritmonitor rates. Despite the uncertainty of the exact payload mass at each trip, the Mercedes Pro validation shows a high level of estimation accuracy that outperforms by an order of magnitude the accuracy achieved by simply considering the homologation consumption rate. The Spritmonitor validation is subject to many more uncertainties, but still offers good correlations with the estimated consumption for most of the vehicles. In some cases the results are either inconclusive or too distinct, so more research in this aspect is necessary. Moreover, the available fleets didn't include neither BEV nor a FCEV, so a further validation of the results with such vehicles is needed.

A comparative analysis of the influence of payloads in the overall fuel consumption has been done, showing that a 600 kg can increase the consumption of a Diesel vehicle more than 10 %. Again, due to the limitations of the fleet, this analysis only includes one vehicle typology despite the ability of the framework to include other alternatives. Thus, the author would like to encourage future works focusing on the impact of the payload in electric transporters.

The framework also includes a methodology to predict the fuel consumption of trips that have not been recorded, or even not driven at all. This function just needs the coordinates of an origin and destination to generate the most likely route between these points and estimate the consumption rate of the route between them. This is achieved by defining a series of speed bins for which an average speed profile is created using the available GPS data. Then, FASTSim calculates the average consumption rate at each bin and creates the so-called consumption curve, which is characteristic for each vehicle. This consumption curve is used together with a routing engine to predict the fuel consumption of any trip. This approach also accounts for the additional payloads. The validation showed high accuracy results that were well positioned within the literature, however the fact of using an average speed profile induced a high variance in the estimations, which should be considered when performing individual simulations. The validation only included ICEV and HEV due to the above mentioned limitations, and also showed that this approach is unable to estimate PHEV. The reason for this is the huge influence of the EMS in the consumption of PHEV. More research is needed in order to find a functional solution for the PHEV estimation as well as for validating the framework with the missing vehicle powertrain technologies.

In this regard, this thesis also includes a short introduction on the generation of synthetic speed profiles, that could improve the accuracy results of the framework for non recorded trips and even include PHEV. Using the average speed of the links provided by the routing engine, a rough

speed profile of the trip is created and directly simulated using FASTSim. The results of this approach were substantially worse than the ones using the consumption curves, as the used speed profile ignores the accelerations, that are particularly important in urban driving. Therefore, this approach underestimated the fuel consumption. However, the author suspects great potential in this approach when combined with a machine learning solution that emulates a more realistic acceleration behavior. Consequently, more research in this field is encouraged.

List of Figures

List of Tables

Bibliography

[1] IEA, „Tracking Transport 2020," IEA, Paris, 2020. Available: https://www.iea.org/reports/tracking-transport-2020 [visited on 11/26/2020].

[2] T. a. Environment, „How European transport can contribute to an EU -55% GHG emissions target in 2030," 2020 European Federation forTransport and Environment AISBL, Brussels and Belgium, 2020.

[3] United Nations. „Adoption of the Paris Agreement," 2021. [Online]. Available: http://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf [visited on 02/16/2021].

[4] A. Grauers, S. Sarasini and M. Karlström, „Why electromobility and what is it?," *91980973*, 2013.

[5] K. Ahn et al., „Estimating Vehicle Fuel Consumption and Emissions based on Instantaneous Speed and Acceleration Levels," *Journal of Transportation Engineering*, vol. 128, no. 2, pp. 182–190, 2002, DOI: 10.1061/(ASCE)0733-947X(2002)128:2(182).

[6] Z. A. Needell and J. E. Trancik, „Efficiently Simulating Personal Vehicle Energy Consumption in Mesoscopic Transport Models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 25, pp. 163–173, 2018, DOI: 10.1177/0361198118798244.

[7] B. T. Morris et al., „Real-Time Video-Based Traffic Measurement and Visualization System for Energy/Emissions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1667–1678, 2012, DOI: 10.1109/tits.2012.2208222.

[8] H. G. R. Kerali, J. B. Odoki and E. E. Stannard, „Overview of HDM-4," *The highway development and management series, Volume one, World Road Association, PIARC. World Bank, Washington DC, USA*, 2000.

[9] W. F. Faris et al., „Vehicle fuel consumption and emission modelling: an in-depth literature review," *Int. J. Vehicle Systems Modelling and Testing*, vol. 6, no. 3/4, 2011.

[10] H. Yue, „Mesoscopic fuel consumption and emission modeling," Tech, Virginia, 2008.

[11] Y. Chen et al., „Data-driven fuel consumption estimation: A multivariate adaptive regression spline approach," *Transportation Research Part C: Emerging Technologies*, vol. 83, pp. 134–145, 2017, DOI: 10.1016/j.trc.2017.08.003.

[12] M. Zhou, H. Jin and W. Wang, „A review of vehicle fuel consumption models to evaluate eco-driving and eco-routing," *Transportation Research Part D: Transport and Environment*, vol. 49, pp. 203–218, 2016, DOI: 10.1016/j.trd.2016.09.008.

[13] B. Othman et al., „Ecological traffic management: A review of the modeling and control strategies for improving environmental sustainability of road transportation," *Annual Reviews in Control*, vol. 48, pp. 292–311, 2019, DOI: 10.1016/j.arcontrol.2019.09.003.

[14] J. Huertas et al., „Driving Cycles Based on Fuel Consumption," *Energies*, vol. 11, no. 11, p. 3064, 2018, DOI: 10.3390/en11113064. Available: https://www.mdpi.com/1996-1073/11/11/3064.

[15] J. L. Jimenez-Palacios, „Understanding and quantifying motor vehicle emissions with vehicle specific power and TILDAS remote sensing," Ph. D. Dept. of Mechanical Engineering, Massachussets Institute of Technology, 1999.

[16] G. O. Duarte, G. A. Gonçalves and T. L. Farias, „A methodology to estimate real-world vehicle fuel use and emissions based on certification cycle data," *Procedia-Social and Behavioral Sciences*, vol. 111, pp. 702–710, 2014.

[17] A. Sciarretta, G. De Nunzio and L. L. Ojeda, „Optimal Ecodriving Control: Energy-Efficient Driving of Road Vehicles as an Optimal Control Problem," *IEEE Control Systems Magazine*, vol. 35, no. 5, pp. 71–90, 2015, DOI: 10.1109/MCS.2015.2449688.

[18] L. Ntziachristos et al., „COPERT: a European road transport emission inventory model," in *Information technologies in environmental engineering* Springer, 2009, pp. 491–504.

[19] „. *COPERT Micro: a tool to calculate the vehicle emissions in urban areas*," 2014.

[20] CARB. „*EMFAC2017 User's guide, v1.0.2*," 2018.

[21] United States Environmental Protection Agency. „MOVES and other mobile source emissions models," 2020. [Online]. Available: https://www.epa.gov/moves [visited on 12/01/2020].

[22] Z. Zhang et al., „Development of a Simplified Model of Speed-Specific Vehicle-Specific Power Distribution Based on Vehicle Weight for Fuel Consumption Estimates," *Transportation Research Record: Journal of the Transportation Research Board*, p. 036119812094741, 2020, DOI: 10.1177/0361198120947415.

[23] Z. Wang, G. Wu and G. Scora, „MOVESTAR: An open-source vehicle fuel and emission model based on usepa moves," *arXiv preprint arXiv:2008.04986*, 2020.

[24] Z. Zhai et al., „Capturing the Variability in Instantaneous Vehicle Emissions Based on Field Test Data," *Atmosphere*, vol. 11, no. 7, p. 765, 2020, DOI: 10.3390/atmos11070765.

[25] AFRAS. „Handbook of Emission Factors for Road Transport," 2021. [Online]. Available: https://www.hbefa.net/e/index.html [visited on 03/30/2021].

[26] T. Markel et al., „ADVISOR: a systems analysis tool for advanced vehicle modeling," *Journal of Power Sources*, vol. 110, no. 2, pp. 255–266, 2002, DOI: 10.1016/s0378-7753(02)00189-1.

[27] CERT. „Comprehensive Modal Emissions Model," 2021. [Online]. Available: https://www.cert.ucr.edu/cmem [visited on 02/21/2021].

[28] G. Scora and M. Barth, „Comprehensive modal emissions model (cmem), version 3.01," 2006.

[29] R. G. Dowling, *Predicting Air Quality Effects of Traffic-flow Improvements: Final Report*, Transportation Research Board, 2005, ISBN: 0309088194.

[30] Z. Kan et al., „Estimating Vehicle Fuel Consumption and Emissions Using GPS Big Data," *International Journal of Environmental Research and Public Health*, vol. 15, no. 4, p. 566, 2018, DOI: 10.3390/ijerph15040566.

[31] H. A. Rakha et al., „Virginia Tech Comprehensive Power-Based Fuel Consumption Model: Model development and testing," *Transportation Research Part D: Transport and Environment*, vol. 16, no. 7, pp. 492–503, 2011, DOI: 10.1016/j.trd.2011.05.008.

[32] S. Park et al., „Virginia Tech Comprehensive Power-based Fuel Consumption Model (VT-CPFM): Model Validation and Calibration Considerations," *International Journal of Transportation Science and Technology*, vol. 2, no. 4, pp. 317–336, 2013, DOI: 10.1260/2046-0430.2.4.317.

[33] C. Fiori, K. Ahn and H. A. Rakha, „Power-based electric vehicle energy consumption model: Model development and validation," *Applied Energy*, vol. 168, pp. 257–268, 2016, DOI: 10.1016/j.apenergy.2016.01.097.

[34] A. Brooker et al., „FASTSim: A Model to Estimate Vehicle Efficiency, Cost and Performance," 2015, DOI: 10.4271/2015-01-0973.

[35] „. *Simplified electric vehicle power train models and range estimation*," IEEE, 2011. DOI: 10.1109/vppc.2011.6043163.

[36] J. G. Hayes and K. Davis, „Simplified electric vehicle powertrain model for range and energy consumption based on EPA coast-down parameters and test validation by Argonne National Lab data on the Nissan Leaf," in *2014 IEEE Transportation Electrification Conference and Expo (ITEC)*, pp. 1–6, ISBN: 1479922625.

[37] TU Graz. „PHEM Light," 2021. [Online]. Available: https://www.ivt.tugraz.at/assets/files/download/PHEMlight_en.pdf [visited on 03/30/2021].

[38] TU Graz. „PHEM," 2021. [Online]. Available: https://www.ivt.tugraz.at/assets/files/download/PHEM_en.pdf [visited on 03/30/2021].

[39] J. M. Lohrer, „Datenbasierte Methoden zur Prädiktion des Trips und des Fahrprofils," Dissertation, Fakultät für Maschinenwesen, München, Technische Universität, 2018.

[40] „. *A new method to calculate emissions with simulated traffic conditions*," IEEE, 2010. DOI: 10.1109/itsc.2010.5625030.

[41] D. Karbowski, S. Pagerit and A. Calkins, „Energy Consumption Prediction of a Vehicle along a User-Specified Real-World Trip," *World Electric Vehicle Journal*, vol. 5, no. 4, pp. 1109–1120, 2012, DOI: 10.3390/wevj5041109. Available: https://www.researchgate.net/publication/267261936_Energy_Consumption_Prediction_of_a_Vehicle_along_a_User-_Specified_Real-World_Trip.

[42] K. Boriboonsomsin et al., „Eco-Routing Navigation System Based on Multisource Historical and Real-Time Traffic Information," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1694–1704, 2012, DOI: 10.1109/tits.2012.2204051.

[43] S. Grubwinkler and M. Lienkamp, eds. „*A modular and dynamic approach to predict the energy consumption of electric vehicles*," 2013.

[44] S. Grubwinkler, T. Brunner and M. Lienkamp, „Range Prediction for EVs via Crowd-Sourcing," in *2014 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2014, pp. 1–6, ISBN: 1938-8756. DOI: 10.1109/VPPC.2014.7007121.

[45] J. Holden et al., „Development of a Trip Energy Estimation Model Using Real-World Global Positioning System Driving Data," National Renewable Energy Lab.(NREL), Golden, CO, 2017.

[46] A. Rousseau, S. Pagerit and D. W. Gao, „Plug-in hybrid electric vehicle control strategy parameter optimization," *Journal of Asian Electric Vehicles*, vol. 6, no. 2, pp. 1125–1133, 2008.

[47] O. Karabasoglu and J. Michalek, „Influence of driving patterns on life cycle cost and emissions of hybrid and plug-in electric vehicle powertrains," *Energy Policy*, vol. 60, pp. 445–461, 2013, DOI: 10.1016/j.enpol.2013.03.047. Available: https://www.sciencedirect.com/science/article/pii/S0301421513002255.

[48] A. Taghavipour et al., „A comparative analysis of route–based energy management systems for PHEVs," *Asian Journal of Control*, vol. 18, no. 1, pp. 29–39, 2016.

[49] C. M. Martinez et al., „Energy management in plug-in hybrid electric vehicles: Recent progress and a connected vehicles perspective," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 4534–4549, 2016.

[50] G. Fontaras, P. Pistikopoulos and Z. Samaras, „Experimental evaluation of hybrid vehicle fuel economy and pollutant emissions over real-world simulation driving cycles," *Atmospheric environment*, vol. 42, no. 18, pp. 4023–4035, 2008.

[51] C. Silva, M. Ross and T. Farias, „Evaluation of energy consumption, emissions and cost of plug-in hybrid vehicles," *Energy Conversion and Management*, vol. 50, no. 7, pp. 1635–1643, 2009.

[52] M. Wittmann et al., „A holistic framework for acquisition, processing and evaluation of vehicle fleet test data," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–7, ISBN: 1538615266.

[53] J. Wahlstrom et al., „Map-Aided Dead-Reckoning Using Only Measurements of Speed," *IEEE Transactions on Intelligent Vehicles*, vol. 1, no. 3, pp. 244–253, 2016, DOI: 10.1109/TIV.2017.2657383.

[54] Lin Zhao et al., „An extended Kalman filter algorithm for integrating GPS and low cost dead reckoning system data for vehicle performance and emissions monitoring," 2009. Available: https://repository.lboro.ac.uk/articles/journal_contribution/An_extended_Kalman_filter_algorithm_for_integrating_GPS_and_low_cost_dead_reckoning_system_data_for_vehicle_performance_and_emissions_monitoring/9441761.

[55] Fisch und Fischl. „Spritmonitor," 2021. [Online]. Available: https://www.spritmonitor.de/ [visited on 03/30/2021].

[56] L. Adenaw et al., „MAGIS – A Geographic Information System for Mobility Data Analysis," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 135–141, DOI: 10.1109/ITSC.2019.8917054.

[57] United States Environmental Protection Agency. „Fuel economy guide," 2021. [Online]. Available: https://www.fueleconomy.gov/feg/download.shtml [visited on 03/22/2021].

[58] European Union. „Vehicle Registration Certificate," 2021. [Online]. Available: https://ec.europa.eu/transport/road_safety/topics/vehicles/vehicle-registration-certificate/germany-vrc-part-1-and-2_en#part1 [visited on 03/30/2021].

[59] NZ Transport Agency. „Carbon emissions," [Online]. Available: https://rightcar.govt.nz/carbon-emissions.

[60] J. Gonder et al. „Future Automotive Systems Technology Simulator (FASTSim) Validation Report," 2021. [Online]. Available: https://www.nrel.gov/docs/fy18osti/71168.pdf [visited on 02/04/2021].

Appendix

A Appendix I

In this appendix the figures and tables that were not added in the main text are presented for documentation purposes. Still, the selection of figures and the presented data is not exhaustive and doesn't include all the plots, maps, calculations and simulations done during the elaboration of the thesis. It is a selection of the most relevant findings.

A.1 Additional virtual vehicles

As mentioned in the text, some vehicles didn't have an active part on the project, but were still created and calibrated. This is the case of the Skoda Fabia and the Hyundai Tucson, presented in Table A.1.

Table A.1: The vehicles in Fleet 1 that where finally not part of the project.

Vehicle Name	Number of vehicles	Powertrain type	Homologation Cycle	Error
Skoda Fabia	1	Gasoline	NEDC	$-0.20\,\%$
Hyundai Tucson	1	Diesel	NEDC	$11.8\,\%$

A.2 Parameters of the vehicles

In this section the final parameters of the virtual fleet creation and calibration are documented. The tables also include the two vehicles that have been not been finally considered in the thesis due to a lack of data.

Table A.2: Parameters of the Mercedes Sprinter CDI I.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	120 kW
Glider Mass	2090 kg
Frontal Area	4.234 m^2
Wheel Radius	0.336 m
Drag coefficient	0.39 (Default)
Rolling resistance R_f	0.008 (Default)
Efficiency improvement	-
Homologation consumption	7.93 l/100km

Table A.3: Parameters of the Mercedes Sprinter CDI II.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	120 kW
Glider Mass	2036 kg
Frontal Area	4.234 m^2
Wheel Radius	0.336 m
Drag coefficient	0.39 (Default)
Rolling resistance R_f	0.008 (Default)
Efficiency improvement	-
Homologation consumption	7.89 l/100km

Table A.4: Parameters of the Mercedes Vito Pro.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	100 kW
Glider Mass	2039 kg
Frontal Area	2.946 m^2
Wheel Radius	0.336 m
Drag coefficient	0.39 (Default)
Rolling resistance R_f	0.007 (Calibrated)
Efficiency improvement	-
Homologation consumption	7.89 l/100km

Table A.5: Parameters of the Opel Astra.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	100 kW
Glider Mass	1087 kg
Frontal Area	2.185 m^2
Wheel Radius	0.316 m
Drag coefficient	0.30 (Calibrated)
Rolling resistance R_f	0.007 (Default)
Efficiency improvement	-
Homologation consumption	4.84 l/100km

Table A.6: Parameters of the Volkswagen Caddy.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	77 kW
Glider Mass	1496 kg
Frontal Area	2.631 m^2
Wheel Radius	0.316 m
Drag coefficient	0.45 (Calibrated)
Rolling resistance R_f	0.010 (Default)
Efficiency improvement	-
Homologation consumption	6.38 l/100km

Table A.7: Parameters of the Hyundai Tucson, a vehicle for which real data was not available.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	101 kW
Glider Mass	1371 kg
Frontal Area	2.434 m^2
Wheel Radius	0.347 m
Drag coefficient	0.29 (Calibrated)
Rolling resistance R_f	0.068 (Calibrated)
Efficiency improvement	4 %
Homologation consumption	4.50 l/100km

Table A.8: Parameters of the Hyundai i30.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Gasoline engine
Engine Power	73 kW
Glider Mass	997 kg
Frontal Area	2.091 m^2
Wheel Radius	0.336 m
Drag coefficient	0.33 (Calibrated)
Rolling resistance R_f	0.012 (Default)
Efficiency improvement	-
Homologation consumption	6.00 l/100km

Table A.9: Parameters of the Skoda Fabia, a vehicle for which not real data was available.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Gasoline engine
Engine Power	77 kW
Glider Mass	810 kg
Frontal Area	1.967 m^2
Wheel Radius	0.336 m
Drag coefficient	0.316 (Real data)
Rolling resistance R_f	0.011 (Default)
Efficiency improvement	-
Homologation consumption	5.35 l/100km

Table A.10: Parameters of the Toyota Prius HEV.

Parameter	Description
Vehicle Type	HEV
Fuel Converter Type	Atkinson engine
Engine Power	71 kW
Glider Mass	1003 kg
Frontal Area	2.582 m^2
Wheel Radius	0.336 m
Drag coefficient	0.27 (Real data)
Rolling resistance R_f	0.006 (Calibrated)
Efficiency improvement	-
Homologation consumption	4.36 /100km
Electric Motor power	in 53 kW
Battery storage Capacity	in 1.3 kWh

Table A.11: Parameters of the Mitsubishi Outlander PHEV.

Parameter	Description
Vehicle Type	PHEV
Fuel Converter Type	Gasoline
Engine Power	99 kW
Glider Mass	1286 kg
Frontal Area	2.452 m^2
Wheel Radius	0.352 m
Drag coefficient	0.3 (Calibrated)
Rolling resistance R_f	0.007 (Calibrated)
Efficiency improvement	-
Homologation consumption	6.38 /100km
Electric Motor power	in 55 kW
Battery storage Capacity	in 13.8 kWh

Table A.12: Parameters of the Opel Movano.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	92 kW
Glider Mass	2204 kg
Frontal Area	4.661 m^2
Wheel Radius	0.349 m
Drag coefficient	0.39 (Default)
Rolling resistance R_f	0.008 (Default)
Efficiency improvement	-
Homologation consumption	7.97 l/100km

Table A.13: Parameters of the Iveco 70C15.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	110 kW
Glider Mass	3578 kg
Frontal Area	5.921 m^2
Wheel Radius	0.372 m
Drag coefficient	0.39 (Default)
Rolling resistance R_f	0.008 (Default)
Efficiency improvement	-
Homologation consumption	-

Table A.14: Parameters of the Iveco 50C15.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	110 kW
Glider Mass	2672 kg
Frontal Area	4.664 m^2
Wheel Radius	0.349 m
Drag coefficient	0.39 (Default)
Rolling resistance R_f	0.008 (Default)
Efficiency improvement	-
Homologation consumption	-

Table A.15: Parameters of the Ford Transit.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	92 kW
Glider Mass	2165 kg
Frontal Area	4.181 m^2
Wheel Radius	0.356 m
Drag coefficient	0.30 (Calibrated)
Rolling resistance R_f	0.007 (Calibrated)
Efficiency improvement	4 %
Homologation consumption	7.59 l/100km

Table A.16: Parameters of the Peugeot Boxer.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	96 kW
Glider Mass	1749 kg
Frontal Area	3.786 m^2
Wheel Radius	0.341 m
Drag coefficient	0.30 (Calibrated)
Rolling resistance R_f	0.007 (Calibrated)
Efficiency improvement	3 %
Homologation consumption	6.16 l/100km

Table A.17: Parameters of the MAN TGE.

Parameter	Description
Vehicle Type	Conventional
Fuel Converter Type	Diesel engine
Engine Power	96 kW
Glider Mass	2402 kg
Frontal Area	4.617 m^2
Wheel Radius	0.357 m
Drag coefficient	0.30 (Calibrated)
Rolling resistance R_f	0.007 (Calibrated)
Efficiency improvement	3 %
Homologation consumption	7.52 l/100km

A.3 Fleet test maps

To contextualize the extension of real data used for this work, following some maps are presented:

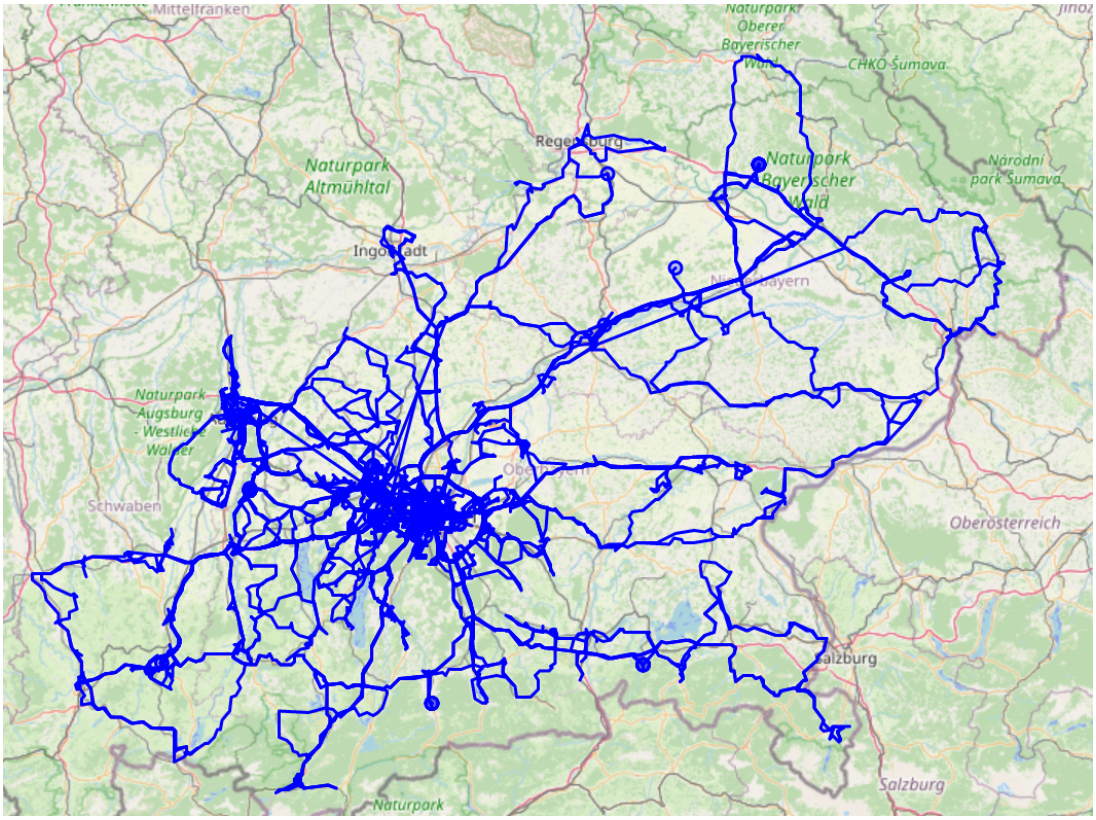


Figure A.1: Map of a selection of trips of Fleet 1 that were recorded during one month.

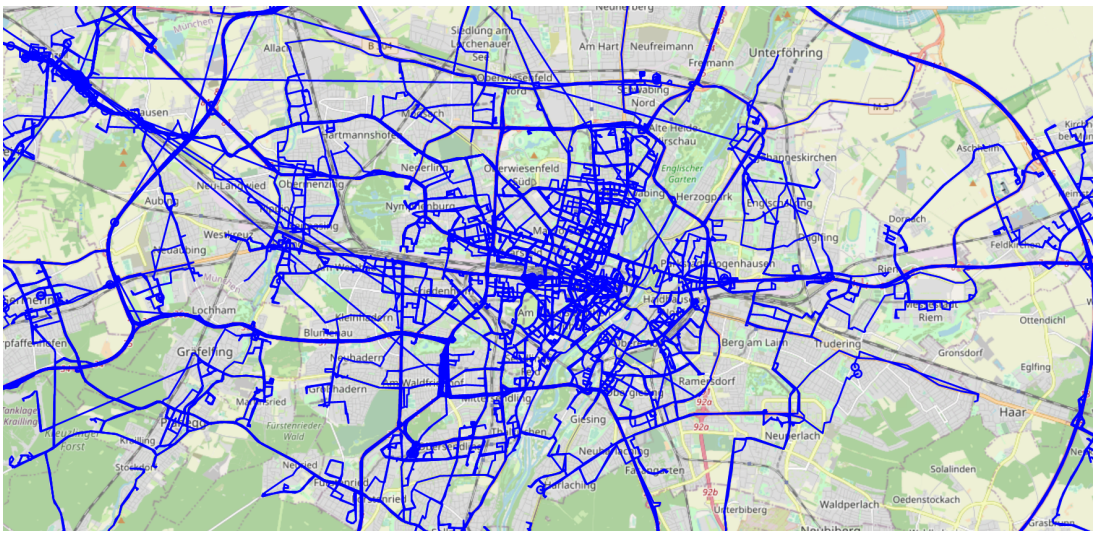


Figure A.2: Zoomed in segment of the map of Fleet 1, centered in the city of Munich.

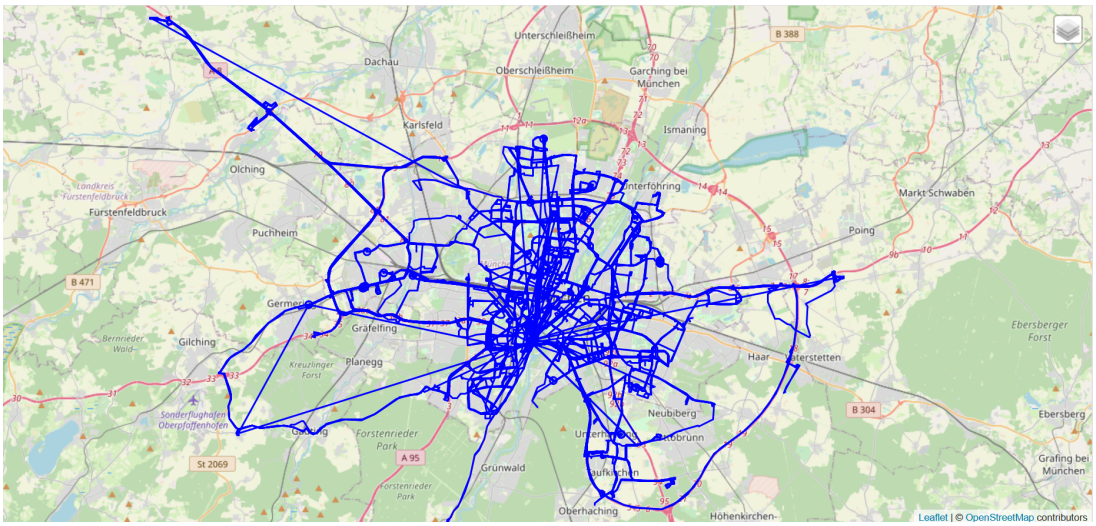


Figure A.3: Map of all the recorded trips of Fleet 2.

A.4 Speed profiles of each speed bin implementation

Each of the bin implementations has a set of speed profiles for all the speed bins. Following, this speed profiles are presented. In Figure A.4 an aleatory example is presented for explanatory purposes. The blue line represents the speed profile, which is usually a long concatenation of tracks. The red horizontal lines are the upper and lower limits of the presented speed bin and the green line is the average speed of the plotted speed profile.

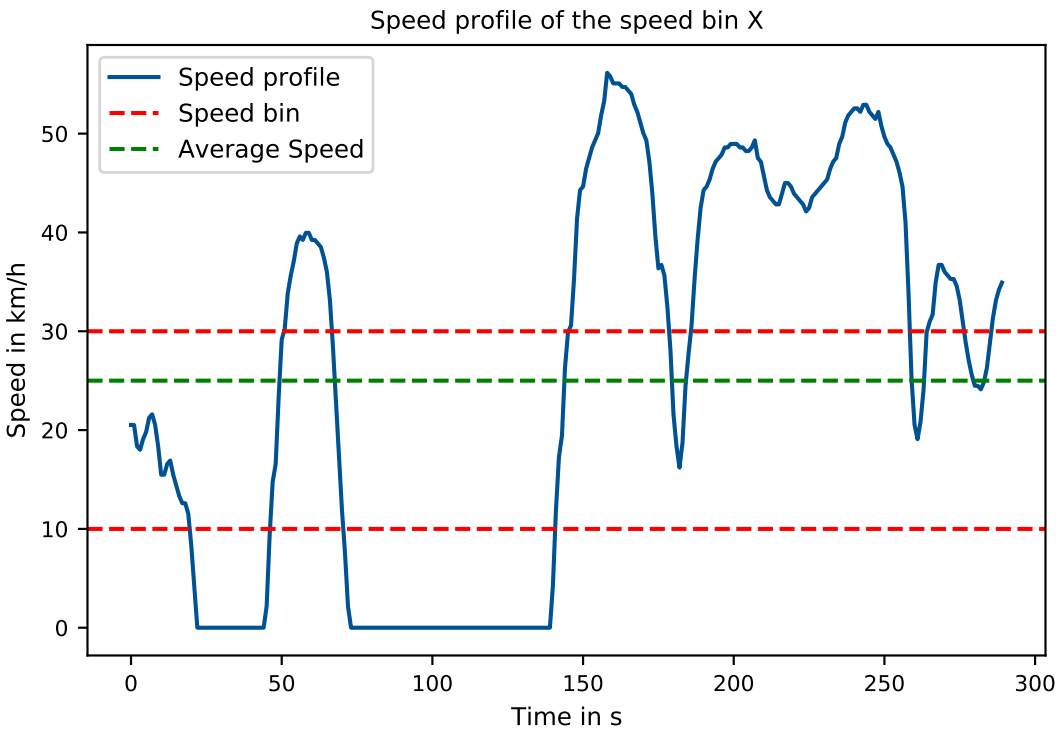


Figure A.4: Example of a speed profile for a given speed bin.

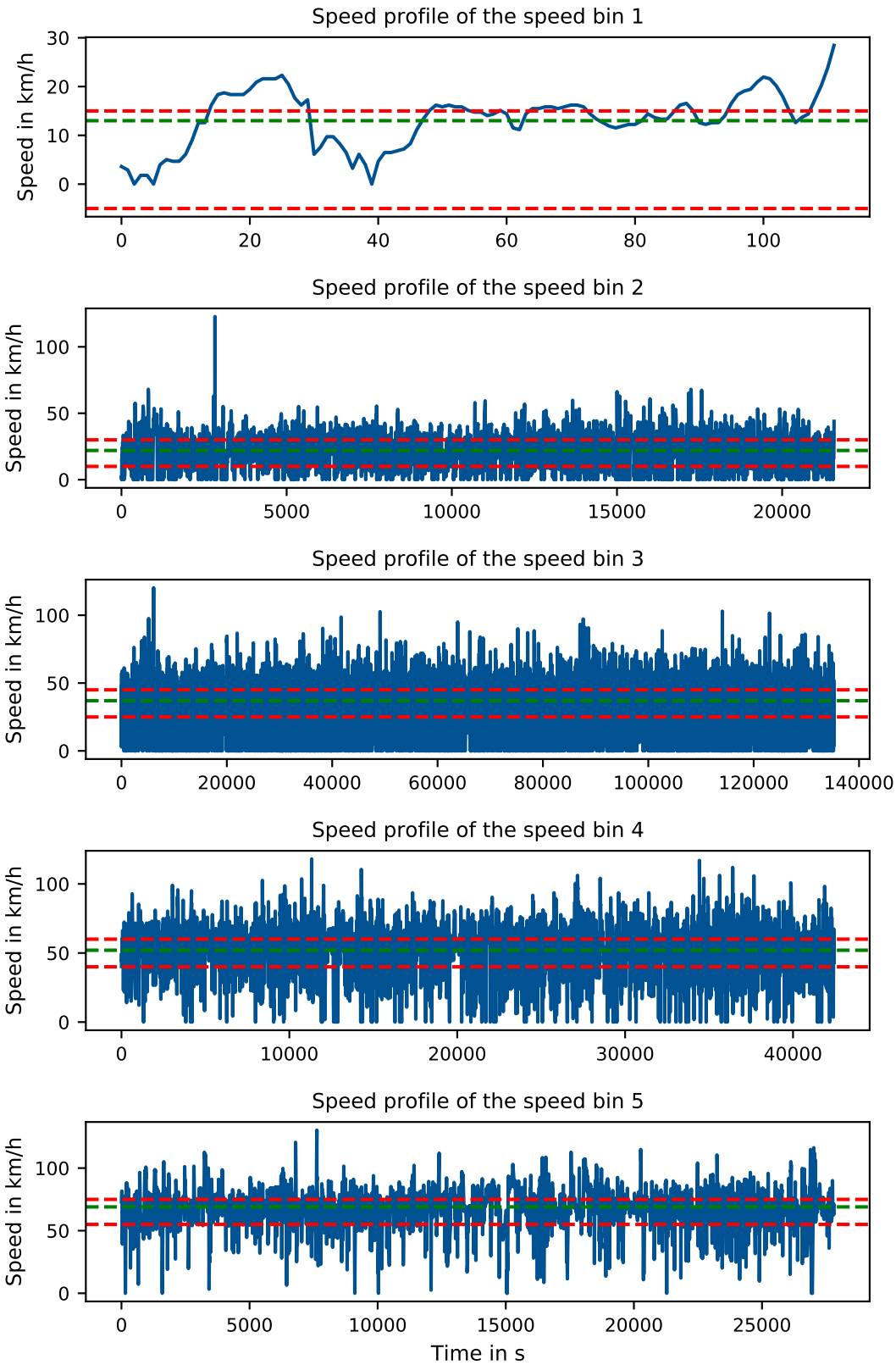


Figure A.5: Speed profiles from the implemented bin configuration: speed bins 1 to 5.

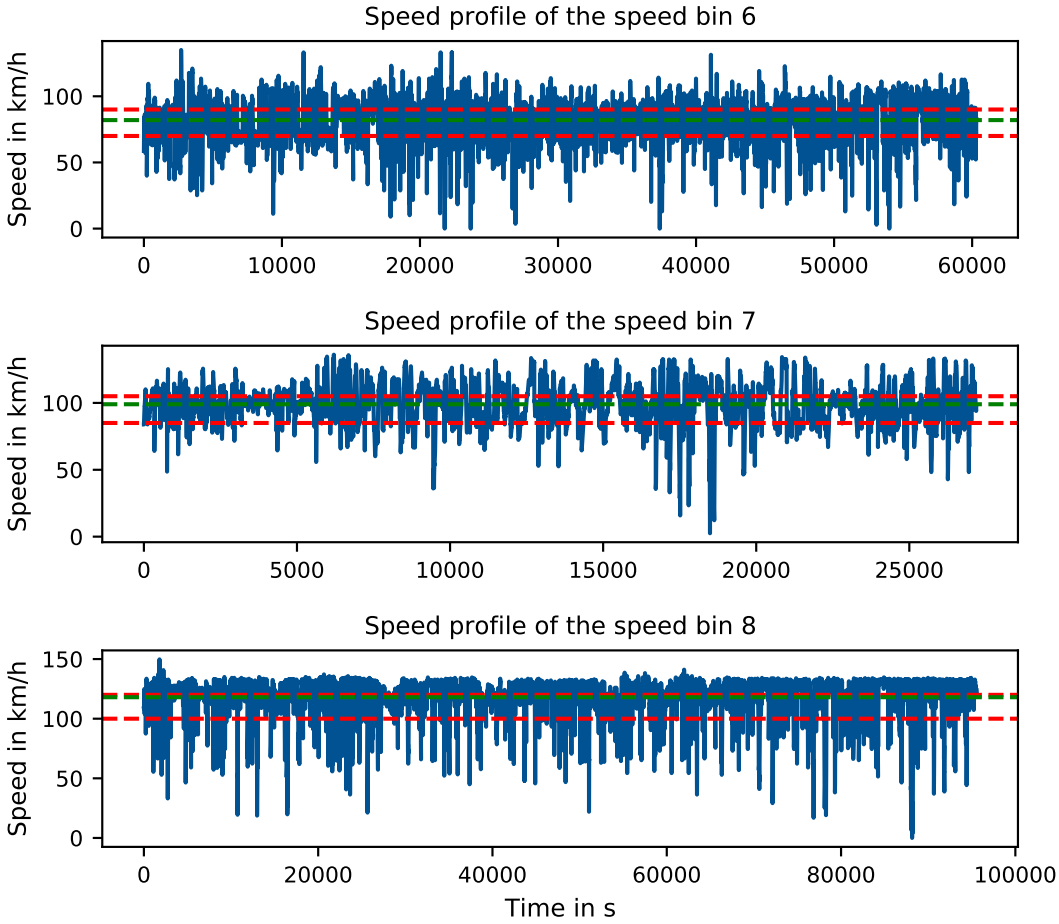


Figure A.6: Speed profiles from the implemented bin configuration: speed bins 6 to 8.

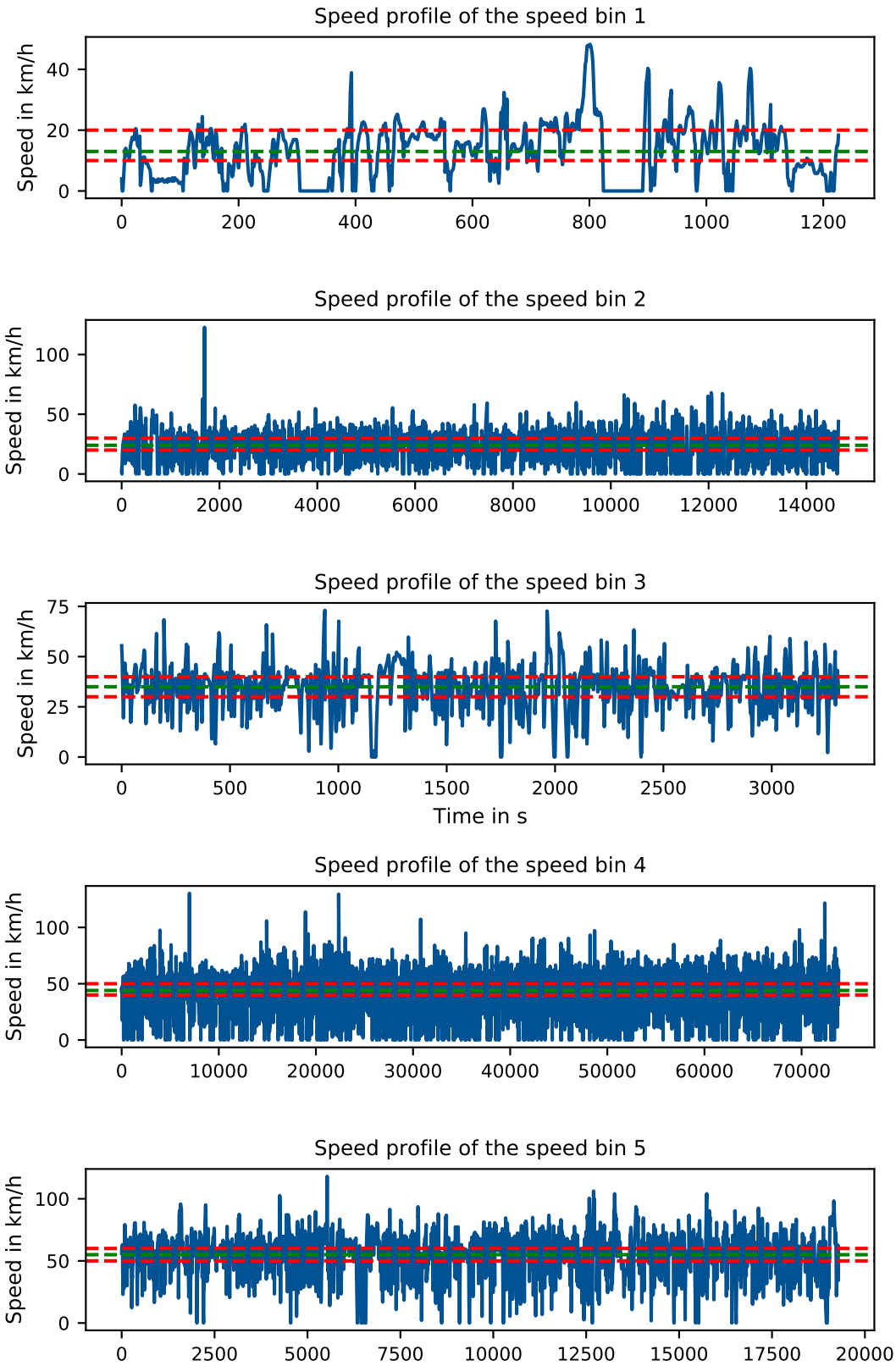


Figure A.7: Speed profiles from the speed bin configuration A: speed bins 1 to 5.

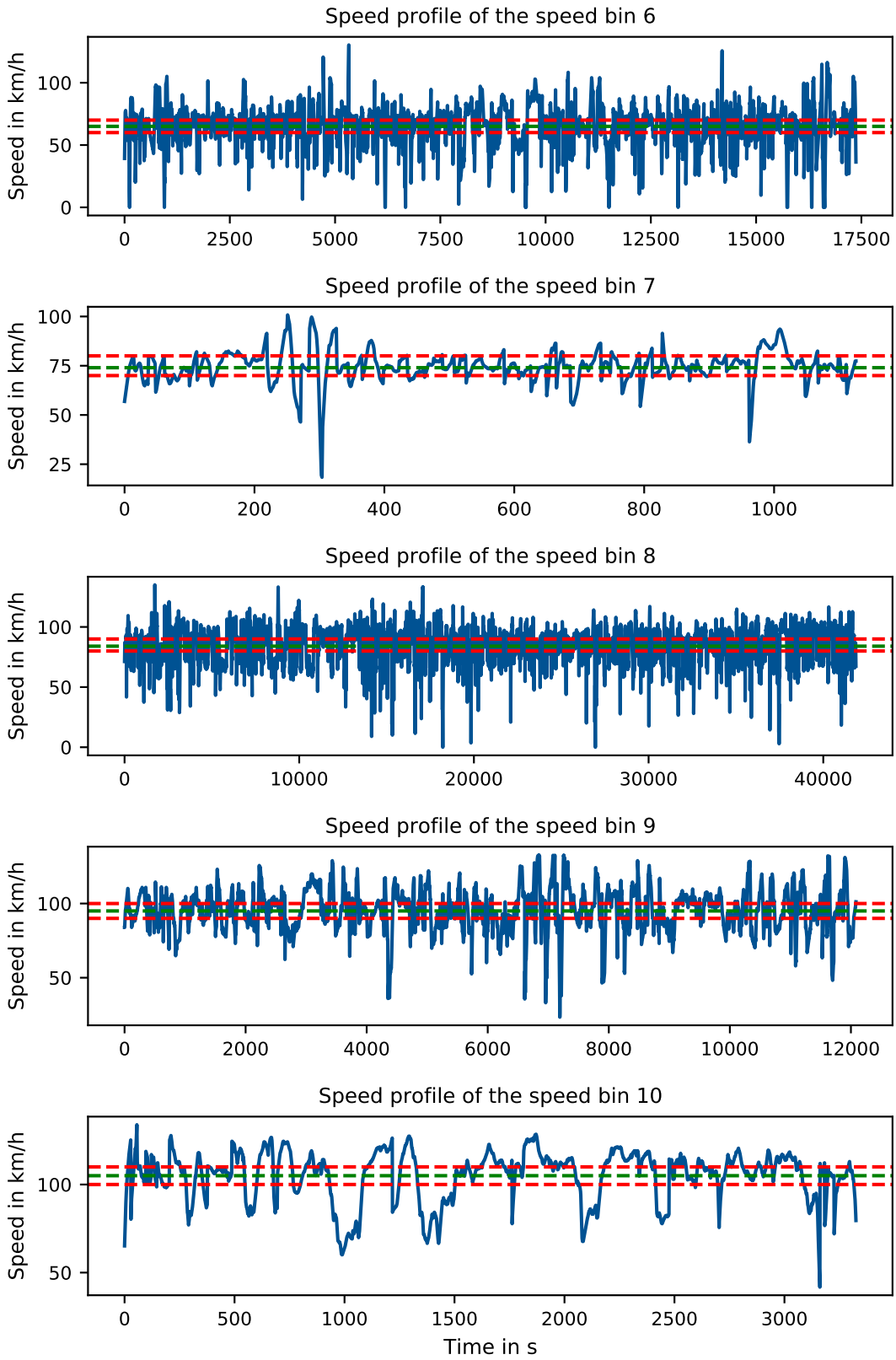


Figure A.8: Speed profiles from the speed bin configuration A: speed bins 6 to 10.

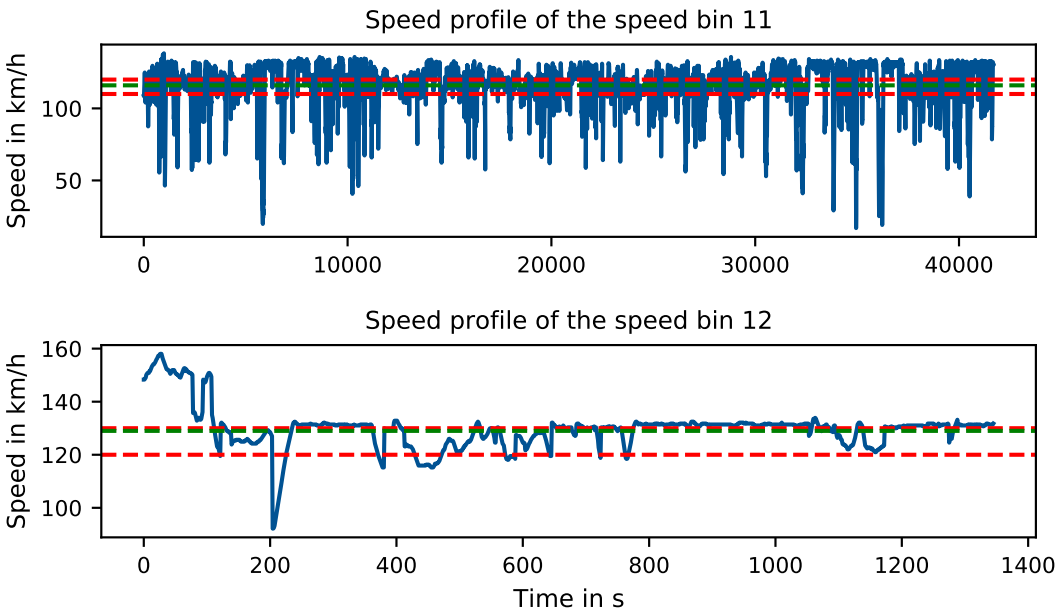


Figure A.9: Speed profiles from the speed bin configuration A: speed bins 11 and 12.

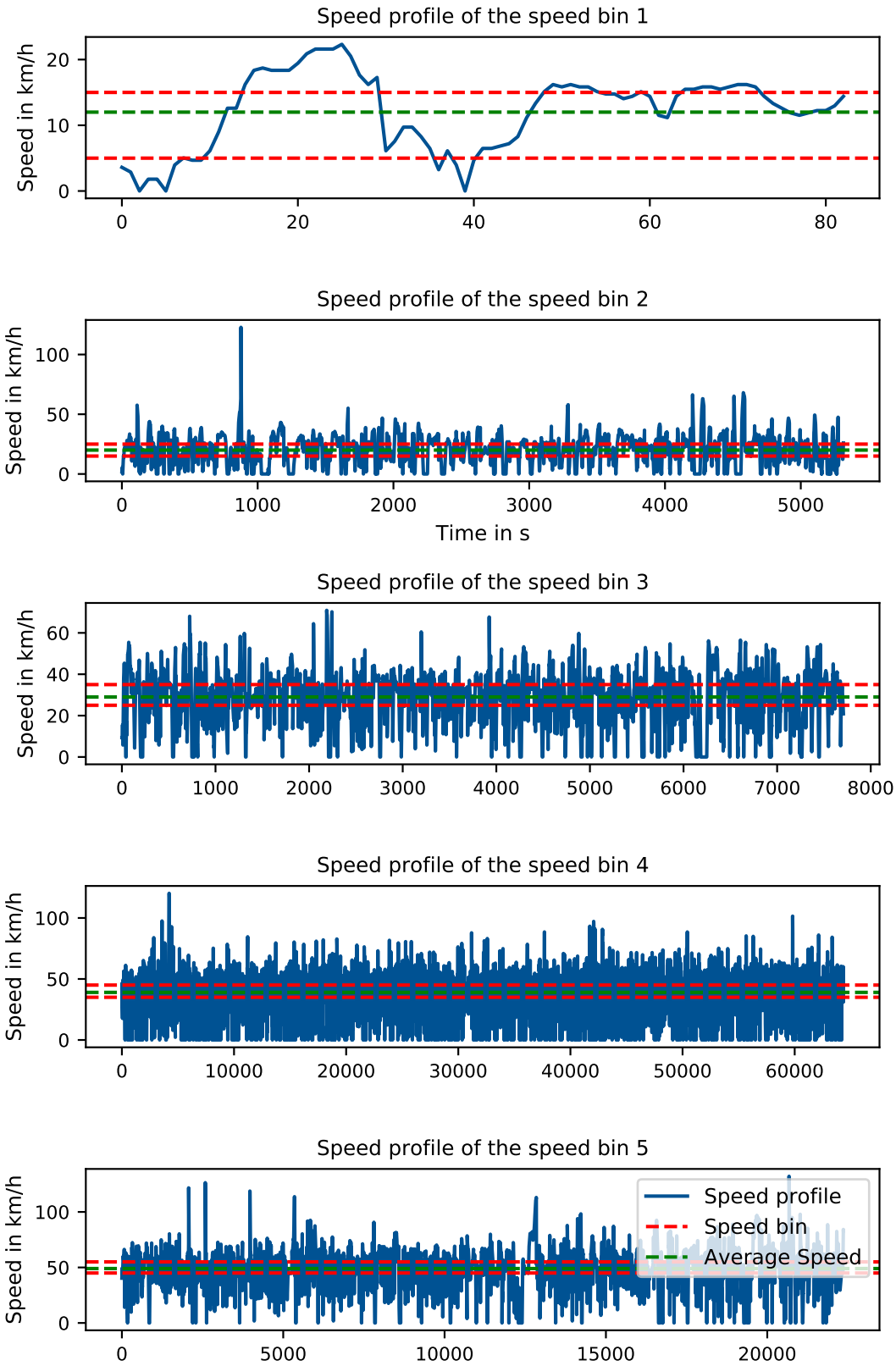


Figure A.10: Speed profiles from the speed bin configuration B: speed bins 1 to 5.

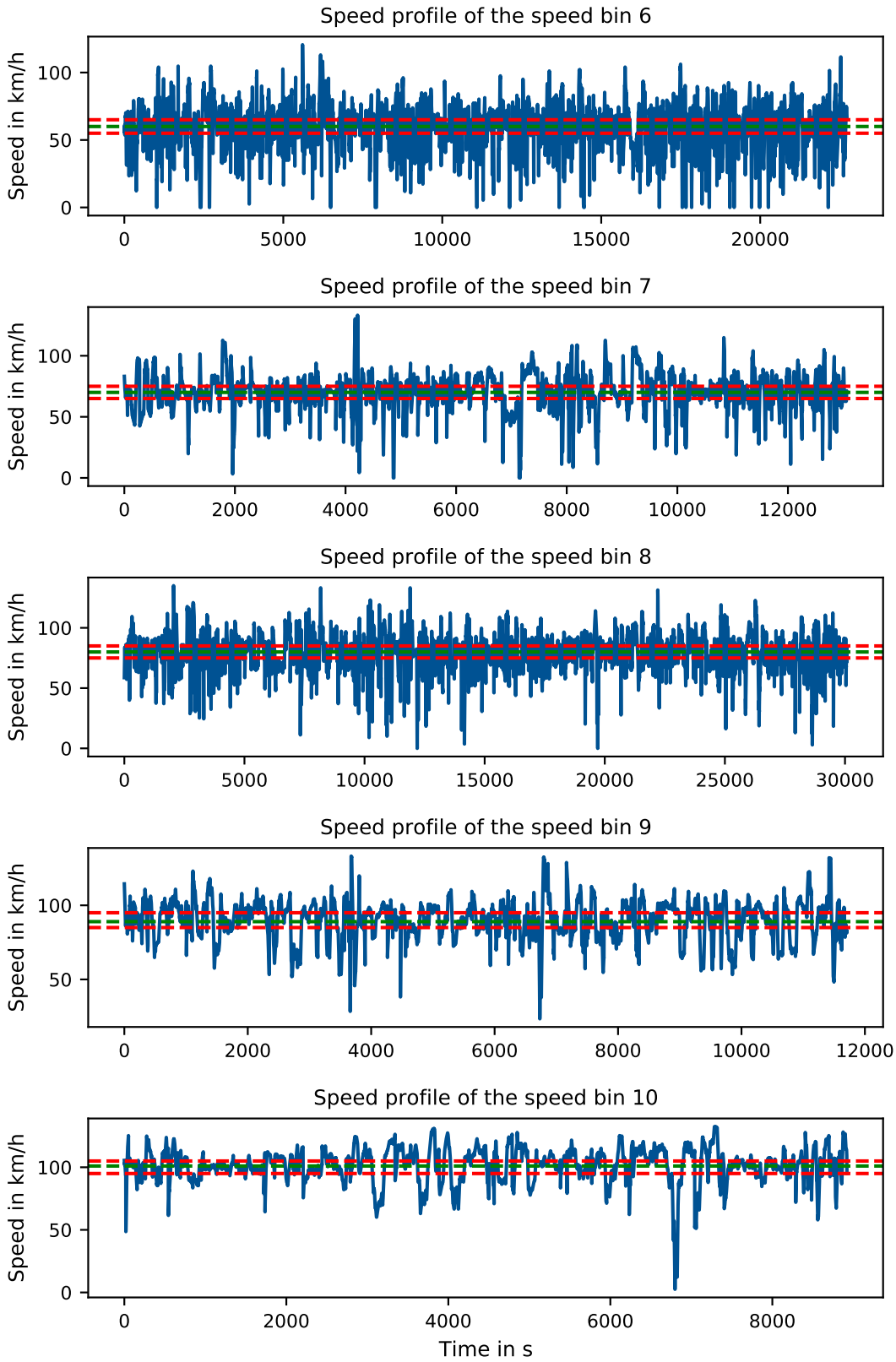


Figure A.11: Speed profiles from the speed bin configuration B: speed bins 6 to 10.

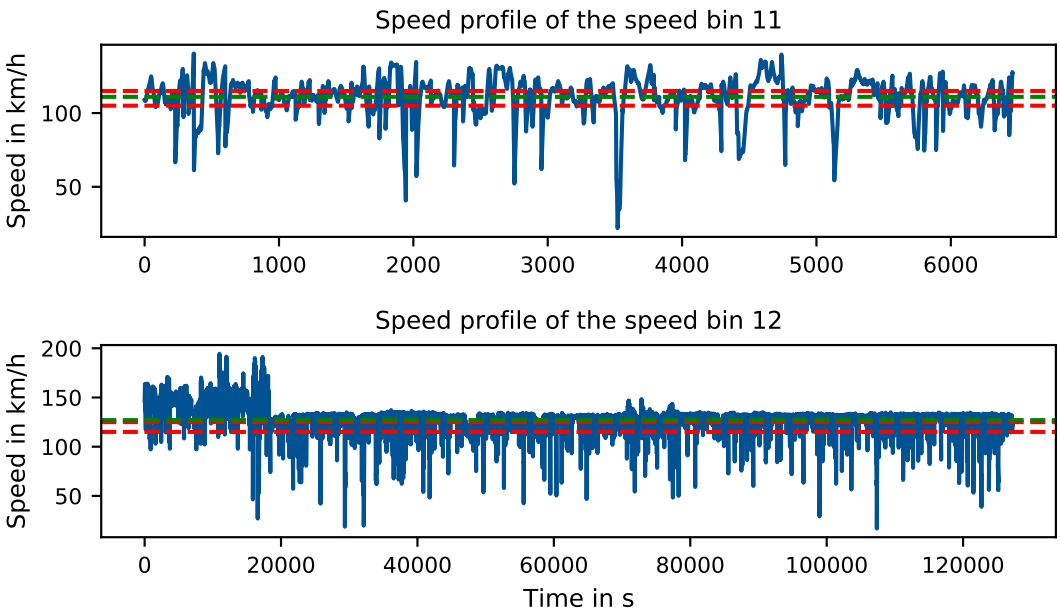


Figure A.12: Speed profiles from the speed bin configuration B: speed bins 11 and 12.

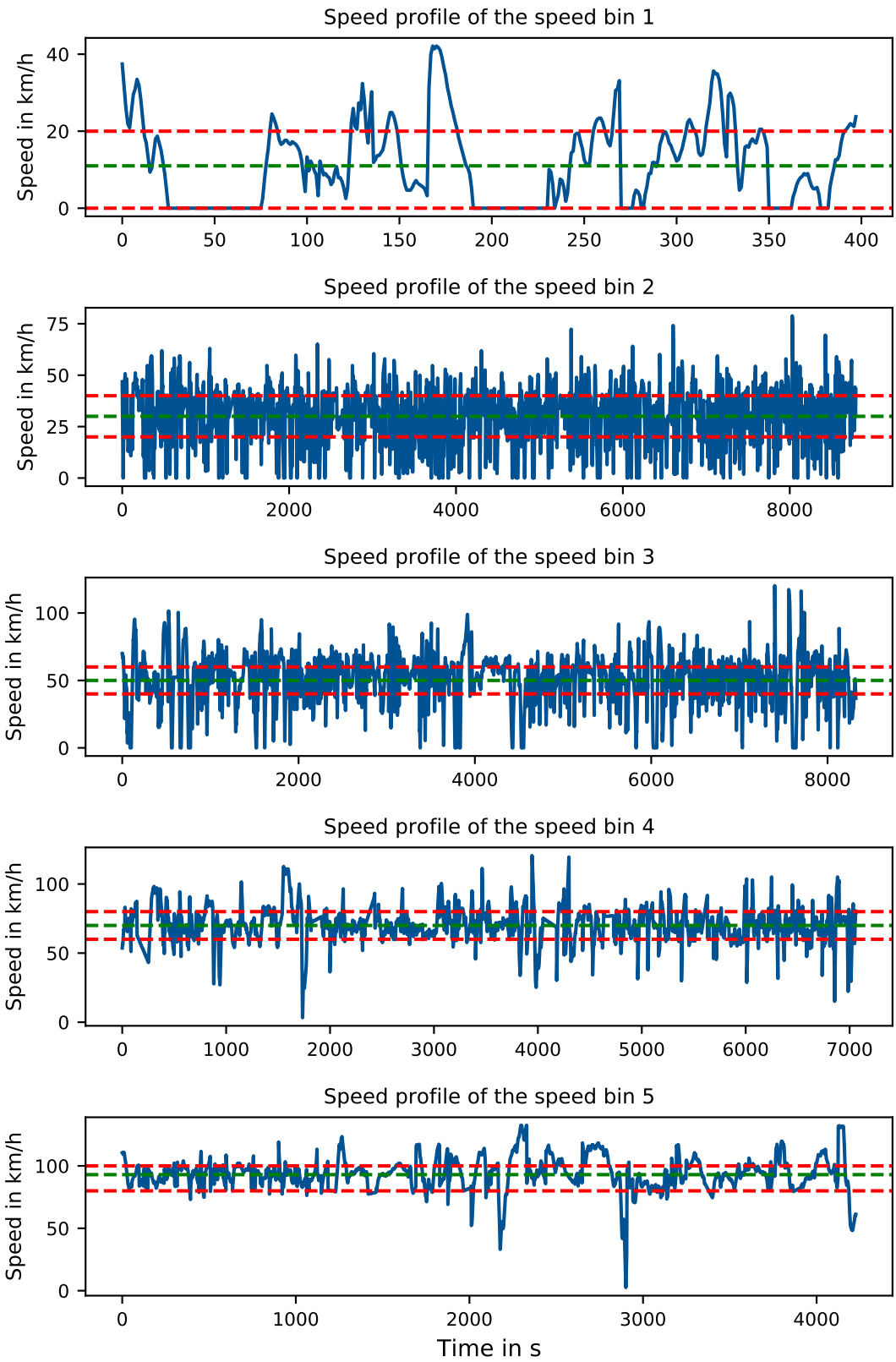


Figure A.13: Speed profiles from the speed bin configuration C: speed bins 1 to 5.

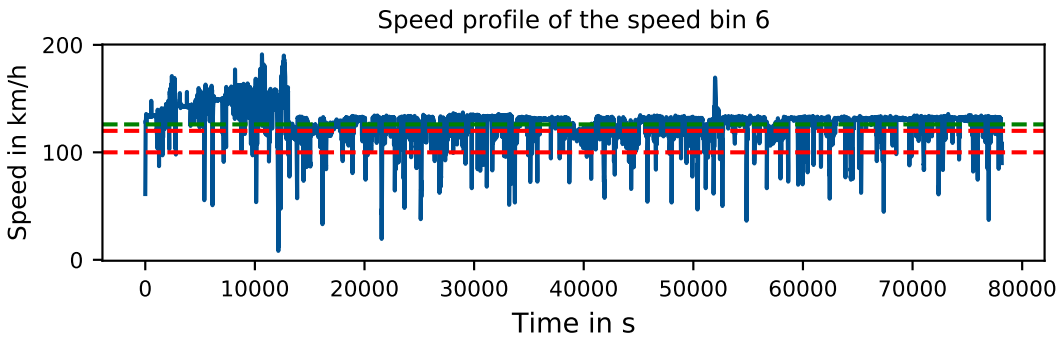


Figure A.14: Speed profiles from the speed bin configuration C: speed bin 6.

B Appendix II: Contents of SD card

1. Thesis
2. Exposé
3. Literature
4. Documents
5. Queries
6. Consumption framework
    - fastsim-2020a
        - cycles (in "imported cycles")
        - docs
            * Results
            * Training data
            * Other functions and notebooks
        - vehdb (see "vehicle data")