Assessing the Impact of Data Availability on Parking Prediction Accuracy: a Case Study Using a Simulation-based Approach

S.J. Hartvelt, University of Twente
s.j.hartvelt@student.utwente.nl

April 24th, 2020

Acknowledgements

The past year of my life has been dedicated to creating the thesis that you currently have in your hands. It has been a tough and at times infuriating process, but I am very happy to have finished it and to be receiving my master's degree in Business Information Technology. I would like to express my sincere gratitude to those who have helped me during this period.

First of all, I would like to thank Ronny Boland for giving me the opportunity to execute this research assignment at Nedap Identification Systems. Ronny was always up for a hot chocolate or a chat about my research and the afternoon walks with the Identification Systems team were a pleasure to be a part of.

Secondly, I'd like to thank Hans Moonen and Geert Heijenk, my supervisors from the University of Twente. Despite us not having regular meetings, whenever we did talk both of you provided me with sharp, constructively critical feedback. After our discussions I was always more motivated to continue working and take the next step.

Next, I'd like to thank Simon Stock from the municipal parking agency of Kortrijk, Parko. From the moment I told Simon about my research, he has been extremely forthcoming with any data he could share with me. This has been invaluable for my research.

Furthermore, I'd like to express a sincere admiration and gratitude for Jason Brownlee of www.machinelearningmastery.com. I came into this research with a very limited understanding of machine learning, and his ebook and the extensive amount of free tutorials were extremely helpful.

My love and gratitude also goes to my family, who have always been supportive. A very special thanks goes out to my grandpa, who has always been very involved in my study progress. I am extremely thankful he can read this thesis and can attend the accompanying presentation.

Finally I'd like to thank Dispuut Yorinf. Through their constant bullying and enquiries about my graduation progress, I was sure to never forget what I was working on. With friends like these, who needs enemies. I love you guys.

Abstract

A case study is performed on the parking situation of the city center of Kortrijk, in Flanders, Belgium, through a simulation of the study area. Using SUMO, a simulation is built using five different data sources about the parking situation in Kortrijk. This simulation creates the possibility to have detailed data on the occupation of every parking space in the study area, by equipping each simulated parking space with a simulated parking occupancy sensor. The dataset generated by the simulation is used to develop a baseline prediction algorithm using Long Short-Term Memory Recurrent Neural Networks. Throughout a series of experiments, variations to the sensor coverage level, the geographical distribution of the sensors and the demand for parking are made. Furthermore, induction loops that count traffic are added in the simulation, and variations regarding the prediction area size are made. The results from these experiments show a linear relationship between the sensor coverage level and the accuracy of the parking occupancy predictions. Adding alternative data sources such as traffic counts further increases the accuracy of the predictions. When using sensor information to predict parking occupancy, the optimal prediction area size is found to be between 25 and 100 parking spaces. To combat false negatives, where no parking spaces are available but the prediction algorithm concludes there are, sensor coverage should be 100% in the most congested areas with the highest demand for parking.

Committee

This thesis was supervised and examined by:

Prof.dr.ir. Geert Heijenk	University of Twente
Dr.ir. Hans Moonen	University of Twente
Ing. Ronny Boland	Nedap N.V.

List of Figures

List of Tables

List of Abbreviations

GIS . . . . Geographic Information System
KOR1 . . . . . . Kortrijk 1, parking tariff zone
KOR2 . . . . . . Kortrijk 2, parking tariff zone
LSTM . . . . . . Long Short-Term Memory
MAE . . . . Mean Absolute Error
MAPE . . . . . . Mean Absolute Percentage Error
OSM . . . . OpenStreetMap
PGIS . . . . . . Parking Guidance and Information System
RNN . . . . Recurrent Neural Network
SUMO . . . . . . Simulation of Urban MObility
SVR . . . . Support Vector Regression
TraCI . . . . . . Traffic Control Interface
VMS . . . . Variable Message Sign

Contents

1 Introduction

1.1 Background

In 2018, 55% of the world population resided in cities. This is forecast to rise to 68% in 2050 [56]. With growing urbanisation, the number of vehicles and, as a result, the demand for parking in cities will grow. Already, traffic cruising around for a parking space makes up a significant amount of the total traffic in central business districts of cities [51]. To combat the growing parking problematics, smart parking solutions have started appearing. Variable Message Signs (VMS) are in use worldwide since the introduction of the first Parking Guidance and Information System (PGIS) in Aachen in the early 1970s [57]. This signalled the first foray into making parking search easier with the help of technology.

The backbone of all smart parking implementations is data about parking occupancy. For parking garages, this is relatively easy. With a limited number of entry and exit points, simply counting the vehicles entering and exiting the facility suffices to keep an accurate count. For on-street parking spaces, this is not an option, as they are not closed off. Many different technologies to monitor on-street parking space occupancy exist, including but not limited to stationary magnetic parking sensors [33, 52, 63, 68], camera-based occupancy monitoring systems [1, 2, 17, 28, 46, 64] and monitoring through probe vehicles equipped with sensors [16, 25, 38, 43]. In an ideal world, the occupancy of all parking spaces would be known and accurate predictions would be made based on that data, allowing commuters and visitors to make informed decisions about where to park, eliminating all traffic cruising for parking. In practice, there are very few real-world applications of these on-street parking monitoring technologies, as the costs for accurate systems are prohibitive [24].

There are a few examples of cities or districts where sensors are rolled out to monitor on-street parking, which are discussed in Section 2.2.1. In these cases, it becomes possible to make predictions about future parking occupancy and the probability of finding a vacant parking space based on historical data. These predictions, further discussed in Section 2.4, are rather accurate. However, they do rely on 100% coverage of parking occupancy sensors in the area that is to be predicted for. This raises the question whether it is possible to make accurate predictions with fewer sensors. If it would be possible to retain accuracy in parking occupancy predictions with only a limited number of sensors, this would lower the costs of smart parking projects significantly.

1.2 Research Goal

The main goal of the research is to assess the impact that stationary parking sensors have on the accuracy of parking predictions. This knowledge will help make informed decisions about investments into smart parking through quantification of the benefits of installing sensors and through insight into the optimal spatial distribution and necessary penetration rate of these sensors.

1.3 Research Questions

In order to achieve the research goal described above, the main research question of this master's thesis is stated as follows:

> *"What is the impact of varying penetration rates and distributions of stationary sensors on the accuracy of parking predictions?"*

To help answer this question, the following sub-questions are formulated:

1. *What different data sources are used in parking prediction?*
2. *How does the availability of each data source influence the accuracy of the prediction?*
3. *Which prediction algorithms are the state-of-the-art?*
4. *What are the effects of varying the spatial distribution and penetration rate of stationary sensors?*
5. *What is the economic trade-off between accuracy and installing stationary sensors for predictive analysis?*

The methods and the approach used to research and answer these questions, as well as the scope of the research, are described in detail in Chapter 3.

1.4 Contribution

The contribution of this research is twofold. Firstly, it quantifies the relationship between parking sensor coverage and parking occupancy prediction accuracy. In this process, insights into the geographical distribution of the parking sensors, the impact of other data sources such as traffic counts, as well as the optimal prediction area size are obtained. Secondly, this thesis proposes a novel research method to research parking prediction. Using a simulation with parking demand modeled on real-world data enables many methods of experimentation, as variables can be easily tweaked and the generated data set is 100% accurate, because the simulated parking sensors are infallible.

1.5 Outline

After introducing the background of the research and its associated questions in Chapter 1, Chapter 2 will explore the work that has already been conducted on this topic by scholars by discussing the history of parking research, smart parking in literature and real-world implementations, parking simulation and the different techniques used to predict parking occupancy. Chapter 3 describes the methods that will be utilised to explore and answer the stated research questions.

The chosen study area is described in detail in Chapter 4. After describing the parking situation in the study area, Chapter 5 discusses all data that was collected from the study area and the approaches taken to process and analyse this data. The findings from this data analysis are then used to build a simulation that is as authentic as possible. This simulation, its components, algorithms, scope and validation process are discussed in Chapter 6.

The simulation output is then used to build a baseline prediction algorithm. Its architecture and performance are described in Chapter 7. This baseline is used as the benchmark throughout the following chapters. In these, variations regarding sensor availability and distribution and simulation parameters are made. These experiments and their results are described in Chapters 8 and 9, respectively.

Rounding up, Chapter 10 discusses the limitations of the research and the real-world implications of the findings, quantifying the trade-off between investments into a smart parking sensor system and its prediction accuracy. Chapter 11 draws conclusions from the previous chapters and finally, Chapter 12 discusses the possible directions for future work.

2 Related Work

In a prior research paper written as a preliminary literature review to prepare for the research discussed in this paper, the smart parking landscape was analysed [27]. Through a structured literature review process, in the end 228 papers were read and analysed. The following is a short summary of the findings of that literature review, tailored to the specific direction of this research.

2.1 History & Cruising for Parking

Since the inception of the automobile, there has been a need for parking when said automobile is not in use. In fact, Bates et al. found that personal vehicles are only in use 3,5% of the time, parked at home 80% of the the time, and parked elsewhere the remaining 16,5% of the time [12]. With an increasing number of vehicles comes the problem of where to store them. As early as 1923, a patent for the vertical storage of vehicles was filed [44], strikingly similar to the vertical parking garages currently ubiquitous in Japan. In 1927, Simpson published a paper called *Downtown Storage Garages* [53] in an issue of *the ANNALS of the American Academy of Political and Social Science* entirely dedicated to city traffic and parking problems. Simpson calls the vehicle storage problem in cities *"the parking evil"*. Simpson found that in two locations in Detroit, traffic cruising while searching for a parking spot downtown accounted for 19% and 34% of the total traffic, respectively. This shows that, by that time, the parking of vehicles in the downtown areas of cities had become a problem worth investigating and solving.

In his overview of the problem with vehicles cruising for a curbside parking spot, Shoup [51] refers to several parking studies conducted throughout the 20$^{\text{th}}$ century. These studies have attempted to quantify the cruising problem in various locations. The results vary wildly, which makes sense as most of the time, none of the traffic will be cruising, but some of the time a lot of the traffic in a city may be cruising while looking for a parking spot, Shoup states [51]. This cruising has been extensively covered in scientific literature, spurred on by Shoup's 2006 paper *Cruising for Parking* [51]. Besides Shoup [41, 50, 51], Anderson et al. [3] and Arnott et al. [4, 5, 6, 8] have written on cruising extensively. In the early 2000s, research into parking has accelerated as a direct result of Shoup's pioneering work [51]. Many attempts to model, simulate and analyse parking have been undertaken, and smart parking with the help of technology started appearing.

2.2 Smart Parking

Smart parking solutions operate on the basis of the availability of data. In order to do any kind of data analysis, guidance to empty parking spaces or occupancy prediction, data about parking space occupation has to be available. A multitude of different sensor technologies exist to detect whether a parking spot is occupied. Lin et al. identified eleven types of stationary sensor technologies, as well as four more mobile technologies [36]. These technologies vary as far as costs, accuracy and ease of installation are concerned. Of these 15 technologies, the most common and most accurate stationary sensor technologies are magnetometers and camera-based sensor systems.

The most common stationary parking detection sensors are magnetometers [36], which work by detecting a change in the Earth's magnetic field as the result of a ferromagnetic object (i.e. a vehicle) being placed in the vicinity of the sensor. This sensing method is passive and thus requires no energy to operate, with an accuracy of over 99% [63]. Magnetometers are the state-of-the-art in parking occupancy detection and are the only sensor technology with notable commercial implementations.

The best alternative to magnetometers are camera-based sensor systems. Driven by the current interest in artificial intelligence, many recent research efforts have been made into camera-based parking occupancy detection [17, 28, 46, 55, 64]. Amato et al. utilise a Raspberry Pi Smart Camera to monitor a parking lot. They use deep learning with Convolutional Neural Networks [1, 2]. Using the same view for training and testing, error rates are under 1%. Using a classifier trained on other views than used for testing, accuracies range from 82.88% to 98.81%, with most results hovering around the 88% mark.

Camera-based sensor systems are able to achieve a high degree of accuracy when they are trained using training data with the same view as the operational cameras. Challenges that still have to be solved include occlusions by objects and shadows, the supply of power to wireless camera systems and security issues.

2.2.1 Smart Parking Projects

The three most notable examples of on-street parking space sensoring are the SFPark project [45], the LA Express Park project [37] and the 4300 in-ground sensors installed in Melbourne [20]. In all of these projects, magnetometers were the sensors of choice. Of these three projects, the data from the SFPark project and the sensors in Melbourne have been published as open data.

2.3 Simulation

Multiple attempts at simulating parking have been undertaken in literature. The aforementioned Arnott [6, 7, 10] has pioneered parking simulation literature, by building on his own mathematical parking models. Dedicated parking simulation software has also been written.

PARKAGENT is introduced in two papers, by Benenson et al. [14] and by Levy et al. [35]. It is an agent-based model in which each driver has a specific origin and destination. The model links a geosimulation approach to a Geographic Information System (GIS) database to simulate a real city, based on three GIS layers; street network, buildings and off-street parking facilities. PARKAGENT simplifies the first part of a driver's journey, and focuses on parking search & choice, the actual parking and driving out again. The validation performed in [35] shows that PARKAGENT closely resembles real-world scenarios. The main contribution of PARKAGENT is that it accounts for the contiguity of parking spaces and the autocorrelation that occurs when the occupancy rate rises and people start looking for parking spaces on the border of congested areas. This has a profound effect on parking when the occupancy rate rises to 95% and higher [35].

Dieussaert et al. describe SUSTAPARK [22]. Similar to PARKAGENT, it is an agent-based spatio-temporal simulation model. Similarly to PARKAGENT, GIS data is used to model the city. The parking choice and search algorithm is less sophisticated than PARKAGENT, and involves the proposal of nearby parking spaces based on their utility to the agent. Dieussaert et al. developed SUSTAPARK for the use in a case study in Leuven [54]. The authors draw many conclusions about parking modeling in general and about SUSTAPARK. Foremost, they conclude that parking is a highly complex problem which relies on many different aspects of city planning and mobility. The authors note that it seems unlikely that all these aspects could be captured in a single model.

MATSim[1] is an open-source framework for implementing large-scale agent-based traffic simulations. Being a traffic simulator foremost, MATSim generally does not simulate parking. It simply assumes drivers park with no delay at the time of their arrival [60]. Waraich et al. extend MATSim with a parking choice model [60]. They do not add a parking search model and lay out many directions for future research. Waraich et al. expand on this research by adding a parking search model in [61]. When comparing MATSim with the SUSTAPARK and PARKAGENT, the main trade-off is simulation detail. Where MATSim is a complex traffic simulator with options like mode choice, SUSTAPARK and PARKAGENT lack detail and functionality in that area. However, they offer more detail in road modelling and have better performance on high resolution networks and with more agents.

SUMO[2], or Simulation of Urban MObility, is an extensive, well-documented open source traffic simulator that was developed in 2001. Out of the box, it does not offer parking choice or parking route models. SUMO does not appear often in the parking literature, but has been used by Lendak et al. [34] and Mejri et al. [39]. In these two papers, SUMO is used to generate and simulate routes and mobility traces, but not for the parking of vehicles. SUMO could be extended to include parking behaviour, but other simulation software offers more specialised options.

[1] https://matsim.org/ - Accessed 2019-04-19
[2] https://sumo.dlr.de/index.html - Accessed 2019-04-19

2.4 Parking Prediction

In a lot of research, the parking predictions are based on historical off-street parking lot occupancy data [11, 18, 19, 42, 65].

Multiple machine learning prediction models for on-street parking occupancy have been proposed. The most recent is the work by Yang et al., who use Graph-Convolutional Neural Networks to extract the spatial relations of traffic flow in large-scale networks, and utilise Recurrent Neural Networks with Long-Short Term Memory to capture the temporal features [66]. The basis for their occupancy prediction is parking meter data, in line with Yang et al.'s earlier work [67]. However, parking meter data suffers from accuracy errors due to underpaid, overpaid and unpaid parking. This accuracy error combined with the error in occupancy prediction leads to a 12% Mean Absolute Percentage Error (MAPE).

Using real-time occupancy data from in-ground sensors installed in Santander, Spain, Vlahogianni et al. use neural networks to estimate parking occupancy 15 minutes into the future with up to 3.6% MAPE [59]. This shows real-time information has a very positive effect on parking occupancy prediction.

Zheng et al. used data from the SFPark project and the sensors in Melbourne as training data for a neural network, a regression tree and support vector regression (SVR) [69]. They conclude that the regression tree method using a feature set that includes the history of the occupancy rates along with the time and the day of the week performs best for parking availability prediction.

Rajabioun et al. also used data from the SFPark project to develop a vector spatiotemporal autoregressive model to predict parking availability at the expected time of arrival of a driver, in order to recommend the parking space with the highest probability to be vacant [42]. The predictions achieve a MAPE of 14% and the system is able to direct a vehicle to a parking area with a free parking space 95% of the time.

Bock et al. propose a 2-step approach to predict parking availability [15]. Their first step handles the data processing, in which SVR's are used to smooth the raw parking data and extract a trend. The idea is that models are easier to train on smooth trend curves than on raw, chaotic parking data. Their second step is performing a standard regression, again using SVR, to predict parking availability. Their approach is shown to outperform the standard regression method of Zheng et al. discussed in the paragraph above, and is able to exploit a longer history, while occupying only 40% of the space of raw data.

Ionita et al. propose a method that estimates parking availability in areas without sensors by comparing their background data from Geographical Information Systems [29]. They compute similarity values between neighbourhoods with and without sensors based on the average visit duration to public amenities in those neighbourhoods. Neighbourhoods were clustered using k-means clustering and several machine learning methods were used for the availability prediction, with extreme gradient boosting being the best performing model. The data from the SFPark project was used for this research.

Pflügler et al. use Neural Networks to predict parking occupancy using publicly available data such as traffic intensity and weather. They found that location, time and weather were the most relevant categories to consider when predicting parking occupancy [40].

Badii et al. use different data sources to train Recurrent Neural Networks (RNN), Support Vector Regressions, Auto Regressive Integrated Moving Average models and Bayesian Regulated Artificial Neural Networks to predict occupancy in parking garages [11]. They find that the time of day and historical occupancy data are the most important factors for predicting occupancy, and notice an improvement to their predictions when including traffic flow information. They reference that their findings are coherent with the findings of Pflügler et al. for on-street parking [40]. They hypothesise that traffic flow information would be less useful for predicting on-street parking prediction.

The research of Evenepoel et al. is, to the author's best knowledge, the only research to investigate the efficiency of a sampling approach for on-street parking sensors [24]. They conclude that city-wide parking totals can be estimated with great accuracy when equipping only a fraction of the parking spaces with sensors, under the assumption that the parking vehicles are uniformly distributed over the city.

2.5 Round-up

Many different techniques have been used in the literature for predicting the occupancy of parking garages and on-street parking. There is no clear best method, but it is clear that the most important features for predicting parking occupancy are the time of day and historical parking data. Weather and traffic flow information play smaller roles but do increase the accuracy of the predictions. This answers the stated sub-questions 1 and 2. Many different methods and algorithms are used to predict parking space occupancy. There is no clear best approach, leaving sub-question 3 partly unanswered.

3 Method

3.1 Approach

The approach that will be taken to answer the research questions described above is as follows. After the structured literature review described in Chapter 2, a city or district will be chosen on which to base the research. A simulation environment will be set up to simulate parking events in that chosen city or district. Then, a prediction algorithm for parking occupancy will be written. Finally, the gathered and simulated data and parking events will be used to gauge the accuracy of the prediction algorithm. Each of these steps is briefly detailed below.

3.1.1 Simulation Environment

Ideally, this research would have been conducted with the data from a parking sensor network that covers every single parking space in a city. Unfortunately, that sensor setup does not exist. Therefore, using as much real-world data as possible, a city or district and all its parking events will be simulated. By using realistic input parameters for the parking vehicles in the simulation, a realistic depiction of the parking situation will be created.

Simulating the parking situation offers something no real-world data set can. The simulation outputs 100% accurate and reliable parking occupancy data for 100% of the parking spaces in the simulated city. This offers great opportunities for prediction and experiments, and will allow the research questions to be answered. The decision regarding the simulation environment and the exact approach regarding the setup of the simulated city is discussed in detail in Section 6.1.

The simulation will undoubtedly have shortcomings that are not present in traditional real-world parking datasets. On the other hand, it offers many opportunities for experimentation. In the simulation, sensors do not fail and are not obstructed or otherwise tampered with. Being a simulation, it offers ample opportunities for experimentation that would not be possible in the real world.

Validation The simulation will be validated by comparing the results found during the data analysis step with the simulation output. If the simulation behaves correctly, the simulated parking events should resemble these data analysis results with respect to characteristics such as the number of parking vehicles, their parking duration, and the average number of vehicles per parking spot.

3.1.2 Study Area

Because the goal was to simulate a city center as accurately as possible, the choice for a study area relied heavily on the availability of data. This research has been carried out in collaboration with Nedap, who develop magnetometers[3]

[3]https://www.nedapidentification.com/products/sensit/ - Accessed 2019-12-18

for smart parking purposes, as discussed in Section 2.2.

Nedap's largest parking project is in the municipality of Kortrijk, where over 1000 sensors are deployed in so-called Shop&Go parking spaces[4]. These are further discussed in Chapter 4. The data from these sensors was made available for this research, providing a good starting point. After contacting the municipal parking agency of Kortrijk, Parko[5], they were able to supply a wealth of data for the research. These datasets are further discussed in Section 5.1.

Alternatively, a decision could have been made to use one of the open parking data sets provided by either the SFPark project [45] or the parking sensor project in Melbourne [20]. These datasets have some data quality concerns however, and given the combination of sensor data and the aggregated datasets by Parko, Kortrijk was a natural choice as the study area for this research.

3.1.3 Prediction Algorithm

During the literature research, many different machine learning techniques were found that can be used to predict parking occupancy. In order to answer the stated research questions, the chosen architecture should support multiple types of inputs. This will allow the impact of different data sources to be researched. This choice of prediction algorithm architecture is described in detail in Section 7.1.

Validation The performance of this prediction algorithm will be compared to two naive parking prediction methods, as well as to the state-of-the-art in the literature. These naive prediction methods are the historical average parking occupancy of the prediction area, and simply using the observation of the parking occupancy at the time of the prediction as the prediction. The proposed baseline prediction method should at least perform better than those naive methods. The goal is not to create a prediction method that outperforms or challenges the state-of-the-art, but rather to create an efficient baseline prediction strategy that is easy to adapt during the execution of the experiments.

3.1.4 Experiments

Finally, after performing the data analysis of the study area, setting up the simulated environment and establishing a baseline prediction algorithm, the experiments that are necessary in order to answer the research questions can be executed. The planned experiments and variations are as follows:

[4]https://www.nedapidentification.com/cases/shop-go-zones-powered-by-parking-sensors/ - Accessed 2019-12-18

[5]http://www.parko.be/ - Accessed 2019-11-06

- Varying spatial distributions and penetration rates of sensors
- Varying cluster size used for prediction
- Combining prediction clusters
- Varying the number of parking vehicles
- Adding other data sources (e.g. traffic counts via induction loops)

Finally, the gathered and simulated data and parking events will be used to gauge the accuracy of the prediction algorithm. In order to answer the research questions, predictions and simulations will be made with the following variations. These experiments are described in detail in Chapter 8.

3.2 Scope

Parking prediction is an extremely multifaceted problem and can take an extensive amount of data sources as input. In order to achieve the research goal in a reasonable time frame, the scope of this research has to be limited. Each topic discussed above will be briefly touched on to describe the boundaries of this research.

Firstly, the input for the simulation will be limited to that which can be deducted from the data sets supplied by the municipal parking agency of the study area and any open data sources. In the case of Kortrijk, this means that some factors that are shown to have a serious impact on parking, such as weather [11, 40], are left out of the scope of this research. The positive impact on parking prediction accuracy of including weather data has already been proven and including it would not help answer the research questions more effectively.

Secondly, regarding the simulation, the scope will be limited to only the vehicles parking in on-street parking spaces. This means omitting all other forms of traffic and all vehicles parking on private property or in parking garages. The impact of this scope limitation is discussed in detail in Section 6.8.

Thirdly, as stated above, the focus during the development of the prediction algorithm will not be on maximum performance, but on efficiency and creating an approach suitable to the planned experiments.

Overall, this research will be limited to the pure technical aspect of parking prediction. Solving the smart parking problem is extremely multifaceted, and requires effort into data collection, data analysis and data processing and prediction. Furthermore, even the best predictions are useless without a way to act on those predictions, a way to efficiently guide traffic to parking spaces, and without creating parking policies based on the insights gained by the smart parking project. This research concerns itself with the first step in the smart parking process, namely how to design a smart parking project as efficiently as possible while sacrificing as little accuracy as possible.

4 Parking in Kortrijk

This chapter will briefly describe the parking situation in the study area, which is the city center of Kortrijk, located in Flanders, Belgium.

4.1 Parking Zones

In the study area, there are 9 parking garages, which are left out of the scope of the research, as described in Section 6.8. There are 4443 on-street parking spaces, which are divided into four categories. There are two paid parking zones, KOR1 and KOR2. In the former, there is a maximum parking duration of 2 hours. In the latter, it is allowed to stay until the end of the day. In addition to the two paid parking zones, there are two ways to park on-street for free in Kortrijk. Firstly, there are the blue zones, in which parking vehicles have to display a parking disc stating their arrival time. The maximum parking duration in these zones is 2 hours. Secondly, scattered around the city are Shop&Go parking spaces. These are special short-stay parking spaces, with a maximum parking duration of 30 minutes. Figure 1 shows the parking zones KOR1, KOR2 and the blue zones in red, yellow and blue, respectively. The figure has been edited manually to reflect the situation at the time of the simulation design, as the zones have changed since the start of this research. Figure 2 shows the location of the Shop&Go parking spaces in the study area. Table 1 summarises the characteristics of the four parking zones in the study area.

Each Shop&Go parking space is outfitted with a magnetometer parking sensor from the SENSIT system, developed by Nedap[6], in order to monitor their occupancy. When a vehicle has not left the parking space after the maximum parking period, a notification is sent to nearby parking attendants who can write a parking ticket, if applicable. The status of the Shop&Go parking spaces can be viewed online on the website of the municipal parking agency, Parko[7]. This is illustrated in Figure 2.

Zone	Time limit	Paid parking
KOR1	2 hours	Yes
KOR2	End of day	Yes
Blue	2 hours	No
Shop&Go	30 minutes	No

Table 1: Characteristics of parking zones in the study area

[6] https://www.nedapidentification.com/cases/shop-go-zones-powered-by-parking-sensors/ - Accessed 2019-12-18

[7] http://www.parko.be/ - Accessed 2019-11-06

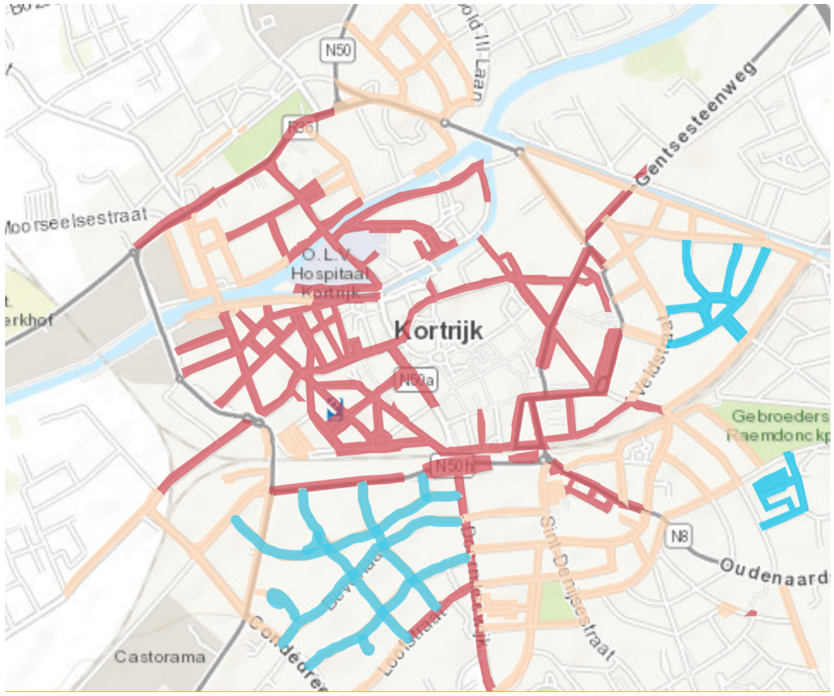


Figure 1: Parking zones in the Kortrijk city center.
Source: https://http://www.parko.be/bezoeken - Accessed 2019-11-06

4.2 Parking Permits

Two types of permits exist that exempt vehicles from paying for parking in Kortrijk. Firstly, a handicap placard allows free parking without a duration limit throughout the city. In the Shop&Go parking spaces, the time limit of 30 minutes was recently reinstated for handicap placard holders, as 50% of all spaces were occupied by placard holders who did not have to adhere to the time limit[8]. In the evaluations of the LA Express Park smart parking project, parked cars using a handicap placard can make up over 90% of all parking time on congested block faces [21]. Clinchant et al., Glasnapp et al. and Zoeter et al. all report high levels of handicap placard use [21, 26, 70]. Non-paying parkers do not only make up a large part of the parking population, but they also park for disproportionally long times, not being bound by parking fees or time restrictions. While of course not all handicap placard use is abuse, it is likely a large proportion is, as argued by Shoup in [47]. Thus, reinstating the time limit for placard holders in Kortrijk might greatly impact the availability of parking spaces in Kortrijk.

[8]https://www.nieuwsblad.be/cnt/blkva_04449304 - Accessed 2019-12-18

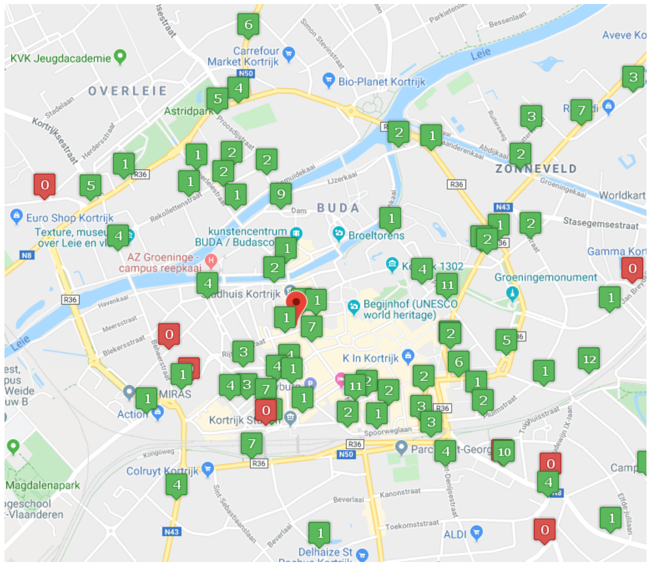


Figure 2: Live occupancy of Shop&Go parking spaces. Snapshot taken 2019-12-18 11:09 Source: https://shop.parko.be/m/#/parking/shop_go - Accessed 2019-12-18

Besides handicap placards, Parko supplies municipal parking permits for inhabitants, health care providers, business owners and employees. These permits vary in cost and are subject to conditions, but when purchased, provide free parking in their designated parking zone without a time limit.

5 Data Sources

In order to achieve a sufficiently realistic simulation to perform the prediction experiments on, data about parking occupancy has to be analysed. The following chapter discusses the data sources that were utilised in this research, their characteristics and the steps undertaken to find the number of parking events and their temporal and spatial distribution.

5.1 Data Sources

For this research, multiple data sources were acquired and utilised. They are as follows.

- All parking areas in the Kortrijk city center in GIS format
- Mobile phone parking ticket sales
  - Monthly averages from January 2012 until June 2019
  - One month (June 2019) of all mobile phone parking transactions
- Parking meter ticket sales
  - Monthly averages from January 2010 until June 2019
- One year (2019) of data from the Shop&Go stationary parking sensors
- 13 months (June 2018 - June 2019) of parking violation data from the Kortrijk parking enforcers

Each of these data sources will be briefly discussed.

5.1.1 GIS data

The municipal parking agency in Kortrijk, Parko[9], is currently working on mapping all parking spaces in the municipality in a Geographic Information System (GIS). When the data was requested, this was still a work in progress. In collaboration with Parko, it was decided to limit the research to the city center area, within the ring road R36. All on-street parking spaces within this area have been mapped and are represented in the GIS data set. As no translation between GIS data and the chosen simulation environment exists, all parking spaces were placed in the simulation by hand. This process is further described in Section 6.4. In total, the GIS data set consisted of 4443 parking spaces that were placed in the simulation.

5.1.2 SMS & Meter ticket sales

There are two ways to pay for on-street parking in Kortrijk. Visitors can use the stationary parking meters to buy a pay and display parking ticket, or use their mobile phone to pay via SMS or a mobile website. Parko has supplied monthly statistics for both these forms of payment, ranging back to January

[9]http://www.parko.be/ - Accessed 2019-09-23

2010 for the parking meter data, and January 2012 for the mobile phone payment data. However, as the municipality of Kortrijk has decided to stimulate underground parking and disincentivise on-street parking by removing parking spaces and raising the parking tariffs, only the last two years of these datasets are representative of the current situation and useful for this research. These datasets consist of monthly total ticket sales per parking tariff zone.

Furthermore, a full month of detailed mobile phone parking sales have been supplied, including start date and time, end date and time, the tariff zone and the monetary parking charge.

5.1.3 Shop&Go sensor data

Nedap has placed 1068 stationary parking sensors in the Kortrijk municipality in their short stay parking spaces. The maximum parking duration in these spaces is 30 minutes, intended for a quick visit to a shop or other service. When a vehicle is parked in one of these spaces for more than 30 minutes, a signal is sent to a parking enforcement officer, who can write the offender a ticket. All sensor data is logged in a database, and all data from 2019 has been made available for this research.

This dataset has, for every parking event, a start time and date, end time and date, parking sensor ID and, if applicable, the overstay time which was spent in the parking space illegally. The dataset is not complete, as information for some sensors is missing. Additionally, not all deployed sensors are in the studied area within the R36 ring road. Finally, not all Shop&Go parking spaces in the supplied GIS dataset are represented in this database.

In total, 493 Shop&Go parking spaces were found with a one-to-one mapping to one of the 778 Shop&Go parking spaces in the GIS dataset that were placed in the simulation.

5.1.4 Parking violation data

Kortrijk employs parking enforcement officers who patrol the city to enforce parking payment. All cars that are checked are registered using a PDA. For each vehicle, one record is registered, which includes a date, time, street, GPS coordinates, license plate and, if applicable, a violation code and comments from the parking enforcement officer. This data set was anonymised by removing the license plate before it was supplied for this research.

5.2 Data Analysis

The datasets described above have formed the basis for the inputs of the simulation. The parking vehicles have been split into four categories, based on their parking space selection. Within the city center of Kortrijk there are four different possible parking zones, as described in Chapter 4. For each of the four categories, distributions have been found for the parking duration and arrival time. Then, the total amount of parking vehicles per category was estimated using a combination of the available data about ticket sales and parking violations. These processes are described below. Table 2 shows a summary of all the found parking characteristics per zone.

5.2.1 Parking zones

KOR1 As described in Chapter 4, the tariff zone KOR1 allows short stay parking with a maximum of two hours. The observed parking durations in the datasets are heavily influenced by this restriction, as 30% of all ticket sales are for the full two hours. This is very distinguishable in the histogram of parking durations in Figure 3. In reality, the people who pay for the full 2 hours will either understay or overstay. In their study of parking occupancy in San Francisco, Yang et al. found that there is a small tendency towards underpaying, and thus overstaying for on-street parking [67]. The ratio between understaying and overstaying varies significantly per district, time period, total parking duration and day of the week. In total, during weekdays in the city center, they find a slight tendency towards underpaying and thus overstaying. It is tough to generalise these findings, as there are many factors that influence the payment behaviour. For the purposes of the simulation, the assumption is made that underpaying and overpaying averages out, as the found trends are not severe. The parking duration in KOR1 is identically distributed throughout the day and no correlation with the arrival time of the vehicle has been found.

The data set is split between parking events with a duration under and over 110 minutes. The former part is found to follow a close to uniform distribution on the interval [0,110]. For generating the parking durations in the simulation, the choice has been made to give any parking car in KOR1 a 30% chance to pay for the full 120 minutes, and to follow a uniform distribution on [0,110] otherwise. If the vehicle pays for the full 120 minutes, its actual parking time is uniformly distributed on the interval [110,130] minutes.

KOR2 As described in Chapter 4, the tariff zone KOR2 allows parking until the end of the day. Because of the way the tariff is set up, payments for this zone have spikes around 60 minutes, 120 minutes and 180 minutes. Furthermore, almost half of the parking events in KOR2 last until the end of the paid parking period. In reality, people will understay or overstay this period. No data about this is available however, so the assumption is made that the people who pay for the full day, i.e. until 19:00, will stay until somewhere between 17:00 and 21:00. For the parking events that do not last the entire day, the same assumption about overpaying and underpaying is made as was made with regards to zone KOR1. When removing the parking events that last until the end of the day, no significant correlation between the arrival time of the vehicle and its parking duration has been found, and the parking durations are identically distributed throughout the day.

Accounting for the spikes at the tariff hikes and removing the parking events that last until the end of the tariff period, the rest of the data follows an exponential distribution with a mean of 102 minutes. For the generation of parking

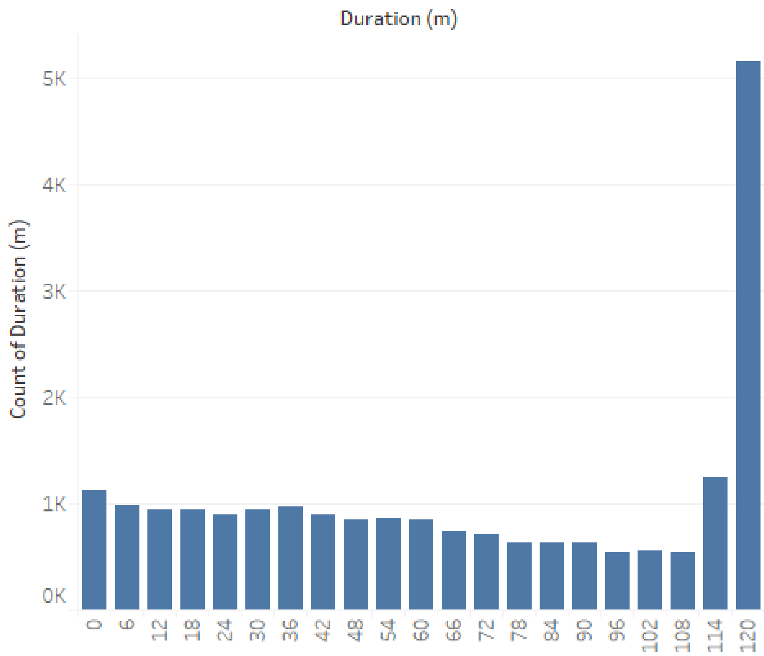


Figure 3: Histogram of parking durations in KOR1, in minutes

durations in the simulation, the assumption was made that the spikes because of the tariff rates in reality even out. Therefore, similar to the approach for KOR1, a split has been made. Each parking vehicle has a chance of 45% to stay until the end of the day and thus leave somewhere between 17:00 and 21:00, and the other vehicles have a parking duration that follows the found distribution.

Blue zone As described in Chapter 4, there is a free parking zone marked by blue lines on the pavement in the Kortrijk city center. In this area, the maximum parking duration is two hours, the same as in KOR1. Due to there being no payment and no parking sensors, there is no specific data on parking duration or parking totals for this area. However, some assumptions can be made on the basis of the parking violation data set. This data set reveals that the cars in the blue zone have the same percentage of parking violations as the cars in KOR1. They are identically distributed through time to the cars in KOR1. Therefore, the assumption has been made that the parking behaviour is identical to the vehicles in KOR1. For the purposes of the simulation, this is good enough.

Shop&Go With all the parking events on the Shop&Go spaces available, no assumptions have to be made in this category. Once again, the timing restrictions on the parking space influence the parking behaviour. When removing

20% of the parking records between 20 and 30 minutes long, the rest of the data set follows a log-normal distribution with parameters μ = 6.16849 and σ = 1.09427. This is illustrated in Figure 4. The records between 20 and 30 minutes make up 10% of the total amount of records. Therefore, for the purpose of the simulation, a similar approach has been taken to KOR1 and KOR2. 2,5% of the generated vehicles will stay between 20 and 30 minutes, uniformly distributed on that interval, and the rest of the vehicles will have a parking duration that follows the found distribution.

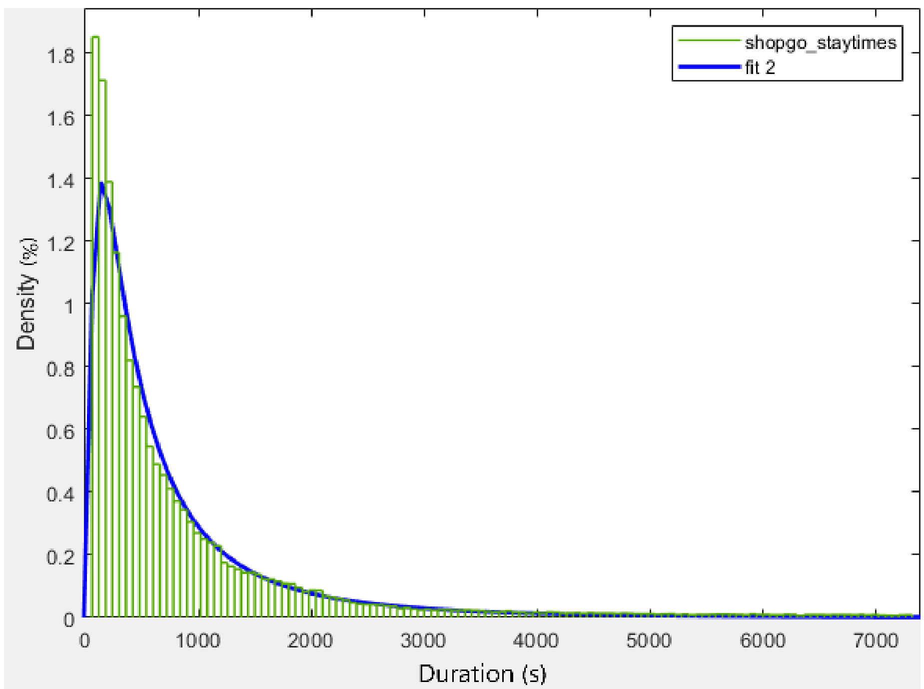


Figure 4: Histogram of parking durations in Shop&Go parking spaces, in seconds, with log-normal distribution fit

5.2.2 Day of the week

The day of the week can severely influence the amount of parking vehicles in a city, as parking tariffs can differ between weekdays and the weekend, and people will work less and shop more in the weekends. In the case of Kortrijk, the parking tariffs are constant from Monday to Saturday, and parking is free on Sunday. The dataset with the month of mobile phone parking ticket sales as well as the dataset of Shop&Go sensor readings were analysed to investigate the effect the day of the week has.

In the phone parking ticket sales, Thursday and Friday are roughly 10% busier than Monday through Wednesday, and Saturdays are 24% less busy than the weekdays. As there are no ticket sales necessary on Sundays, these are not represented in the data set. In the Shop&Go dataset, Wednesdays and Fridays stand out as 7,5% busier than Mondays, Tuesdays or Thursdays, which are all equally busy. Saturdays are 13% less busy than the weekdays. Sundays are 55% less busy than the weekdays.

While some trends can be seen from this data, such as parking occupancy increasing throughout the week and falling on Saturday, there is not enough data to draw serious conclusions. To answer the research questions posed in Chapter 1, whether Thursday or Friday are a few percent busier than other weekdays does not matter. Therefore, all weekdays are considered to be identical and the weekend is not simulated. The effect of busier or less busy days on the accuracy of predictions will be evaluated by varying the total amount of vehicles in the simulation during the experiments, which will implicitly answer the question whether the predictions are also accurate on the weekends.

5.2.3 Parking totals

The dataset of parking violation data encompasses 807.580 total records. Of these, 16.533 are Shop&Go violations. Shop&Go spaces are only checked after a notification of a parking offence is received. In order to estimate the total amount of parking events in the city, these are removed from the data set. 791.047 records of checked vehicles remain. Of these, 384327 possess a city parking permit, and 33.971 were found to display a handicap placard. The records of these vehicles in the parking violation dataset are spatially and temporally similarly distributed to the paying vehicles. They have to follow the same rules and are thus estimated to behave similarly to paying vehicles.

Disregarding these permit holders and handicap placard holders, 372.749 records of checked vehicles remain. Of these vehicles, 48.149 were found to be in violation, meaning an offence rate of 12,93%. Of these violations, 7.271 did buy a ticket, but it had already expired at the time of the check. As these vehicles are already counted in the data discussed above, these are subtracted, leaving 40.878 vehicles that did not buy a ticket. The offence rates between parking zones are similar. The parking violation factor is defined as follows.

$$\frac{\text{Total \# Vehicles}}{\text{Total \# Vehicles - Vehicles without a ticket}}$$

This parking violation factor is $1,1232$. The total amount of vehicles in the categories KOR1, KOR2 and the blue zone will therefore be multiplied with this factor in order to obtain the actual amount of parking vehicles.

The hitherto disregarded vehicles with a permit or handicap placard also play a significant role in the dataset. The parking permit factor is defined as follows and is found to be $2,1222$.

$$\frac{\text{Total \# Vehicles}}{\text{Total \# Vehicles - Vehicles with a permit or placard}}$$

Thus, to find the total amount of monthly parking events for KOR1 and KOR2, the average monthly total over the last two years was taken from the data sets of SMS and parking meter ticket sales. This number is then multiplied by the parking violation factor and the parking permit factor.

The amount of parking vehicles in the blue zones is estimated by finding the ratio of total observations of vehicles parked in a blue zone to the total observations of vehicles parked in KOR1 and KOR2. The amount of parking vehicles in the Shop&Go zones is simply taken as a monthly average from the actual complete dataset. After this process, the characteristics found for each parking zone are as follows.

Zone	# vehicles	Mean parking duration	Distribution	Exception
KOR1	6193	60 minutes	Uniform	30% stay 2 hours
KOR2	2433	102 minutes	Exponential	45% stay until end of day
Blue	1389	60 minutes	Uniform	30% stay 2 hours
Shop&Go	8982	6 minutes	Lognormal	2.5% stay 20-30 minutes

Table 2: Overview of results of data analysis

5.2.4 Arrival times

From the month of detailed SMS ticket sales, hourly arrival rates were deduced for KOR1 and KOR2. As mentioned earlier, the blue zone is assumed to behave identically to KOR1 in terms of its temporal distribution. Hourly Shop&Go arrival rates were taken from the data set. For each hour and each category, it was verified that the inter-arrival times are exponentially distributed and thus behave as Poisson processes. For the purposes of the simulation, each category will generate arriving vehicles based on a Poisson process with a mean based on these hourly averages.

5.2.5 Spatial distribution

Two datasets contain useful information to analyse the spatial distribution of parking events through the city center of Kortrijk. Figures 5 and 6 contain heatmaps of parking records in Shop&Go spaces and checked vehicles by parking enforcement officers, respectively. Several things stand out.

Firstly, two main hot spots can be identified. The first is clearly visible in Figure 6, and is situated near the train station. This location especially stands out in this data set, as according to the municipal parking agency Parko, the parking enforcement is more active in this area. They are more active mostly because there is more parking activity. The heatmap therefore overestimates the difference between this location and the other areas, but that it is more busy is clear. The second hot spot is the rightmost red area in Figure 5. This is the location of the large shopping mall "K in Kortrijk". The Shop&Go

spaces around the shopping mall have an average occupancy comparable to the Shop&Go spaces around the train station.

These heatmaps give a slightly skewed perspective on the parking situation in the city center, as the parking space density is not uniform throughout the city center. When analysing the occupancy of the Shop&Go parking spaces closer and accounting for the density of those spaces per area, six distinct areas can be identified in which the parking occupancy differs substantially from one another. Figure 7 shows these geographic partitions and the location of the Shop&Go sensors in these districts. Besides the two already identified hot spots around the station and the shopping mall in green and yellow respectively, the city center was further divided in the area north of the river in pink, the area south of the train tracks in red, the area in the west and center between the train tracks and the river that acts as the central business district and shopping district of Kortrijk, blue in Figure 7, and finally an area in the north east that has fewer businesses in orange. The average amount of parking events per parking space in a district ranges from 5 vehicles per parking space per day to 15, and are distributed as follows. On average, for every two parking vehicles in the area north of the river, two will park south of the train tracks, four in the northeast, five in the central business district and six in both the areas around the station and the shopping mall.

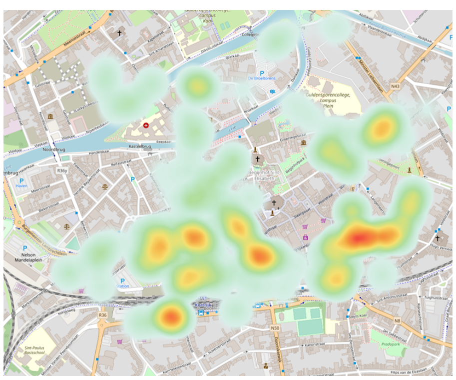


Figure 5: Heatmap of parking events in Shop&Go parking spaces

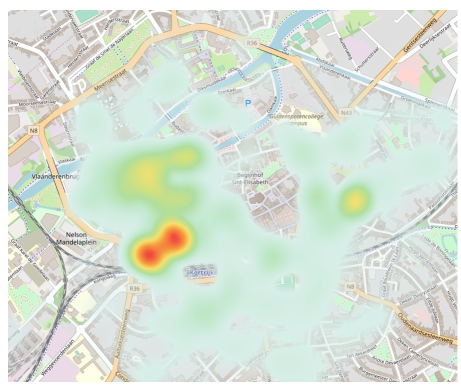


Figure 6: Heatmap of parking records in parking enforcement dataset

The arrival times and parking durations do not show any relevant differences between the districts, apart from the northernmost district, where the parking duration distribution of the Shop&Go parking spaces has a mean of 20 minutes instead of 12 minutes. The parking duration generation formula for the vehicles visiting a Shop&Go space in the simulation is updated accordingly. No data about the geographical distribution of parking durations for vehicles in the other parking zones is available, and are assumed to be identically distributed over the city, as there is no evidence this should be different. While the found parking demand per district is a rough approximation, the exact distribution of

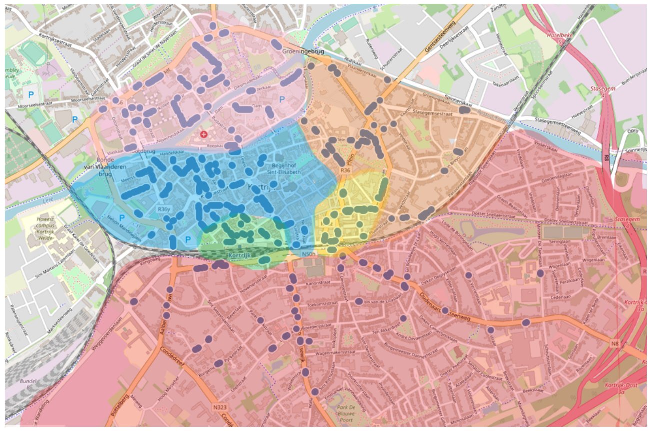


Figure 7: The partitioned city districts and their corresponding Shop&Go sensor locations

this parking demand is not of vital importance. For the purpose of predicting the parking occupancy of a certain area, it is important that vehicles behave realistically with respect to their arrival times and parking durations. The exact parking pressure for that area is of lesser importance, and as long as there is enough variation in parking pressure between streets and areas and the prediction algorithm performs well on streets with varying parking pressures, this approximation suffices for the purposes of distributing parking traffic across the city in the simulation.

6 Simulation

This chapter will describe the simulation environment that was used in the research, the steps taken to construct and validate the simulation, and its scope.

6.1 SUMO

SUMO[10], or Simulation of Urban MObility, is an extensive, well-documented open source traffic simulator that was developed in 2001 by Krajzewicz et al. [32] and described in detail by Behrisch et al. [13]. As a traffic flow simulator first and foremost, SUMO does not appear often in the parking literature. The few occurrences it has are by Lendak et al. [34] and Mejri et al. [39]. In these two papers, SUMO is used to generate and simulate routes and mobility traces, but not for the actual parking of vehicles.

For this research, SUMO was chosen in spite of more dedicated parking simulators such as PARKAGENT [14] or SUSTAPARK [22], as no working versions of these simulators were available from their respective developers. SUMO on the other hand, while not explicitly a parking simulator, is well-documented and has an active community with an openly published mailing list with questions and answers throughout its lifetime.

6.2 TraCI

TraCI, as described by Wegener et al. [62], is short hand for Traffic Control Interface. It uses a TCP-based client/server architecture to provide access to SUMO, by giving access to the simulated objects, offering value retrieval and behaviour manipulation. It offers a python interface that was used in the research described in this paper.

TraCI's main use case for this research is to generate vehicles based on the simulation time, assign parking and reroute vehicles to other parking spaces. SUMO is designed to use pre-generated routes or traffic flows, in which stops have to be explicitly set. For a realistic parking simulation, the assignment of parking spaces has to happen dynamically, and TraCI offers this opportunity. This is further discussed in Section 6.5

6.3 Network

The first step in building a simulation scenario for the city center of Kortrijk was importing the network. SUMO offers a tool called NETCONVERT which converts data from OpenStreetMap[11] (OSM) into a road network compatible with SUMO. This proved a good starting point, but the generated network was not completely free of problems.

Firstly, the streets outside of the city center for which no parking space data was provided, were removed from the network to keep its size manageable

[10] https://sumo.dlr.de/index.html - Accessed 2019-11-08
[11] https://www.openstreetmap.org/ - Accessed 2019-11-08

and not slow the simulation down unnecessarily. Secondly, several junctions were not fully connected or not connected properly. These have been adjusted manually. Thirdly, some roads that are dead ends did not allow for U-turns, meaning vehicles would get stuck in simulations. Something similar happened at two-lane roundabouts, where the inner lane had no connection to the outer lane, so cars would get stuck going around in circles on the inner lane. These connections have been added manually.

After these adjustments, the simulation was tested with random routes for an extended period of time, and further adjustments were made while necessary, until no vehicles would get stuck in the simulation any more.

6.4 Parking Spaces

As discussed in Section 5.1, the location of all parking spaces in the city center of Kortrijk were provided as GIS data. No tool exists to translate this data to a usable format within the SUMO environment. Therefore, the parking spaces were placed manually. They were each given a name in the format "Zone_Street_ID" in order to distinguish them within TraCI. The parking spaces are defined as ParkingArea objects within SUMO, each with a capacity of one. Their occupation status can be requested through TraCI. In total, 4443 on-street parking spaces were placed in the simulation. These are split into 778 Shop&Go spaces, 1354 spaces in zone KOR1, 1675 spaces in zone KOR2 and 636 in the blue zone.

6.5 Routing

SUMO is designed to use pre-generated routes or traffic flows, in which stops have to be explicitly set. By using TraCI, stops can be set dynamically after a vehicle is generated. The routing of vehicles in the simulation is performed as follows.

When a vehicle is generated in the simulation, it has to have an assigned route. To this extent, ten routes are defined. These routes start at one of the ten extreme edges in the road network, and end at the same edge. The route in between the edges is computed dynamically via the Dijkstra routing algorithm [23], which is the standard routing algorithm in SUMO[12]. The generated vehicle is given a category (KOR1, KOR2, Blue, Shop&Go) based on the arrival rates discussed in Chapter 5. Then, a parking space from that category is randomly chosen from the full set of possible parking spaces. The vehicle is then rerouted to that parking space. This means the vehicle will enter the simulation on one of the extreme edges, as if it were driving into the simulation environment, drive to its assigned parking space, stay there for an amount of time that is generated via the relevant distribution, discussed in Section 5.2, and then drive back to its entry edge and leave the simulation there. Figure 8 shows an screenshot of the simulation, with multiple parked cars and cars on their way to their destination.

[12] https://sumo.dlr.de/docs/Simulation/Rerouter.html#rerouter - Accessed 2019-11-21

Figure 8: Detail of the running simulation

There is the possibility that the assigned parking space is already taken when the vehicle arrives. If this is the case, the rerouting algorithm is engaged. First, it is checked whether there is a free parking space in the street that allows the parking duration the vehicle is planning to stay. For instance, if the vehicle has a set parking duration of 1550 seconds in a KOR1 parking space, it can also park in a Shop&Go space in the same street, and vice versa. Would the set parking duration be 3600 seconds, it can only park in KOR1, KOR2 or Blue zones. When a vacant space that satisfies the parking duration is found, the vehicle reroutes to this parking space and will attempt to park there.

If there is no vacant parking space on the original destination edge, the vehicle will choose a random edge from the set of connected edges to its current position. If a vacant parking space on that edge is found, the vehicle will drive to that edge and park there. Otherwise, it will repeat the process recursively from the previously chosen edge. If a vacant parking space is not found within 100 visited edges, the vehicle is assumed to be stuck and will leave the simulation. To ensure the total amount of parking events mirrors reality, a new vehicle of the same type is generated with a new destination. If a vacant parking space that is found during cruising is already occupied when the vehicle arrives, the process will start anew.

6.6 Runtime

Two concurrent runs of the simulation were performed, each simulating 40 week-days. SUMO and TraCI are not optimised for multi-core processing, and run on a single core. The runtime on an i5-4300U @ 2.5GHz dual core processor was 29 hours and 5 minutes, for an average of 43 minutes and 37 seconds per simulated day. The runtime on an i7-2600k @ 4.5GHz quad core processor was 20 hours and 54 minutes, for an average of 31 minutes and 21 seconds per simulated day. The majority of the simulation time is spent on the recursive rerouting algorithm, which has been set to only run once per 60 simulation steps, i.e. once per simulated minute, in order to keep the runtime of the simulation reasonable. This does mean that traffic may get stuck behind a car that is waiting to be rerouted for up to 59 seconds, reducing the total traffic throughput. 1,9% of all vehicles encountered the problem of being stuck cruising, as described in the paragraph above.

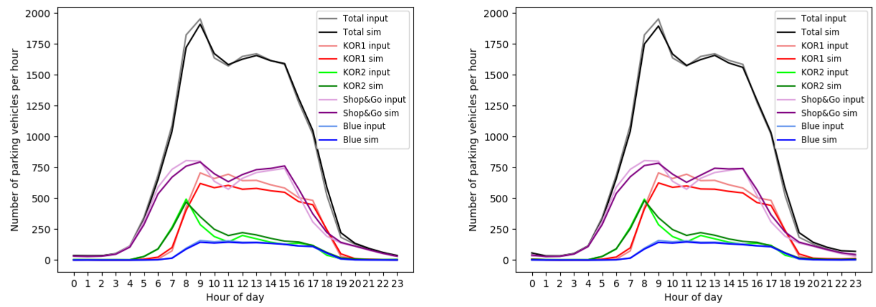
6.7 Validation

Without actual occupancy data available for the majority of the parking spaces in Kortrijk, validating the simulation is difficult. The output of the simulation can be verified by comparing the amount of duration of the simulated parking events to the specified input described in Section 5.2. The Shop&Go dataset does offer the opportunity to validate the simulation by comparing the simulated occupancy and vacancy percentages to the actual occupancy and vacancy percentages for different sets of parking spaces.

To validate the amount of parking events per zone and their arrival times, two simulation runs of 40 weekdays were conducted, together totalling to 16 weeks of weekdays. Figures 9a and 9b show the average amount of vehicles starting to park at each hour of the day, divided into the four vehicle categories. The light colours show the input, and the dark colours show the simulation output. Two things can be noted. Firstly and most apparently, the simulated occupancy of parking events in KOR1 parking spaces is lower than the input between 09:00 and 18:00, whereas it is higher in KOR2 and Shop&Go spaces. This is explained by the fact that the cruising algorithm explained in Section 6.5 allows vehicles that find their assigned parking space occupied to reroute to a parking space of a different type if it fits the vehicle's parking duration criterion. These vehicles that were originally assigned a KOR1 parking space can relocate to a Shop&Go parking space or a KOR2 parking space. Secondly, the simulated values appear to lag behind the input values a little bit. This is explained by the fact that the vehicles do not park immediately at the moment they are generated by the input function. After entering the simulation, they take some time to drive to their parking space, meaning all parking events take place later than the generation of the vehicles.

Figure 10 shows histograms for the simulated parking durations per vehicle category, for a simulated run of 40 weekdays. Clearly visible are the vehicles that stay the full 120 minutes in zone KOR1 and the blue zone and the vehicles that stay until the end of the day in zone KOR2. These histograms are synchronous to the input distributions specified in Section 5.2.

As the variance in the occupancy of single parking spaces is enormous, the validation was performed on groups of Shop&Go parking spaces that are close to each other. Figure 11 shows both the simulated and actual occupancy of the 17 Shop&Go parking spaces on Minister Tacklaan, located south of the train tracks. Although the difference in absolute occupancy at any given moment



(a) Simulation run one (b) Simulation run two

Figure 9: Validation of amount and arrival times of parking events

may differ a lot, the total amount of occupied spaces per minute over the entire day is very similar, at 1,87 and 1,96 for the simulated environment and the sensor readings from the database, respectively.

When comparing the simulation results to real sensor readings, e.g. in Figures 11 and 12, it is clear that the variance in occupancy and thus the parking behaviour is accurately modeled. As the total occupancy percentage of the parking spaces as a factor of time in the simulation also mirrors reality and the distribution of parking durations mirrors the distributions found in Section 5.2, the simulation is realistic enough to perform the intended experiments.

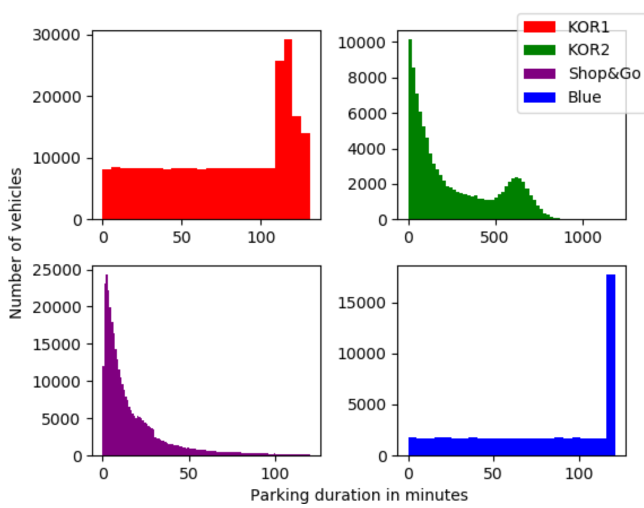


Figure 10: Histograms of simulated parking durations per zone

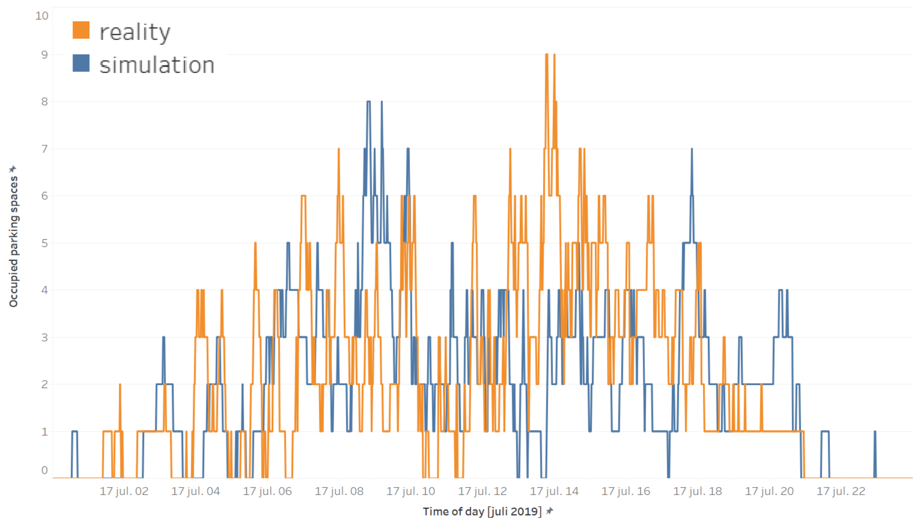


Figure 11: One day of parking occupancy of Shop&Go spaces located on Minister Tacklaan

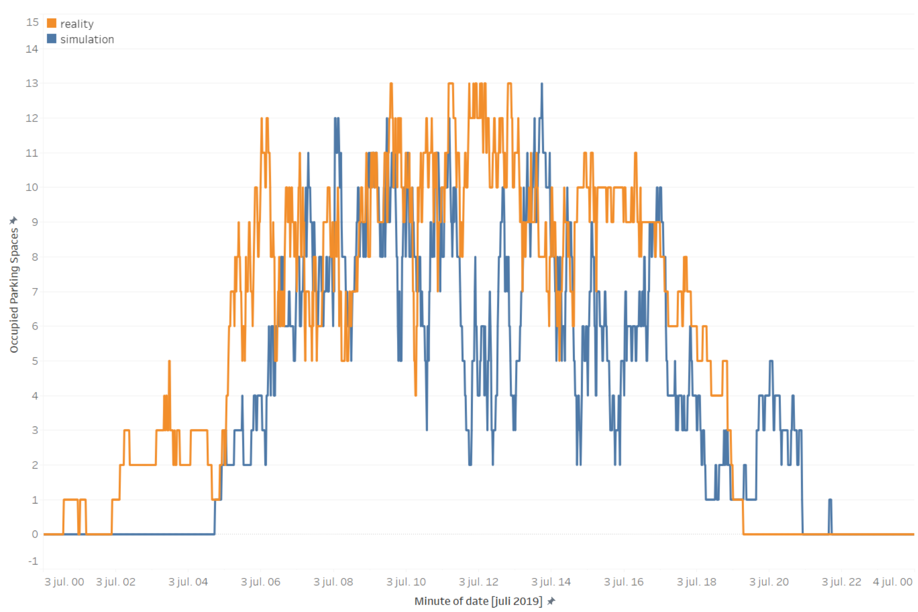


Figure 12: One day of parking occupancy of Shop&Go spaces located on Doorniksestraat

6.8 Scope

The scope of the simulation model is intentionally limited. It focuses on on-street parking exclusively, omitting all parking events in parking garages and closed-off open surface parking lots. In a series of papers, Arnott et al. have shown that on-street parking and off-street parking are not perfect substitutes for one another [5, 8, 9]. In fact, Kobus et al. show that drivers are willing to pay a premium for on-street parking as opposed to garage parking [31]. The data that was available for this research is implicitly influenced by the availability of parking garages and their occupancy, and the simulated amount of vehicles parking on-street mirrors reality even in the absence of simulated parking garages. As this research is focused on trying to reduce the number of sensors for on-street parking occupancy prediction, off-street parking was left out of the scope.

The simulation omits not only the parking that happens in these off-street parking garages and lots, but also the traffic that that would generate. Furthermore, traffic that parks in a private parking space or garage, as well as the through traffic that does not park in the city center but only drives through it, are not simulated. For the latter category, most of this traffic would stick to the ring road anyway, diminishing its effect. However, overall, the vehicles that are simulated for on-street parking will suffer from less congestion than in real life, as some categories of vehicles are not present in the simulation. As the main goal of the study is not to study traffic patterns, but parking events, this is not seen as a problem. The same is true for all other forms of transport, including but not limited to trains, trams, buses, cyclists and pedestrians. These will all have an impact on the traffic situation, but are left out of the scope of the simulation for the same reason.

There is a near infinite amount of reasons a parking space could be inaccessible for a vehicle. Examples one could think of are delivery trucks, trash containers or bicycles taking up that space, a road or road segment being blocked off for construction, an accident or an on-street event. Simulating all this is nigh impossible, but mostly not of enough importance. By limiting the scope to only simulating the vehicles that park in the city center, the research and the experiments stay manageable at very little cost to the simulation accuracy.

Even with the limited scope, the simulation offers something no real-world data set can. It outputs 100% accurate and reliable parking occupancy data for 100% of the parking spaces in the city center. This offers great opportunities for prediction and experiments, and will allow the main research question to be answered.

7 Prediction Algorithm

In order to gauge the impact of the variations that will be implemented during the experiments fairly, a baseline prediction has to be developed of which the performance is similar to that of the state-of-the-art prediction methods. The following chapter describes the architecture of the chosen baseline prediction model and its performance.

7.1 Architecture

Predicting the occupancy of a set of parking spaces is, in essence, a time series forecasting problem. There are many different approaches that work with time series, but not all can be applied to this specific problem. It has a very specific seasonality, with next to no relation between a certain day and the next. While the parking occupancy will correlate strongly from day to day, the occupancy throughout any day does not affect the occupancy of other days in any way.

Recurrent Neural Networks have certain properties that make them very applicable to the planned experiments in this research. With RNN's, it is easily possible to make predictions when removing certain input data, in this case the readings from specific sensors. It is also possible to add another category of sensor data, such as traffic counts. There are multiple examples of parking prediction using RNN's with an LSTM architecture in the literature, which are discussed in Section 2.4. The Long Short-Term Memory architecture for RNN's has feedback connections which help capture the temporal features in a data set, which is very relevant for parking data. Furthermore, it is specifically designed to work with sequence data, which is a natural way to represent time series. For these reasons, the decision was made to use Recurrent Neural Networks with an LSTM architecture.

Predicting the occupancy of a single parking space is neither easy to do nor very useful. Predicting the total occupancy of a set of parking spaces that are geographically close to each other is much more achievable and useful. In the literature, these clusters of parking spaces differ in size, from the block level of 8 - 30 parking spaces per cluster [66], to the regional level with more than 100 parking spaces per cluster [59]. For the baseline prediction model, the predictions are made and evaluated on the street level. The parking spaces were grouped based on the street on which they are physically located. This resulted in 133 sets of parking spaces. The smallest set has 2 parking spaces, whereas the largest has 147. The sets have an average of 33.41 parking spaces. The impacts this has on the performance of the prediction are discussed in detail in Section 9.6.

To prepare the data for training, a 40-day simulation output was converted to data sets with a value for each sensor for each minute, either a 1 if occupied, or a 0 if unoccupied. The 40 days of simulation are split into a training set of 30 days, and a test set of 10 days. This means the training set for each group of parking spaces has 43200 data points, one for each minute in those 30 days. Initially the networks were trained on all sensor data, being a sequence of 1's and 0's for each

minute. However, the networks were found to perform better on simpler input, when training on aggregated sensor totals. Therefore, the input was converted to the sum of the sensors per prediction set. Additionally, the time of the day was given as input, to help the network train the aforementioned seasonality in the data. To capture the cyclical nature of time, the time of day was converted to a two-dimensional feature using a sine and cosine transformation as described here[13]. These transformations are as follows, with *sin_time* and *cos_time* being the two features used as input for the network, *minutes* being the amount of minutes past midnight at the time that is being transformed, and *minutes_in_day* being the amount of minutes in a day, 1440.
$sin\_time = sin(2 * \pi * minutes/minutes\_in\_day)$
$cos\_time = cos(2 * \pi * minutes/minutes\_in\_day)$

When inputting the raw number of minutes that have passed in a day as a feature in the neural network, it seems like 23:55 (1435 minutes past midnight) and 0:05 (5 minutes past midnight) are 23 hours and 50 minutes apart, whereas in reality they are only 10 minutes apart. Using only a sine transformation solves this problem and makes the time cyclical. However, it does result in the same value for two different times of the day. The second dimension of the cosine transformation solves this problem.

Parameter	Value
Input minutes	10
Output minutes	30
LSTM layers	1
LSTM layer neurons	20
Optimiser	Adam
Learning rate	0.001
Epochs	50
Early Stopping Patience	3

Table 3: Prediction network hyperparameters

Each of the 133 prediction areas is unique as far as its composition of parking spaces is concerned. Prediction areas can have multiple types of parking spaces in them, making a generic prediction model for all prediction areas very ineffective. Therefore, an individual model for each of these 133 prediction area has to be trained. Given this fact, and recognising that an order of magnitude more models will be trained during the experiments, it is important to find a neural network architecture that performs reasonably well and is trainable in a reasonable time frame. The hyperparameters of the network were set using a grid search approach, where all combinations of a set amount of hyperparameter settings were considered. Table 3 shows the settings that proved to perform the best.

[13] https://ianlondon.github.io/blog/encoding-cyclical-features-24hour-time/ - Accessed 2019-11-08

With these hyperparameters, training models for all prediction sets takes between 4 and 8 hours on a workstation with an i7 2600k @ 4.5GHz, 16GB DDR3 RAM and a mildly overclocked RTX2070 GPU. Given 100% sensor availability, training times are towards the lower end of that spectrum, increasing as the sensor penetration rate decreases and the uncertainty increases. The network uses an input of the last 10 minutes to predict the next 30 minutes. While 10 minutes is not a lot, no significant improvements were made by increasing this up to 60 or even 240 minutes, while training times did increase. A prediction horizon of up to 30 minutes is the standard in parking prediction literature.

A network architecture with a single LSTM layer of size 20 followed by a dense layer of size 30, as there are 30 minutes ahead to predict, was found to perform the best. Adding more LSTM layers, even when accompanied by dropout layers to prevent overfitting, did not improve the network performance. Yang et al. also find that fewer layers work better when training LSTM networks on parking occupancy data [66]. Increasing or decreasing the size of the LSTM layers did not improve the performance either. The networks are trained with the Adam [30] optimiser for 50 epochs, with an early stop if there is no improvement in 3 epochs, to prevent overfitting and reduce runtime. In each epoch, the entire training set of 30 days is passed through the network once. Figure 13 shows the training and validation losses of one full set of 133 prediction models being trained. Each blue line represents the loss value on the training dataset and has a corresponding orange line that represents the loss value on the test dataset. The point at which these lines terminate shows the early stopping point that was used based on the early stopping patience value described above.

7.2 Performance

To assess the performance of the baseline prediction, it is compared to two naive prediction methods, as well as a similar experiment in literature. The two naive prediction methods are the historical average and the value of the last observation. In the former, the predicted value is the historical average over the training set for that time of the day, rounded to a whole number. In the latter, the predicted value for t+x is the observation at t. Furthermore, the baseline prediction method is evaluated against the best performing LSTM from the 2019 Yang et al. paper [66].

In comparison with the two naive methods, the proposed baseline significantly outperforms them at a prediction horizon of 30 minutes, as shown in Table 4. Figure 14 shows the performance of the proposed baseline and the two naive methods at prediction horizons of 5, 10, 15, 20, 25 and 30 minutes.

When compared to aforementioned best performing LSTM by Yang et al., a lower Mean Absolute Error and higher Mean Absolute Percentage Error can be observed. As Yang et al. state, "In general, lower prediction errors are received on blocks with larger parking capacities" [66]. Out of 133 prediction sets, 58 have under 20 parking spaces and perform worse than sets with more parking spaces, relatively. Section 9.6 goes deeper into the impact of prediction area size.

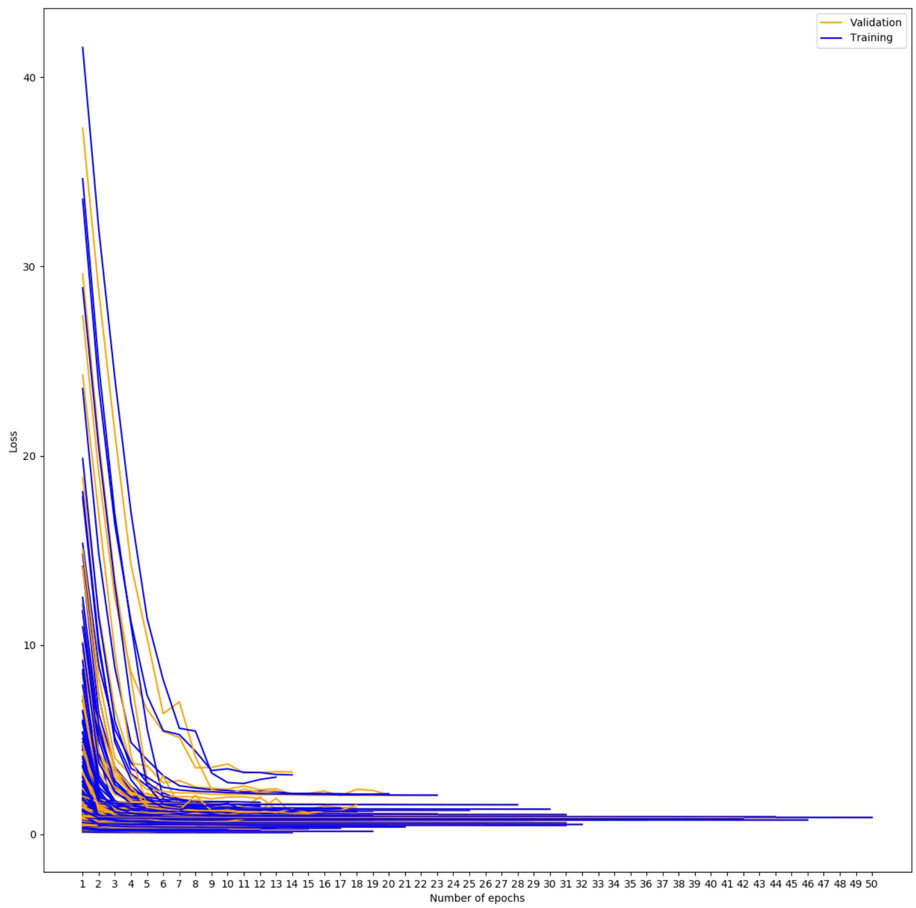


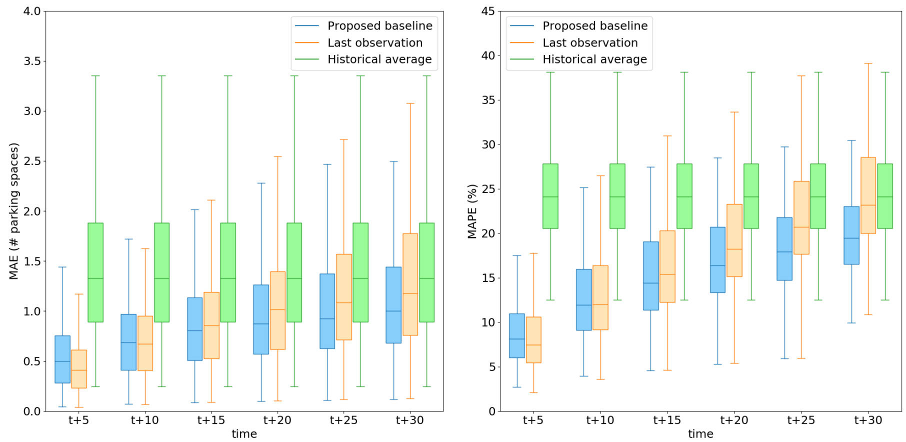
Figure 13: Model training result

Besides the average error of the prediction, an important performance metric is the performance near the occupancy limit. Table 5 shows confusion matrices of the baseline and two naive predictions over the 133 test sets of 14400 minutes. The proposed baseline has a false negative rate of 29.59%, where it predicts there will be a free space even though parking on that street is saturated. It outperforms both naive prediction methods on this metric. When analysing these false negatives in-depth, they naturally only happen on streets where the occupancy during the day is very close to the amount of parking spaces. During the peak parking hours, these streets have an occupancy level of over 98%.

These situations close to full saturation is also where false positives occur, where there is a vacant parking space, but the prediction result is that there are none. Figure 16 shows the relationship between the occurrence of false predictions and the occupancy of a parking area during peak parking hours,

Model	Test MAE	Test MAPE
Proposed baseline	1.18	19.96%
Last observation	1.43	24.34%
Historical average	1.44	24.50%
2-layer LSTM [66]	1.87	14.6%

Table 4: Baseline performance comparison at t+30



(a) Absolute

(b) Relative

Figure 14: Comparison of baseline prediction to naive prediction methods

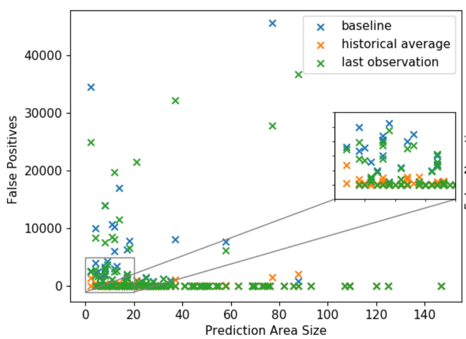
from 09:00 - 19:00. False positives happen a small percentage of the time, as most streets spend most of their time not close to full occupancy. Both false positives and false negatives occur more in prediction sets with a high occupancy percentage, and in prediction sets with a small amount of parking spaces. These correlations will be further discussed in Sections 9.4 and 9.6, respectively.

False positives are less problematic than false negatives. When directing traffic based on a false positive, parking lots may not be fully utilised. On the other hand, when directing traffic based on a false negative, vehicles will not be able to find parking in the street they are directed to. Combating these false negatives is further discussed in Section 10.2.

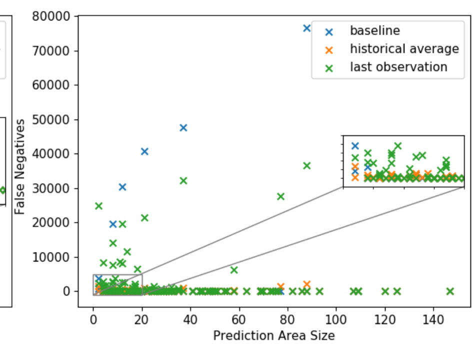
Overall, outperforming the naive predictions and approaching the performance of prediction models in the literature, the proposed baseline performs well enough to gauge the impact of the experiments that are described in Chapter 8.

			Actual value	
			Saturated	Not saturated
Predicted value	Proposed baseline	Saturated	70.41%	0.39%
		Not saturated	29.59%	99.61%
	Last observation	Saturated	63.01%	0.49%
		Not saturated	36.99%	99.51%
	Historical Average	Saturated	47.86%	0.17%
		Not saturated	52.14%	99.83%

Table 5: Baseline confusion matrix

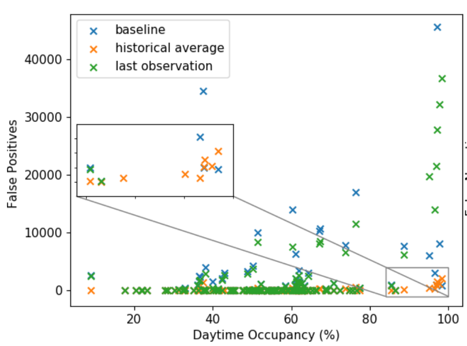


(a) False Positives

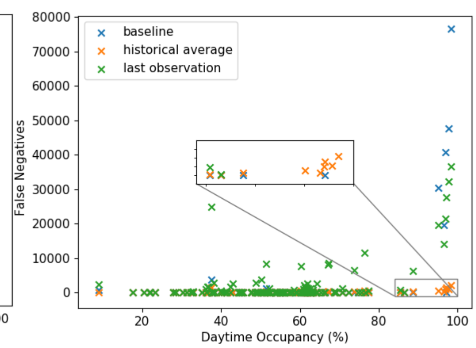


(b) False Negatives

Figure 15: False positives and negatives compared to prediction area size



(a) False Positives



(b) False Negatives

Figure 16: False positives and negatives compared to occupancy

8 Experiments

With a dataset generated in Chapter 6 and a baseline prediction method developed in Chapter 7, the following chapter will discuss the setup of the performed experiments concerning that generated dataset and the developed prediction method.

8.1 Sensor Penetration Rate

As part of the main research question, the sensor penetration rate will be varied. To achieve this, a percentage of the sensors will not be used when training the prediction models, as if they were switched off or not there in the first place. Each sensor has an equal probability of being disabled for this experiment. This experiment will be performed with a coverage of 50%, 25%, and 10%, and evaluated against the baseline of 100% coverage.

8.2 Geographical Distribution

To find an optimal geographical distribution of sensors, the most important question is what parking metrics influence the effectiveness of a sensor on the prediction. To explore this, models will be trained using sets of sensors with the longest parking duration, shortest parking duration, most parking events and least parking events (i.e. highest and lowest turnover, respectively). The coverage level will be chosen based on the results of the experiment described in Section 8.1. These models will then be evaluated against the models from the aforementioned experiment, that have an uniform geographical distribution.

8.3 Spatial Correlation

The baseline prediction method only uses data from the set of parking spaces it is trying to forecast. Adding data from other parking areas that are close by geographically may improve the accuracy of the prediction. Multiple cases will be hand picked based on the road network topology. These will be evaluated against the baseline models. These cases will also be tested with different sensor penetration rates, including having no sensors in the prediction area, but only using those from the areas around the prediction area.

8.4 Parking Pressure

To assess the impact of the amount of parking vehicles on the prediction accuracy, the simulation will be re-run with a modifier on the amount of generated vehicles. Both half the original amount, and 1.5 times the original amount will be tested. Then, new models will be trained according to the baseline method, which are evaluated against the original baseline models.

8.5 Induction Loops

SUMO offers the ability to count traffic passing specific points in the road network using simulated induction loops. The road network of the simulation will be modified to accommodate these induction loops near the parking areas. This data will be used as additional training data for the prediction models, whose performance will be evaluated against the regular baseline models. This experiment will be repeated with different parking sensor penetration rates, and evaluated against the models from the experiment described in Section 8.1

8.6 Prediction Area Size

Finally, variations regarding the prediction area size will be made. As discussed in Section 7.2, the amount of parking spaces in a prediction area is relevant for the performance of the prediction model for that area. Smaller, geographically co-located prediction areas will be combined into one data set for which one model will be trained and one occupancy value will be predicted. The performance of that model will be evaluated against the sum of the individual models. This experiment will also be repeated for different parking sensor penetration rates.

9 Results

This chapter will discuss the results of the experiments stated in Chapter 8. Throughout this chapter, several figures depict box plot graphs. The elements in each box plot are the 133 different prediction areas described in Chapter 7. Each prediction area has its own MAE, measured in amount of parking spaces, and MAPE. The box plots show the minimum, 25^{th} percentile, median, 75^{th} percentile and maximum of these errors. The average errors are given in the text and the accompanying tables of results.

9.1 Sensor Penetration Rate

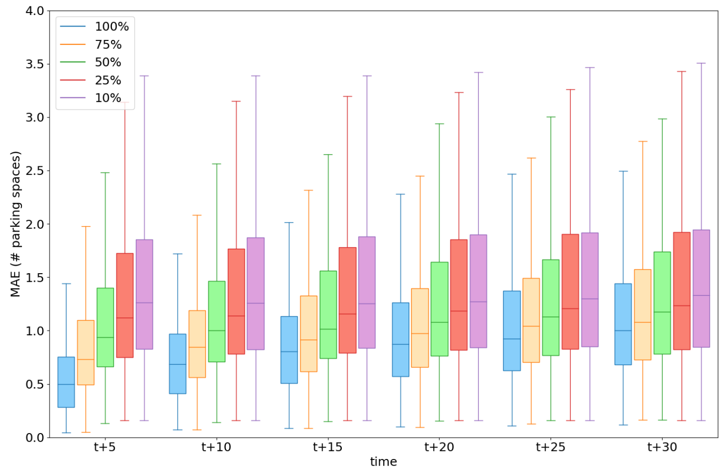
Figure 17 shows the absolute and relative errors of the models with different penetration rates. Performance quickly drops off as the number of sensors is decreased, with the sets under 50% coverage not outperforming the naive prediction methods on a 30 minute prediction horizon, as can be seen in Table 6. On a shorter prediction horizon, all models expect those with 10% coverage outperform the historical average, while being outperformed by the last observation naive prediction method.

Table 6 shows the average performance of the models on a prediction horizon of 30 minutes. When lowering the sensor coverage, the occurrence of false negatives rises quickly. Curiously, at the lowest coverage level, the amount of false negatives decreases again. Given the relatively low amount of prediction areas where false negatives occur in the first place, this is likely to be an artifact of the imperfect training progress. Overall, a linear reduction in error rates can be seen as the coverage level is increased. Figure 18 shows this linear relationship between the coverage level and the error rate.

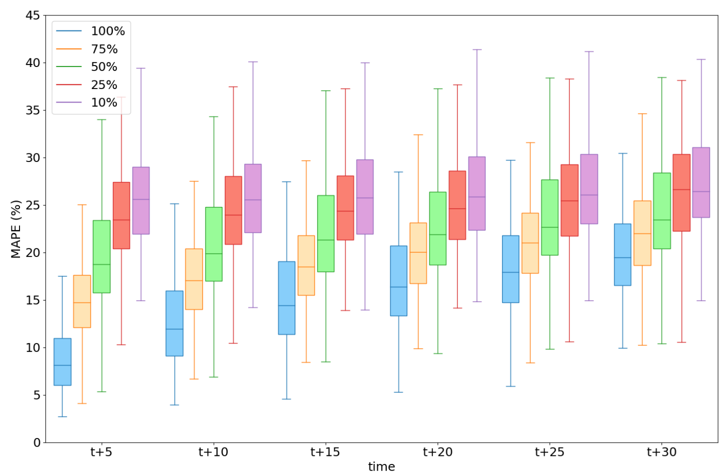
Important to note about these naive prediction methods is that they are generated from sensor data with 100% coverage. It is impossible to take the value of the last observation without knowing the exact amount of parked vehicles at any point. The historical average can be generated from parking meter ticket sales, but this will bring an extra degree of uncertainty as there are varying degrees of people underpaying, overpaying and parking without a ticket, as discussed earlier in Section 5.2.1.

Model	Test MAE	Test MAPE	False Positives	False Negatives
100% coverage	1.18	19.96%	0.39%	29.59%
75% coverage	1.28	22.38%	0.35%	33.22%
50% coverage	1.38	24.36%	0.32%	41.22%
25% coverage	1.46	26.57%	0.37%	42.24%
10% coverage	1.52	27.49%	0.31%	35.24%
Last observation	1.44	24.50%	0.49%	36.99%
Historical average	1.43	24.34%	0.17%	47.86%

Table 6: Performance of different penetration rates at t+30



(a) Absolute



(b) Relative

Figure 17: Comparison of performance of different sensor penetration rates

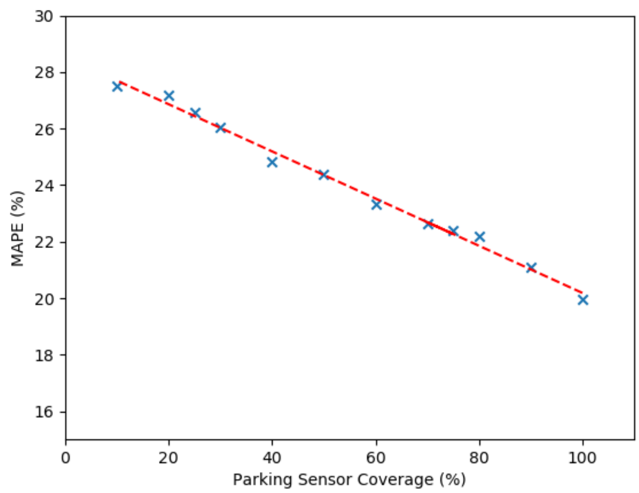


Figure 18: Relationship between coverage level and error rate at t+30

9.2 Geographical Distribution

Models were trained and tested using 25 and 50% of the sensors that experienced the longest parking duration, the shortest parking duration, the highest turnover and the lowest turnover. None of these experiments resulted in an improvement on the average prediction accuracy of all streets. The performance of each set of trained models was very close to the baseline of that sensor coverage level from Section 9.1. The prediction areas which had the most parking spaces inside the experiment set of parking spaces, whether they were high or low on turnover or duration, performed better than the baseline, and the prediction areas with fewer parking spaces performed worse. The geographical distribution has a minimal effect, but the penetration rate of sensors has a profound effect. The best all-round performance is found using a uniform distribution of sensors over a city.

The question how sensors should be distributed geographically within a prediction area cannot be answered, as the utilisation of parking spaces within each prediction area is uniform in the simulation. The geographical distribution within a prediction area will heavily rely on the destinations within that area of the occupants of the parking vehicles. This requires either a very microscopic simulation or a real-world case study, and is left to future research.

9.3 Spatial Correlation

In the simulated dataset, no situations were found where adding data from parking sensors on adjacent streets improved the performance of the prediction model, at no sensor penetration level. When using zero sensors in the prediction area, and information from 100% sensor coverage in the adjacent streets, the models were not able to outperform the historical average prediction.

Other work in literature is able to improve on their predictions by adding spatially correlated data such as parking meter transactions and traffic speed on adjacent roads [66]. The data to replicate that approach is not present in the case study of Kortrijk, but it proves that improvements can be made using spatially correlated data.

9.4 Parking Pressure

Simulations were run using the parking totals found in Chapter 5 modified with a factor 0.5 and 1.5. The simulation with factor 1.5 had a run time of over two weeks, due to excessive clogging of the road network. When analysing the output, it became apparent that due to the extreme congestion, there were more vehicles that were unable to perform their scheduled parking stop than vehicles that performed their scheduled stop. Nevertheless, insights can be gained from analysing the performance of the prediction algorithm on the output of the simulation run with half parking pressure, and from taking a detailed look at the prediction areas that approach the maximum occupancy level in the original dataset.

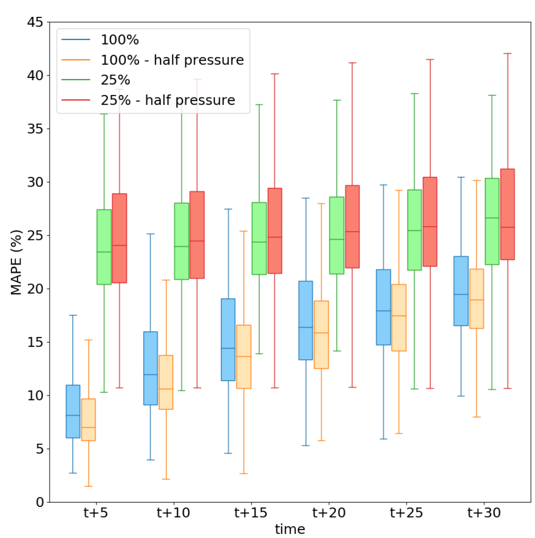


Figure 19: Comparison of performance with half parking pressure

As one would expect, the MAE of the predictions on the dataset with half the regular parking pressure are significantly lower, as the absolute parking totals are lower, and thus the possible deviations from that. When comparing the MAPE, Figure 19 shows that the medians of the mean absolute percentage errors per street are generally lower. However, the averages are very similar, as can be seen in Table 7.

Model	Test MAPE regular	Test MAPE 0.5 pressure
100% coverage	19.96%	19.36%
25% coverage	24.36%	26.53%
Last observation	24.34%	21.61%
Historical average	24.50%	25.07%

Table 7: Performance with lower parking pressure at t+30

When analysing the performance of the prediction method close to the occupancy limit, there is no deterioration of performance when getting closer to saturated parking. Figure 20 shows a scatter plot comparing the model performance per prediction area compared to that area's daytime occupancy level. If anything, a slight improvement can be seen as the occupancy percentage increases, but there is not enough data to draw that conclusion. As discussed in Section 7.2, the absolute amount of false positives does increase when approaching the occupancy limit. However, the relative amount of false negatives as a percentage of the total amount of observations does not, when occupancy levels rise. False negatives are more likely to occur closer to the occupancy limit, but the ratio at which they occur does not increase.

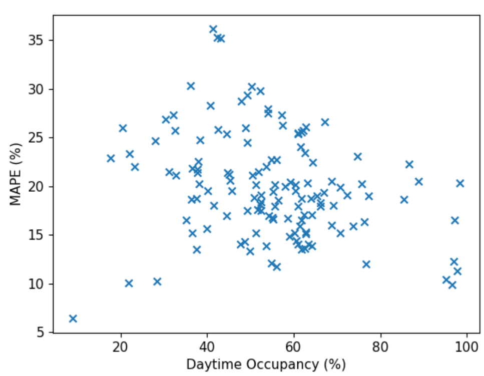


Figure 20: Prediction performance compared to occupancy percentage

9.5 Induction Loops

The road network used for the simulation was edited to include traffic counts for every street segment with parking spaces. Induction loops have been manually placed at the entrances to each street segment with parking spaces, to count the traffic on that space. Figure 21 shows a few street segment with its parking spaces in purple and induction loops in yellow. The traffic counts are then given as an extra variable when training the prediction models.

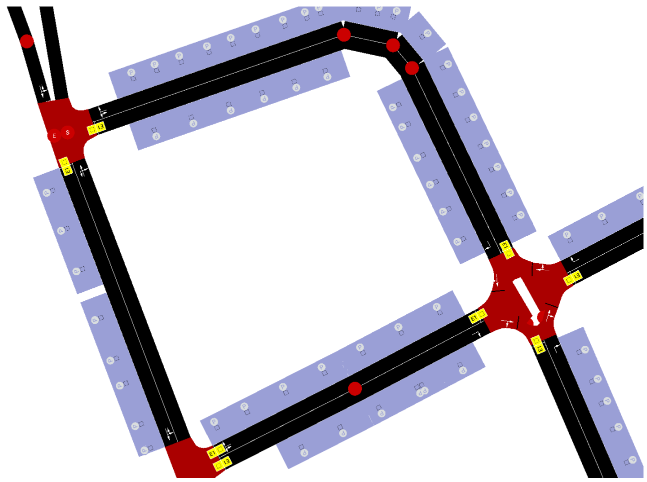


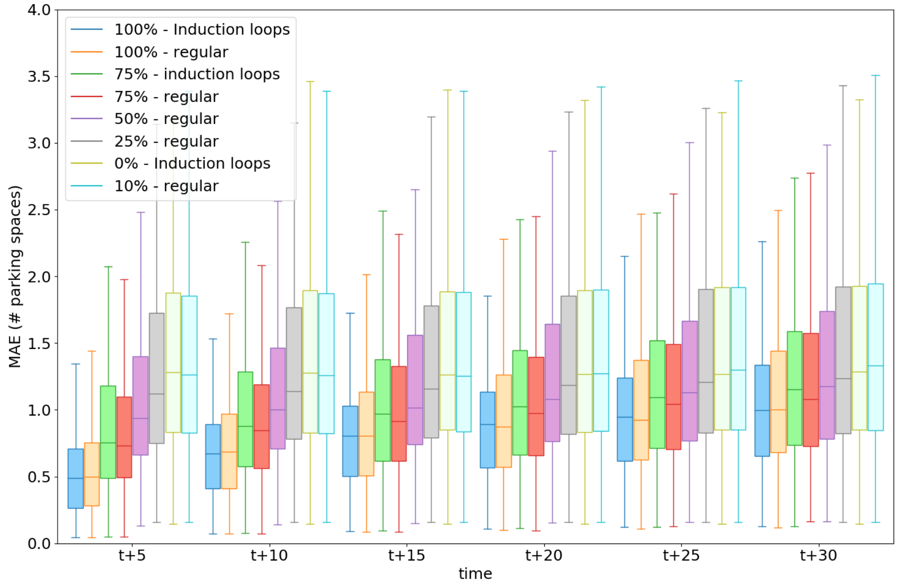
Figure 21: Detail of simulation network including induction loops

This slightly improves model performance, especially at higher sensor coverage levels and at shorter prediction horizons. Even though the median of the set with 75% coverage and induction loops is higher than the set without traffic count data, the MAPE at t+30 is lower at 22.21 versus 22.43. Interestingly, using no parking sensor data but only traffic count data as input for the prediction algorithm leads to better results than using purely the data from a 10% coverage level of parking sensors. Table 8 and Figure 22 show the results of this experiment.

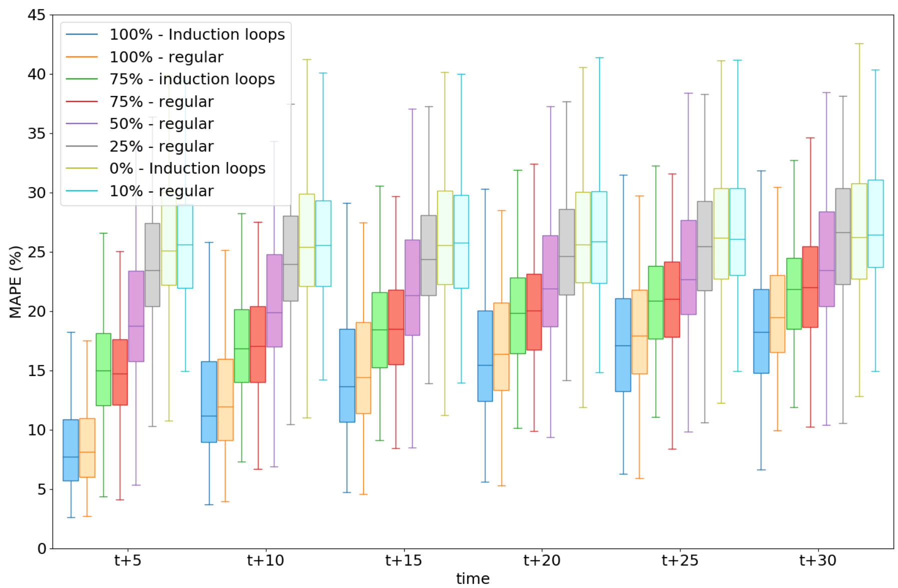
Multiple other scholars report a positive influence on the prediction accuracy when adding data from traffic volume sensors, such as Badii et al. [11] and Yang et al. [66], strengthening the validity of these results.

Model	Test MAE	Test MAPE	False Positives	False Negatives
100% coverage	1.09	18.96%	0.39%	29.59%
75% coverage	1.28	22.21%	0.33%	41.20%
0% coverage	1.54	26.84%	0.35%	38.26%

Table 8: Performance of models with traffic counts at t+30



(a) Absolute



(b) Relative

Figure 22: Comparison of performance of different sensor penetration rates with and without traffic counts

9.6 Prediction Area Size

To assess the impact of the prediction area size, combinations of adjacent streets were made for which a single prediction model was trained. While there are quite a few streets with fewer than 10 parking spaces, these are not often geographically close to each other. One example of adjacent small parking areas is at the station. When combining 24 sensors on three streets into a single model instead of three, a small improvement in prediction performance is found on longer prediction horizons. This is illustrated in Figure 23. It shows the mean absolute error of the prediction models for the three streets individually, those seperate models combined in purple, and the newly trained combined model in orange. The improvement is minimal, and disappears when reducing the sensor coverage level.

When combining sets of larger parking spaces, the prediction accuracy suffers. Training models for larger prediction sets is more difficult, and the architecture described in Section 7.1 is not sufficient to train on these large prediction areas. The error rates go up when the prediction area size exceeds 200 parking spaces. The lack of geographical correlation between the parking spaces may play a role in this.

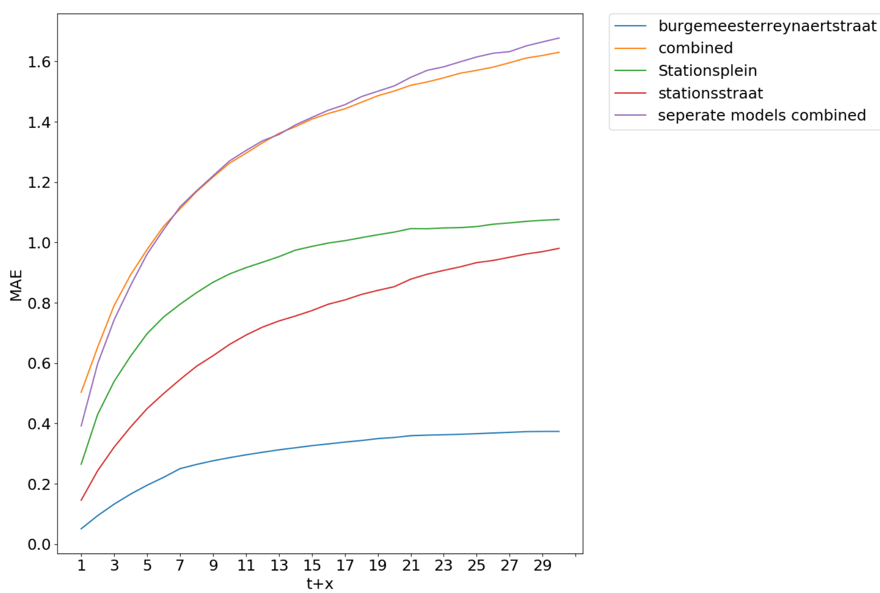


Figure 23: Prediction accuracy for combined prediction areas at the station

When looking at the average error as a function of the prediction area size, such as in Figure 24, it becomes apparent that very small prediction areas are detrimental for the prediction accuracy. Furthermore, as shown in Section 7.2, small prediction areas have a much higher rate of false negatives. Combining smaller prediction areas that are inherently prone to false negatives can thus reduce the rate at which false negatives occur, at the expense of some geographical accuracy, as the prediction is effectively for a larger area.

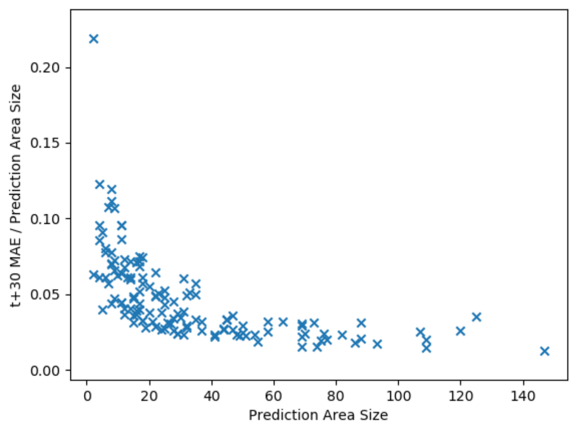


Figure 24: Prediction accuracy compared to prediction area size

9.7 Round-up

When considering the deployment of a smart parking system in a city center, there are a lot of factors to consider. The higher the amount of sensors you install, the more accurate the prediction algorithm you can develop using the gathered data will be. The performance remains constant with varying degrees of parking pressure and parking space utilisation. With a coverage of under 25% and not using any additional data sources, the prediction models are unable to outperform two naive prediction methods. Of these naive prediction methods however, the last observation method relies on 100% sensor coverage, and while the historical average approach can be approximated using different data sources such as parking ticket sales, it will be less accurate.

The prediction method performs best with a uniform geographical distribution of sensors over the study area. Additional data sources can improve the prediction accuracy, compared to exclusively using the sensor data within the prediction area. A positive impact was found when enriching the parking sensor data with traffic count data from induction loops in the road network. In a real-life smart parking deployment, other data sources such as floating cellular data or surveillance cameras can be used to substitute the induction loops. In this case study, no positive impact was found when enriching the data with data from adjacent parking areas. However, findings from literature indicate improvements can be made to the prediction by using a combination of data sources regarding the adjacent road network, such as traffic speed, traffic volume and the amount of parked vehicles.

The size of the prediction areas is very relevant for the performance of the prediction. Small prediction areas lead to relatively higher error rates and more false negatives. Combining adjacent areas to create areas of at least 25 parking spaces will reduce the amount of false negatives at the expense of a small bit of geographical accuracy. Conversely, there are diminishing returns on increasing the prediction area size beyond 50 parking spaces. Thus, prediction areas larger than 100 should be split into multiple prediction areas in order to increase the geographical accuracy of the predictions.

10 Discussion

This chapter will discuss the limitations of the performed research and delve into the real-world implications of this research one has to consider when designing a smart parking system.

10.1 Limitations

The conducted research has multiple limitations. Firstly, the scope of the simulation model is limited. The scope and the limitations to the accuracy of the simulation were extensively discussed in Section 6.8. Furthermore, by using a simulation, the impact of data sources such as the weather conditions, that are found to be relevant for parking prediction, could not be analysed. The impact of the weather on parking has to be quantified in a real-world scenario. Researching the effect of weather on prediction accuracy in a simulation does not make sense without knowing the relationship between weather and parking behaviour. Besides the limitations inherent to the simulation, the research has been conducted as a single case study on a single city. This is a limitation, and the question is how well the results generalise to other cities.

Secondly, the prediction architecture is not fully designed for maximum accuracy, but also takes efficiency and runtime into account. This allowed for the training of many models to conduct many experiments, but does not always achieve the maximum result. The architecture and models can be improved on, but the goal of this research was to explore the impact of the variations and conducted experiments within a reasonable time frame, not to achieve the maximum possible accuracy. For this purpose, the prediction architecture sufficed.

Finally, during the experiments, the limitations of the simulation became apparent. Enriching data sets with spatially correlated data did not improve the prediction accuracy, even though other scholars did find improvements by adding this data. This may be a characteristic of the parking situation in Kortrijk, but it is more likely it is a limitation of the simulation. Furthermore, the question how sensors should be distributed geographically within a prediction area cannot be answered, as the utilisation of parking spaces within each prediction area is uniform in the simulation. Both these limitations can be overcome with a more microscopic simulation where each simulated agent has a specific destination in the city. This can be achieved by using a dedicated population activity model. Finally, the simulation was not able to handle increased parking pressure, leaving one experiment partly unfinished.

10.2 Real-world implications

To fairly assess the real-world implications of the results of this research, one has to consider the costs of a smart parking network. The deployment of a smart parking network is expensive, as each sensor costs between €350 and €400 to purchase, install and connect. This is excluding yearly network costs for their communication. With an estimated lifespan of 10 years and including operating costs and possibly necessary repairs, the costs of outfitting just the 4443 parking spaces in the city center of a small city like Kortrijk with a smart parking system for 10 years already exceed €2,000,000. In larger cities or cities with more parking spaces, costs can rise quickly. Reducing the necessary coverage therefore is a very interesting method to reduce the total costs of the network. Cutting down the sensor penetration rate to 25% can already save €1,500,000 in a small city like Kortrijk as opposed to a system based on 100% sensor coverage.

As shown in the experiments, reducing the sensor coverage leads to an increase in the occurrence of false negatives, where the models predict a vacant parking space, whereas there is no vacancy in reality. Reducing the amount of false negatives is the most important metric in a real-world deployment of a smart parking network, as predicting a false negative means drivers are sent to an area without vacant parking spaces. False negatives occur exclusively near the occupation limit, and with a lower penetration rate, false negative ratios are close to naive prediction methods. To reduce the amount of false negatives, 100% sensor coverage is necessary in the busiest parking areas. To prevent sending drivers to areas without vacant parking spaces, user interfaces for smart parking systems can be adapted to not show a vacancy, not only when the predicted amount of vacancies is 0, but also when it is 1.

Enriching the data with additional data sources such as the weather and traffic counts can improve the accuracy of the predictions, especially when the sensor coverage is reduced. In this research, traffic counts from induction loops were used, but there are other data sources that can provide similar data. Cities may have access to traffic counts based on surveillance cameras, pneumatic traffic counters and/or floating cellular data, which can all play a similar role in enriching the data and improving the prediction accuracy.

When experimenting with a reduced sensor coverage, the models were trained with the ground truth data that was obtained from the simulation. In a real deployment of parking sensors, this ground truth data will not be available. Options include monitoring the prediction area to get aggregated parking counts for training purposes, or simulating the events based on parking ticket sales similar to the simulation conducted in this research. Regardless, this poses a challenge in deploying smart parking systems with less than 100% sensor coverage.

In summary, when deploying parking sensors in a smart parking project with the goal of predicting the occupancy of parking areas, the following guidelines should be followed. Divide the city geographically in prediction areas so that each area has between 25 and 100 parking spaces. Lower numbers will lead to increased false negatives, whereas higher negatives will lead to decreased geographical accuracy. In the areas where the parking utilisation is the highest and the most congested, 100% sensor coverage is necessary to prevent false negatives. In the lesser congested areas that always have vacancies, an approach using historical averages based on parking meter ticket sales data will suffice, as the larger error rate is unlikely to lead to false positives or false negatives. The areas that fall in between those two categories, which sometimes have no vacancies but are not completely saturated daily, can benefit from a smart parking system with a lower sensor coverage. Depending on the budget and the required service level, the sensor coverage can be anywhere from 25% to the full 100%. Less than 25% coverage does not offer enough benefits. For all prediction areas, adding additional data sources will improve the quality of the prediction. Municipalities and cities investing in smart parking are likely to have access to traffic counts, but can also use floating cellular data or other data sources.

11 Conclusion

This thesis has explored an extensive case study on parking prediction in Kortrijk. By combining many data sources, a realistic simulation of the parking situation was developed. Then, an architecture for a prediction algorithm based on long-short term memory recurrent neural networks was developed and its performance on a simulated data set of parking events evaluated. Finally, several experiments were conducted to research how the prediction approach and the simulation behave under different circumstances and the effects that has on the prediction accuracy.

In Section 1.3, five sub-questions were stated. The first question, *What different data sources are used in parking prediction?*, was answered through literature research. The main data sources that are used for parking prediction are historic parking data, live parking data, weather, and traffic counts and traffic speed on the road network close to the prediction area. The second question, *How does the availability of each data source influence the accuracy of the prediction?*, is difficult to answer quantitatively. Different researchers find different values for the impact of the different data sources. It is clear however, that adding more data sources improves the accuracy of the prediction. Most important are historic and live parking data, with the weather, traffic counts and traffic speeds playing a smaller role. The answer to the third research question, *Which prediction algorithms are the state-of-the-art?*, is that it depends heavily on the availability of the data sources. Scholars find a specific machine learning architecture that works best on their data sources. There are many different options that can all work well, and experimentation with different architectures is necessary to find the one that fits the data available for a specific use case.

The fourth research question, *What are the effects of varying the spatial distribution and penetration rate of stationary sensors?* has been partly answered during the experiments. While no best practice for the spatial distribution of the stationary sensors has been found, a clear linear relationship between the coverage level of parking sensors and the error rate of the prediction was found. Furthermore, reducing the coverage level leads to an increased amount of false negatives in the predictions. These observations were also made when varying the demand for parking and the geographical distributions of the sensors.

The final sub-question, *What is the economic trade-off between accuracy and installing stationary sensors for predictive analysis?* ties into the previous question and depends on the required service level of the smart parking system the sensors are deployed in. As a linear relationship between the coverage level and the prediction accuracy has been established and the costs of a smart parking project scale linearly when increasing the number of sensors, there is also a linear relationship between the costs of a smart parking project and its performance. Adding other data sources such as traffic counts or weather data can be a cheaper alternative to increasing the coverage level.

Based on the findings from the experiments, when deploying parking sensors in a smart parking project with the goal of predicting the occupancy of parking areas, the following guidelines should be followed. Divide the city geographically in prediction areas so that each area has between 25 and 100 parking spaces. In the areas where the parking utilisation is the highest and the most congested, 100% sensor coverage is necessary to prevent false negatives. In the lesser congested areas that always have vacancies, an approach using historical averages based on parking meter ticket sales data will suffice, as the larger error rate is unlikely to lead to false positives or false negatives. The areas that fall in between those two categories, which sometimes have no vacancies but are not completely saturated daily, can benefit from a smart parking system with a lower sensor coverage. Depending on the budget and the required service level, the sensor coverage can be anywhere from 25% to the full 100%. Less than 25% coverage does not offer enough benefits. For all prediction areas, adding additional data sources will improve the quality of the prediction. Municipalities and cities investing in smart parking are likely to have access to traffic counts, but can also use floating cellular data or other data sources.

12 Future Research

Throughout the duration of this research, multiple possible directions for future research were discovered, either through a deliberate restriction of the scope of the research, or through limitations of the data set or the taken approach.

Firstly, there are many directions one could go to improve the simulation. The first step is to extend it to include variations between weekdays and simulate weekends. Then, including other forms of traffic and off-street parking will improve the fidelity of the routing and congestion in the simulation. This also means that traffic counts can be more realistically simulated, and floating cellular data can be extracted from the simulation. The routing algorithm used in the simulation can also be improved, to include actual destinations in the network instead of having a certain parking space as a destination. The actual destination can then be used to simulate a more realistic cruising behaviour when the desired parking spot is occupied.

Due to the time and effort required to build the simulation, this research has been conducted as a single case study on one city. Using the same approach on a different city, with a different road network, parking space distribution and parking regulations would be an interesting future research direction to validate how well the results of this research can be generalised. Furthermore, the results of this research can be validated by replicating the experiments on real-world data.

During the data analysis, a serious impact on the occupancy of parking spaces was found by people using handicap placards. Kortrijk has reported handicap placard abuse and changed its policies accordingly [14]. In other cities, handicap placard abuse is likely to play a large role in the parking problem, as also argued by Clinchant et al., Glasnapp et al., Zoeter et al. and Shoup [21, 26, 47, 70]. Research into new policies to combat handicap placard abuse can help alleviate the parking problem in cities.

Related to this is the issue of parking pricing. This topic is discussed in a lot of parking literature, dating back to Vickrey's pioneering work in 1954 [58], with Shoup being the most vocal advocate for dynamic pricing [49, 51, 48]. Pilots using dynamic pricing have been run, most notably in the SFPark project [45]. How real-time dynamic pricing can balance the demand for parking throughout a city supported by parking sensors is a very interesting direction for further research. This would also require research into the political feasibility of such a pricing scheme.

During the experiments, the question how sensors should be distributed geographically within a prediction area could not be answered, as the utilisation of parking spaces within each prediction area is uniform in the simulation. The geographical distribution within a prediction area will heavily rely on the destinations within that area of the occupants of the parking vehicles. This requires either a very microscopic simulation or a real-world case study. A microscopic simulation would also help answer the stated questions about the impact of

[14]https://www.nieuwsblad.be/cnt/blkva_04449304 - Accessed 2019-12-18

combining data from spatially correlated parking areas to improve predictions. No such relation was found in this research, but other scholars do report increased accuracy when using spatially correlated data [66]. This more detailed simulation or a real-world case study is left to future research.

Finally, despite being out of scope for this research, the question of how to actually use parking predictions remains and has been a topic of discussion often. Ideally, navigation applications would have access to predictions based on sensor data and guide their users accordingly. A few of these applications exist, most notably ParkNav[15], but the areas in which they are active are limited, and their userbase and thus impact are small. Designing a method to get parking information and predictions to drivers and autonomous vehicles efficiently is a big challenge and a good direction for future research.

[15] https://www.parknav.com/ - Accessed 2020-04-19

References

[1] G. Amato, F. Carrara, F. Falchi, C. Gennaro, C. Meghini, and C. Vairo. Deep learning for decentralized parking lot occupancy detection. *Expert Systems with Applications*, 72:327–334, apr 2017.

[2] G. Amato, F. Carrara, F. Falchi, C. Gennaro, and C. Vairo. Car parking occupancy detection using smart camera networks and deep learning. In *2016 IEEE Symposium on Computers and Communication (ISCC)*. IEEE, jun 2016.

[3] S. P. Anderson and A. De Palma. The economics of pricing parking. *Journal of Urban Economics*, 55(1):1–20, 2004.

[4] R. Arnott. On the optimal target curbside parking occupancy rate. *Economics of Transportation*, 3(2):133–144, 2014.

[5] R. Arnott and E. Inci. An integrated model of downtown parking and traffic congestion. *Journal of Urban Economics*, 60(3):418–442, 2006.

[6] R. Arnott, E. Inci, and J. Rowse. Downtown curbside parking capacity. *Journal of Urban Economics*, 86:83–97, 2015.

[7] R. Arnott and J. Rowse. Modeling parking. *Journal of urban economics*, 45(1):97–124, 1999.

[8] R. Arnott and J. Rowse. Downtown parking in auto city. *Regional Science and Urban Economics*, 39(1):1–14, 2009.

[9] R. Arnott and J. Rowse. Curbside parking time limits. *Transportation Research Part A: Policy and Practice*, 55:89–110, 2013.

[10] R. Arnott and P. Williams. Cruising for parking around a circle. *Transportation research part B: methodological*, 104:357–375, 2017.

[11] C. Badii, P. Nesi, and I. Paoli. Predicting available parking slots on critical and regular services by exploiting a range of open data. *IEEE Access*, 6:44059–44071, 2018.

[12] J. Bates and D. Leibling. Spaced Out: Perspectives on Parking Policy. http://www.racfoundation.org/assets/rac_foundation/content/downloadables/spaced_out-bates_leibling-jul12.pdf. Accessed: 2019-02-26.

[13] M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz. Sumo–simulation of urban mobility: an overview. In *Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation*. ThinkMind, 2011.

[14] I. Benenson, K. Martens, and S. Birfir. Parkagent: An agent-based model of parking in the city. *Computers, Environment and Urban Systems*, 32(6):431–439, 2008.

[15] F. Bock, S. Di Martino, and A. Origlia. A 2-step approach to improve data-driven parking availability predictions. In *Proceedings of the 10th ACM SIGSPATIAL Workshop on Computational Transportation Science*, pages 13–18. ACM, 2017.

[16] F. Bock, D. Eggert, and M. Sester. On-street parking statistics using lidar mobile mapping. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pages 2812–2818. IEEE, 2015.

[17] K. C. T. Bong, D. B. L. and K. C. Lai. Integrated approach in the design of car park occupancy information system (coins). *IAENG International Journal of Computer Science*, 35(1), 2008.

[18] F. Caicedo, C. Blazquez, and P. Miranda. Prediction of parking space availability in real time. *Expert Systems with Applications*, 39(8):7281–7290, 2012.

[19] A. Camero, J. Toutouh, D. H. Stolfi, and E. Alba. Evolutionary deep learning for car park occupancy prediction in smart cities. In *International Conference on Learning and Intelligent Optimization*, pages 386–401. Springer, 2018.

[20] City of Melbourne. Melbourne Parking Data. https://data.melbourne.vic.gov.au/Transport-Movement/On-street-Parking-Bay-Sensors/vh2v-4nfs. Accessed: 2019-02-25.

[21] S. Clinchant, C. Dance, T. de Ruijter, P. Ghent, and O. Zoeter. Using analytics to understand on-street parking: the impact of special permit use and the benefit of demand based rates over zones. In *22nd ITS World Congress*, 2015.

[22] K. Dieussaert, K. Aerts, T. Steenberghen, S. Maerivoet, and K. Spitaels. Sustapark: an agent-based model for simulating parking search. In *AGILE International Conference on Geographic Information Science, Hannover*, pages 1–11, 2009.

[23] E. W. Dijkstra et al. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959.

[24] S. Evenepoel, J. Van Ooteghem, S. Verbrugge, D. Colle, and M. Pickavet. On-street smart parking networks at a fraction of their cost: performance analysis of a sampling approach. *Transactions on Emerging Telecommunications Technologies*, 25(1):136–149, 2014.

[25] K. Gkolias and E. I. Vlahogianni. Convolutional neural networks for on-street parking space detection in urban networks. *IEEE Transactions on Intelligent Transportation Systems*, 2018.

[26] J. Glasnapp, H. Du, C. Dance, S. Clinchant, A. Pudlin, D. Mitchell, and O. Zoeter. Understanding dynamic pricing for parking in los angeles: Survey and ethnographic results. In *International Conference on HCI in Business*, pages 316–327. Springer, 2014.

[27] S. Hartvelt. Real-time parking information and guidance: Research topics. *University of Twente*, 2019.

[28] H. Ichihashi, A. Notsu, K. Honda, T. Katada, and M. Fujiyoshi. Vacant parking space detector for outdoor parking lot by using surveillance camera and FCM classifier. In *2009 IEEE International Conference on Fuzzy Systems*. IEEE, aug 2009.

[29] A. Ionita, A. Pomp, M. Cochez, T. Meisen, and S. Decker. Where to park?: Predicting free parking spots in unmonitored city areas. In *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics*, page 22. ACM, 2018.

[30] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[31] M. B. Kobus, E. Gutiérrez-i Puigarnau, P. Rietveld, and J. N. Van Ommeren. The on-street parking premium and car drivers' choice between street and garage parking. *Regional Science and Urban Economics*, 43(2):395–403, 2013.

[32] D. Krajzewicz, G. Hertkorn, C. Rössel, and P. Wagner. Sumo (simulation of urban mobility)-an open-source traffic simulation. In *Proceedings of the 4th middle East Symposium on Simulation and Modelling (MESM20002)*, pages 183–187, 2002.

[33] S. Lee, D. Yoon, and A. Ghosh. Intelligent parking lot application using wireless sensor networks. In *CTS*, pages 48–57, 2008.

[34] I. Lendák and K. Farkas. How many drivers does it take to spot an openspot? In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 000401–000406. IEEE, 2016.

[35] N. Levy, K. Martens, and I. Benenson. Exploring cruising using agent-based and analytical models of parking. *Transportmetrica A: Transport Science*, 9(9):773–797, 2013.

[36] T. Lin, H. Rivano, and F. Le Mouël. A survey of smart parking solutions. *IEEE Transactions on Intelligent Transportation Systems*, 18(12):3229–3253, 2017.

[37] Los Angeles Department of Transportation. LA Express Park. http://www.laexpresspark.org/. Accessed: 2019-02-25.

[38] Q. Luo, R. Saigal, R. Hampshire, and X. Wu. A statistical method for parking spaces occupancy detection via automotive radars. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2017.

[39] N. Mejri, M. Ayari, R. Langar, and L. Saidane. Con2pas: A constrained congestion game for parking space assignment. In *2016 International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN)*, pages 1–6. IEEE, 2016.

[40] C. Pflügler, T. Köhn, M. Schreieck, M. Wiesche, and H. Krcmar. Predicting the availability of parking spaces with publicly available data. *Informatik 2016*, 2016.

[41] G. Pierce and D. Shoup. Getting the prices right: an evaluation of pricing parking by demand in san francisco. *Journal of the American Planning Association*, 79(1):67–81, 2013.

[42] T. Rajabioun and P. A. Ioannou. On-street and off-street parking availability prediction using multivariate spatiotemporal models. *IEEE Transactions on Intelligent Transportation Systems*, 16(5):2913–2924, 2015.

[43] C. Roman, R. Liao, P. Ball, S. Ou, and M. de Heaver. Detecting on-street parking spaces in smart cities: performance evaluation of fixed and mobile sensing systems. *IEEE Transactions on Intelligent Transportation Systems*, 19(7):2234–2245, 2018.

[44] S. Rupert. Automobile garage, May 8 1923. US Patent 1,454,348.

[45] San Fransicso Municipal Transportation Agency. SFpark. http://sfpark.org. Accessed: 2019-02-25.

[46] R. J. L. Sastre, P. G. Jimenez, F. J. Acevedo, and S. M. Bascon. Computer algebra algorithms applied to computer vision in a parking management system. In *2007 IEEE International Symposium on Industrial Electronics*. IEEE, jun 2007.

[47] D. Shoup. The access almanac: Ending the abuse of disabled parking placards. *ACCESS Magazine*, 1(39), 2011.

[48] D. Shoup. The right price for curb parking. *RETOOLING METROPOLIS*, page 9, 2016.

[49] D. C. Shoup. Cashing in on curb parking. *Journal of the American Planning Association*, 1994.

[50] D. C. Shoup. The high cost of free parking. *Journal of planning education and research*, 17(1):3–20, 1997.

[51] D. C. Shoup. Cruising for parking. *Transport Policy*, 13(6):479–486, 2006.

[52] E. Sifuentes, O. Casas, and R. Pallas-Areny. Wireless magnetic sensor node for vehicle detection with optical wake-up. *IEEE Sensors journal*, 11(8):1669–1676, 2011.

[53] H. S. Simpson. Downtown storage garages. *The Annals of the American Academy of Political and Social Science*, 133(1):82–89, 1927.

[54] K. Spitaels, S. Maerivoet, G. De Ceuster, G. Nijs, V. Clette, P. Lannoy, K. Dieussaert, K. Aerts, and T. Steenberghen. Optimising price and location of parking in cities under a sustainability constraint (sustapark), final report. *Brussels: Belgian Science Policy*, 104, 2009.

[55] Stanislav Vítek and Petr Melničuk. A distributed wireless camera system for the management of parking spaces. *Sensors*, 18(2):69, Dec 2017.

[56] United Nations. World Urbanization Prospects: The 2018 Revision. https://population.un.org/wup/Publications/Files/WUP2018-KeyFacts.pdf, 2018. Accessed: 2019-06-18.

[57] M. M. B. Vianna, L. da Silva Portugal, and R. Balassiano. Intelligent transportation systems and parking management: implementation potential in a brazilian city. *Cities*, 21(2):137–148, 2004.

[58] W. Vickrey. The economizing of curb parking space. *Traffic engineering*, 25(2):62–67, 1954.

[59] E. I. Vlahogianni, K. Kepaptsoglou, V. Tsetsos, and M. G. Karlaftis. A real-time parking prediction system for smart cities. *Journal of Intelligent Transportation Systems*, 20(2):192–204, 2016.

[60] R. Waraich and K. Axhausen. An agent-based parking choice model, paper presented at the 91st annual meeting of the transportation research board, 2012.

[61] R. A. Waraich, C. Dobler, and K. W. Axhausen. Modelling parking search behaviour with an agent-based approach. In *International Conference on Travel Behaviour Research (IABTR)*, 2012.

[62] A. Wegener, M. Piórkowski, M. Raya, H. Hellbrück, S. Fischer, and J.-P. Hubaux. Traci: an interface for coupling road traffic and network simulators. In *Proceedings of the 11th communications and networking simulation symposium*, pages 155–163. ACM, 2008.

[63] J. Wolff, T. Heuer, H. Gao, M. Weinmann, S. Voit, and U. Hartmann. Parking monitor system based on magnetic field senso. In *2006 IEEE Intelligent Transportation Systems Conference*, pages 1275–1279. IEEE, 2006.

[64] Q. Wu, C. Huang, S. yu Wang, W. chen Chiu, and T. Chen. Robust parking space detection considering inter-space correlation. In *Multimedia and Expo, 2007 IEEE International Conference on*. IEEE, jul 2007.

[65] J. Xiao, Y. Lou, and J. Frisby. How likely am i to find parking?–a practical model-based framework for predicting parking availability. *Transportation Research Part B: Methodological*, 112:19–39, 2018.

[66] S. Yang, W. Ma, X. Pi, and S. Qian. A deep learning approach to real-time parking occupancy prediction in spatio-termporal networks incorporating multiple spatio-temporal data sources. *arXiv preprint arXiv:1901.06758*, 2019.

[67] S. Yang and Z. S. Qian. Turning meter transactions data into occupancy and payment behavioral information for on-street parking. *Transportation Research Part C: Emerging Technologies*, 78:165–182, 2017.

[68] S.-e. Yoo, P. K. Chong, T. Kim, J. Kang, D. Kim, C. Shin, K. Sung, and B. Jang. Pgs: Parking guidance system based on wireless sensor network. In *2008 3rd International Symposium on Wireless Pervasive Computing*, pages 218–222. IEEE, 2008.

[69] Y. Zheng, S. Rajasegarar, and C. Leckie. Parking availability prediction for sensor-enabled car parks in smart cities. In *2015 IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, pages 1–6. IEEE, 2015.

[70] O. Zoeter, C. Dance, S. Clinchant, and J.-M. Andreoli. New algorithms for parking demand management and a city-scale deployment. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1819–1828. ACM, 2014.