

Intelligent Parking Assist for Trucks with Prediction

December
2018

A Research Report from the National Center
for Sustainable Transportation

Petros Ioannou, University of Southern California

Filipe de Almeida Araujo Vital, University of Southern California



About the National Center for Sustainable Transportation

The National Center for Sustainable Transportation is a consortium of leading universities committed to advancing an environmentally sustainable transportation system through cutting-edge research, direct policy engagement, and education of our future leaders. Consortium members include: University of California, Davis; University of California, Riverside; University of Southern California; California State University, Long Beach; Georgia Institute of Technology; and University of Vermont. More information can be found at: ncst.ucdavis.edu.

U.S. Department of Transportation (USDOT) Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the United States Department of Transportation's University Transportation Centers program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Acknowledgments

This study was funded by a grant from the National Center for Sustainable Transportation (NCST), supported by USDOT through the University Transportation Centers program. The authors would like to thank the NCST and USDOT for their support of university-based research in transportation, and especially for the funding provided in support of this project. The authors gratefully acknowledge the company Pilot Flying J for the data available online that was used in this project.



Intelligent Parking Assist for Trucks with Prediction

A National Center for Sustainable Transportation Research Report

December 2018

Petros Ioannou, Department of Electrical Engineering, University of Southern California

Filipe de Almeida Araujo Vital, Department of Electrical Engineering, University of Southern California

[page left intentionally blank]

TABLE OF CONTENTS

EXECUTIVE SUMMARY	iv
Introduction	1
Background	1
Motivation	3
Structure of Report	5
Literature Review	6
Modeling and Prediction	6
Resource Allocation	8
Infrastructure	8
Problem Description	12
Objective	12
Approach	12
Evaluation	13
Data Description	14
Models Description	18
Nonhomogeneous Poisson (NHP)	18
Multivariate Spatiotemporal Model (MSM)	18
Curves Similarity (CS)	21
Forecasting (F / wF)	22
Decision Model	22
Evaluation Results	29
Experiments	29
Results	31
Conclusion and Recommendations	49
Implementation	50
References	51

List of Figures

Figure 1. Truck parking spaces and truck traffic in the state of California. Plot taken from [3]	2
Figure 2. Commercial Vehicle Truck Parking Spaces per Daily 100,000 Miles of Combination Truck Vehicle Miles of Travel (VMT). Plot taken from [3].	3
Figure 3. General structure of the system	12
Figure 4. All USA Truck Stops, California stops in red.....	15
Figure 5. Pilot Flying J Truck Stops, California stops in blue [68].....	15
Figure 6. Relative availability of each truck stop (1 to 14) in CA from October 5, 2017 to October 14, 2017	16
Figure 7. Average Relative availability during the day of each CA truck stop (1 to 14) over the whole collected data.....	17
Figure 8. ROC of experiment 1.1.....	24
Figure 9. ROC of experiment 1.2.....	24
Figure 10. ROC of experiment 1.3.....	25
Figure 11. ROC of experiment 2.1.....	25
Figure 12. ROC of experiment 2.2.....	26
Figure 13. ROC of experiment 2.3.....	26
Figure 14. ROC of experiment 3.1.....	27
Figure 15. ROC of experiment 3.2.....	27
Figure 16. ROC of experiment 3.3.....	28
Figure 17. MSE of the availability estimate of experiment 1.1	32
Figure 18. MSE of the availability estimate of experiment 1.2	32
Figure 19. MSE of the availability estimate of experiment 1.3	33
Figure 20. Classification performance of experiment 1.1	34
Figure 21. Classification performance of experiment 1.2	35
Figure 22. Classification performance of experiment 1.3	36
Figure 23. MSE of the availability estimate of experiment 2.1	37
Figure 24. MSE of the availability estimate of experiment 2.2	38
Figure 25. MSE of the availability estimate of experiment 2.3	38
Figure 26. Classification performance of experiment 2.1	40
Figure 27. Classification performance of experiment 2.2	41
Figure 28. Classification performance of experiment 2.3	42

Figure 29. MSE of the availability estimate of experiment 3.1	43
Figure 30. MSE of the availability estimate of experiment 3.2	44
Figure 31. MSE of the availability estimate of experiment 3.3	44
Figure 32. Classification performance of experiment 3.1	46
Figure 33. Classification performance of experiment 3.2	47
Figure 34. Classification performance of experiment 3.3	48

Intelligent Parking Assist for Trucks with Prediction

EXECUTIVE SUMMARY

Truck parking has been identified as a major issue both in the USA and E.U. and has been selected by the American Transportation Research Institute (ATRI) as the most important research need for the trucking industry in 2015 [1]–[5]. The lack of appropriate and convenient parking locations has been the cause of several safety issues over the past years as drivers might be forced to either drive while tired and increase the risk of accidents or park illegally in unsafe locations, which might also pose a safety hazard to them and other drivers. Additionally, the parking shortage also impacts the shipment costs and the environment as the drivers might spend more fuel looking for parking or idling for power when parked in inappropriate locations.

The project's objective is to study the truck parking problem, generate useful information and parking assist algorithms that could assist truck drivers in better planning their trips. By providing information about parking availability to truck drivers, we expect to induce them to better distribute themselves among existing rest areas. This would decrease the peak demand in the most popular truck stops and attenuate the problems caused by the parking shortage.

In this project, several parking availability prediction algorithms are tested using data from a company's private truck stops reservation system. The prediction MSE (mean squared error) and classification (full/available) sensitivity and specificity plots are evaluated for different experiments. It is shown that none of the tested algorithms is absolutely better than the others and has superior performance in all situations. The results presented show that a more efficient way would be to combine them and use the most appropriate one according to the situation. A model assignment according to current time of the day and target time for prediction is proposed based on the experiment data.

Introduction

Background

It was estimated that, in the year of 2013, trucks were responsible for carrying around 70% (in weight) of USA's total freight shipments, without considering multimodal shipments that use trucks at some point [6]. It is expected that this value will still be as high as 66% by the year of 2040, despite substantial increases in multimodal and rail shipments[6]. This shows just how important trucks are to the USA economy. However, the increasing demand for trucks comes with a need for supporting infrastructure and legislation.

A survey by the American Transport Research Institute(ATRI) has pointed the top issues in trucking industry, among which are the Hours-of-Service (HOS) rules, Compliance, Safety and Accountability (CSA) scores and Truck Parking [1]. These issues are strongly linked. The HOS rules caused an increase in the demand for parking as the drivers cannot exceed a certain number of hours driving. That increase in demand made the already existent problems of truck parking availability even more pronounced, making some drivers opt to park illegally, leading to a decrease in safety conditions.

Over the past years some states have done analysis of their current situation regarding truck parking and the impacts of a shortage in parking locations. The state of California is one of the states with the largest quantity of parking spaces. However, due to also having a large highway network and heavy truck traffic, these parking spaces are too sparse compared to the real necessities of the state. Figure 1 shows information on the parking spaces in California's highways and its traffic. This mismatch between number of parking spaces and road extension/traffic can be better visualized on Figure 2 that shows the ratio of Commercial Vehicle Truck Parking Spaces per Daily 100,000 Miles of Combination Truck Vehicle Miles of Travel (VMT) for several states and shows California among the last ones in this aspect. As of 2000 California had estimated their total number of parking spaces, space shortage and increase in the demand by 2020 as 7500, 8000 and 53% respectively for public rest areas and 1100, 6100 and 100% respectively for private truck stops respect [7].

A 2015 report by the Virginia Department of Transportation (VDOT) calculated a statewide deficit of nearly 5000 parking spaces, which means that they only satisfy around 60% of their calculated demand (12500 spaces)[8]. According to the U.S. DoT (U.S. Department of Transportation), 36 states are experiencing shortages in rest areas, either public or private, which negatively affect truck parking [3]. The lack of parking availability occurs mainly at peak hours of demand which suggests that a better balancing of parking availability in space and time may help reduce the demand at peak hours.



Figure 1. Truck parking spaces and truck traffic in the state of California. Plot taken from [3]

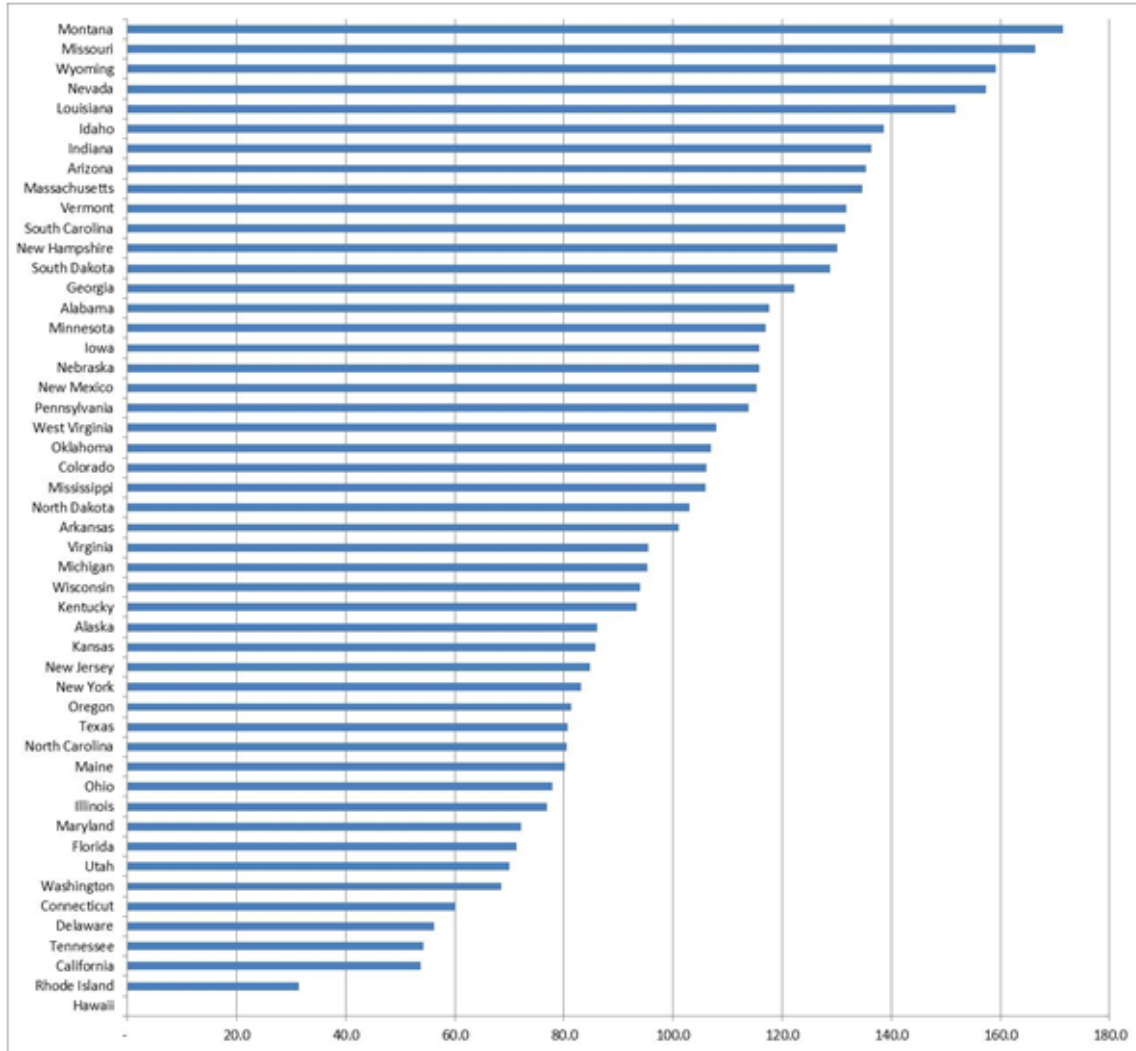


Figure 2. Commercial Vehicle Truck Parking Spaces per Daily 100,000 Miles of Combination Truck Vehicle Miles of Travel (VMT). Plot taken from [3].

Motivation

This project focuses on the truck parking situation, which has been identified as a major issue both in the USA and E.U. and has been selected by the ATRI as the most important research need for the trucking industry in 2015 [1]–[5]. The lack of appropriate and convenient parking locations has been the cause of several safety issues over the past years as drivers might be forced to either drive while tired and increase the risk of accidents or to park illegally in unsafe locations, which might also pose a safety hazard to them and other drivers. Additionally, the parking shortage will also impact the shipment costs and the environment. The following paragraphs comment on these 4 unwanted consequences of the current inadequate parking infrastructure.

Illegal Parking

Surveys carried out by some states have identified several hundred illegal or unofficial parking locations such as freeway shoulders, freeway entrance and exit ramps, roadways accessing freeway ramps, local streets and commercial areas [3]. These practices pose serious safety hazards to other motorists and truck drivers themselves and expose drivers to become targets of ill-intentioned people. In addition it was found that, between 2008 and 2012, 25% of the 4117 truck related accidents in the major corridors in the state of Virginia happened in entrance or exit ramps [8].

Unsafe Driving

With the driving time limits imposed by the HOS rules, a driver unable to find a suitable parking location may choose to either park illegally or drive illegally and tired. A study by the AAA Foundation for Traffic Safety found that 21% of all accidents in which a person was killed involved a drowsy driver [9]. Although the data used was not specific to trucks it shows how dangerous drowsy driving can be. A study by the California DoT shows a reduction in the percentage of fatigue related collisions up to 30 miles downstream of rest areas compared to areas further away [10].

Environmental Impact

The shortage of parking spaces forces drivers to drive around looking for parking and/or park at inappropriate locations. Both actions result in an increase in the environmental impact by generating more pollution and increasing the cost of operations of trucks due to wasted fuel consumption as well as time spent. While in some truck stops the drivers are able to plug in their vehicles to the grid and avoid idling, no illegal parking location will have this kind of service available, forcing the truck to idle for several hours. Idling is a large source of emissions, fuel expenditure and engine wear and many states already have laws and incentives for idle reduction [11]. If the drivers often need to find parking in the local streets, they might even end up impacting the air quality and health of the nearby communities [12].

Cost increase

As mentioned before, the shortage of parking can have a substantial impact on fuel consumption, be it because of the time spent looking for parking or the time the engine will have to spend idling for lack of proper infrastructure. A study by the University of California, Davis has estimated idling time to be responsible for 8.7% of the total fuel consumption of long-haul trucks [13]. Fuel is responsible for a large share of the operational costs in the trucking industry, making the overall cost highly dependent on fuel costs. During the period of 2009 to 2015 the fuel's share in the operational cost has varied a lot due to fluctuations in fuel price, oscillating within the range of 25% to 39% of the total operational cost [14]. Other than the fuel consumption there is still the cost related to vehicle maintenance (10% of total cost in 2015) [14], which can be increased by almost \$2,000 a year due to idling [15]. Insurance premiums are another possibility of impacted costs as they can be affected by the number of accidents

and robberies involving this kind of vehicle. In 2015 insurance premiums counted for 6% of total operational cost [14].

Structure of Report

The following sections are organized as follows: Section Literature Review presents a literature review of work relevant to the problem. Section Problem Description describes the problem to be studied along with the objectives pursued and measures used to evaluate the performance. Section Data Description describes the data used in the experiments. Section Models Description presents the different parking prediction models used in the project. Section Evaluation Results describes the experiments realized and presents the results for each of the tested models. Section Conclusions and Recommendations concludes with a comparison of the different results along with a recommendation of how to approach the problem.

Literature Review

Prior research on intelligent parking for commercial vehicles is scarce. However, there is plenty of research on intelligent parking for passenger vehicles in urban environments. Although the problems are not equivalent, they hold a certain degree of similarity. Therefore, methods for urban parking were also considered when reviewing past work on the topic. The research can be roughly divided into 3 subjects: modeling & prediction of demand/availability; resource allocation; and infrastructure.

Modeling and Prediction

A demand model for commercial vehicle parking on highway segments was developed in a study sponsored by the USA Federal Highway Administration [16]. A simplified version of the model predicts the demand along a highway segment based on the proportion of commercial trucks in the total traffic, the annual average daily traffic, the length of the segment, the average truck speed and the average parking time per truck-hour of travel. The final model still considers factors to account for peak traffic, the ratio between long-haul and short-haul trucks, loading/unloading time and other variables.

In [17], Heinitz chose a region of a highway and divided its upstream region in isochronal 60-min rings. He estimated the parking lot occupancy by measuring the historic traffic flow on these isochronals and used it to estimate how many trucks will pass by that section of the highway and their elapsed driving time. He then used a choice model to estimate where each truck is going to park. He did not limit occupancy to the official capacity, so this work considers the possibility of over-capacity in the lots and illegal parking in the vicinities of each lot when they are full.

In [18], [19] Bayraktar installed a pilot smart parking-management system in a Florida parking lot and used the acquired data to feed a Kalman-filter based occupancy prediction model. The root mean squared error analysis performed by the author showed that the presented Kalman filter prediction performed better than a linear regression model.

Due to the larger number of models available for urban parking prediction there is also a large array of opportunities to look for ways to adapt some of these methods or at least the general idea behind them to truck parking.

In [20], Rajabioun and Ioannou proposed an autoregressive model that takes into account temporal and spatial correlations on parking availability. Instead of using an autoregressive model of the occupancy, an autoregressive model of the difference between the average occupancy variation and the actual occupancy variation was used. This predicted difference is added to the average variation of the following time unit and then summed with the current occupancy in order to estimate the future occupancy.

In [21] the author used genetically optimized neural networks for short term parking occupancy prediction. Genetic algorithms are used to optimize the learning rate and momentum of back-

propagation training algorithm, as well as the structure of the hidden layer and look-back time window. This approach was based on a previous work on short-term traffic-flow prediction and the detailed description is given in [22]. The measured errors for predictions of up to 30 minutes in advance showed that this method has an acceptable level of accuracy.

In [23] a combination of neural networks and classification methods are used to predict if the parking lots will be full or not, according to time, events happening in the vicinity and weather data. The first part of the model tested both generalized regression neural network (GRNN) and multilayer feed-forward network (MLFN) to predict the parking availability, achieving better results with the GRNN. The second part of the model is the classification method that is used to reduce the number of false negatives (not full) generated when the parking lot has high occupancy. The methods tested were a naive Bayes classifier and a classification and regression tree (CART). Better results were achieved using the CART method, leaving the final suggestion as being to use a GRNN for prediction coupled with a CART for classification.

In [24] a method to predict the vehicles' arrival and departure times and therefore the future parking availability in real-time is developed in the context of parking reservations. This system used a discrete choice model with on-line and historical information to simulate the arrivals, choose their allocation and their departure times. It needs to be previously calibrated with information on parking preferences, duration of stay, arrival and departure process and the static capacity of the parking lots.

Other works use Poisson random processes to model parking availability. [25], [26] model the arrival rates of the parking lots as Poisson random processes and the duration of stay as an exponential variable and model the parking lot using a continuous time Markov Chain. Reference [25] represented the parking lot as a homogeneous Markov model. The inter-arrival and parking times are exponentially distributed. There is a state for each of the possible number of currently parked vehicles, for a total of $m+1$ states, where m is the parking lot capacity. The information about certain parameters was obtained from the municipality where the study was realized and by the vehicles through a vehicular ad hoc network. This study used a simulation environment built on the traffic simulator VISSIM coupled with a network simulator and MATLAB to evaluate the performance of the prediction model. [26] builds up on [25] and presents an efficient method to calculate the probability distribution of the occupancy. [27] models the lot's availability as a Poisson random process, estimating the λ parameter using historical data.

[28] used simulation methods in its predictions. An agent-based model (ABC) was combined with Markov Chain Monte Carlo (MCMC) in order to achieve more accurate results than either method applied separately. The simulated data generated through ABC was used to create the proposal distribution for the MCMC. The method was tested using data from the campus of the University of Central Florida (UCF).

Resource Allocation

Besides using prediction algorithms to help plan the best stops beforehand without interacting with other drivers or establishments, another method is reserving spots or negotiating with the nearby vehicles in order to organize the stops in a way that benefits everyone and uses the total parking capacity to its fullest.

In [29], [30], the authors used a multiagent system for the management of parking reservations among requesting trucks. When the vehicle enters the road network it sends its origin, destination and preferred parking to the system manager. If the rest area has available spots a temporary reservation will be made. If the rest area does not have available spots the negotiation protocol is initiated. Each driver receives a list of possible rest areas to be graded according to his/her preferences. Each driver's vote is weighted according to how close they are to reaching their legal driving limit. The scores for each driver are summed for each feasible solution and the solution with the largest score is selected.

This resource allocation problem was also treated before in the context of urban parking. In [31] the resource allocation problem is defined as a sequence of Mixed Integer Linear Programming problems solved over time subject to a set of fairness constraints. [32] uses interval scheduling algorithms to try to optimally allocate the parking spaces.

Infrastructure

The truck parking problem has been the subject of a few initiatives, mainly European, to improve the current situation of commercial vehicles parking infrastructure. These projects usually aim to either provide better information to truck drivers as to the current occupancy in the parking lots or to increase the quantity and quality of parking locations. These were the objectives of the SETPOS project [33] co-funded by the European Commission. This project defined a set of standards for secure parking, built some pilot secure parking areas and also established a platform for information, guidance and reservation. EasyWay Programme [34], [35] is another European project, which aimed to deploy Europe-wide ITS (Intelligent transportation systems) core services. These programs generated a few pilot sites that aimed to test the performance of different types of ITP (Intelligent Truck Parking). Among the countries that tested built pilot sites (not necessarily through these programs) are Austria, France, Germany, Hungary, Denmark and the USA [19], [34]–[43]. Most of these pilot projects used sensors on the entrance/exit to count the numbers of vehicles that enter and leave the lot or installed sensors in each individual parking space to detect its occupancy and use it both for availability information and vehicle guidance inside the lot. However, these are not the only ways to measure occupancy. Examples of sensors used to acquire occupancy data are:

- Magnetic Sensors
- Laser Sensors
- Video image processing
- GPS data analysis

- Crowdsensing
- Vehicle Ad-hoc Networks (VANET)

Magnetic: This kind of sensor has been used in several pilot projects for smart truck parking, both for in/out counting and individual spot sensing. It is used as the basis for some urban smart parking systems, such as [44]–[47]. According to these studies the sensors are very energy efficient and can last for several years before running out of battery. Some sensors use more than one type of sensing at the same time to improve detection. The ones used in [19], [38] use both magnetic and infrared sensing to detect vehicles. Siemens Wimag sensors [48] combine magnetic field sensors with MicroRadar to improve detection and [45] uses an optically triggered magnetic sensor to lower energy consumption. In [49], Haghani et al. developed a parking sensor with magnetic sensors, using a temperature sensor and software to compensate for the temperature drift.

Laser: The Denmark pilot project [35] uses this kind of sensor in order to verify the remaining spaces in each column of the parking lot. The Siemens Sitraffic Conduct+ tested in Germany at the A9, uses laser scanners at the entrance of the parking lot to identify the type of vehicle that is entering.

Video Image processing: This method of detection consists of using images from video surveillance to determine whether there is a vehicle or not in each parking space, or if a vehicle has entered the lot. While the approach is more computationally expensive than other approaches it is attractive due to the fact that surveillance cameras are already installed on parking lots for other reasons. Another advantage of cameras when compared with individual spot sensing methods is that a small quantity of cameras can cover a wide area.

In the area of urban parking reference [50], proposed a vehicle detection algorithm based on gray-level segmentation, accumulator agents and threshold that uses edge detection as a way to avoid the false detection caused by nearby cars' shadows. [51] used the detection of shapes previously drawn on each parking spot to determine if there is any vehicle parked on them. [52]'s sensor fusion-based method used image processing to determine the parking slots' positions and then used ultrasonic sensors to detect occupancy. In [53] the author used convolutional neural networks to detect the parking spaces occupancy on regular color camera images. [54] proposed the use a team of unmanned aerial vehicles to inspect parking lots cooperatively while using a generative adversarial network model to detect vacant and occupied parking spots on the acquired color images.

In the case of truck parking and trucks, [39], [40] used several images of the same area from different points of view to reconstruct a 3D representation of that scene. The 3D representation is then matched with real-world coordinates and used to check if there is a vehicle occupying the position of each parking space. This method avoids the problems that 2D approaches have with shadows and occlusion.

GPS data analysis: GPS data has been used for freight performance [55] and parking utilization analysis [56], [57] with data obtained by the ATRI.

In the ATRI report [57], the author developed an expansion factor for every hour in the observation period by counting how many trucks passed by the weigh-in-motion traffic counting station and comparing this number with the GPS counting for the same period. This factor was later used as a multiplier to estimate the real occupancy of a nearby parking lot based on the GPS data for the parking lot at each hour.

In [56], the authors compare 4 different models for truck parking utilization: a Poisson, a negative binomial, a Poisson with propensity, and a Poisson with propensity and threshold specific constant. Variables such as time of day, number of vehicles detected on/off-ramp, average speed close to the parking lot and truck volume were used to estimate the parameters for these models. The models were compared using the Bayesian Information Criterion and the Generalized Poisson count model with propensity and threshold was found to have a better data fit.

The GPS data can provide the number of vehicles in the general area of the parking lots and also information about the current truck traffic flow in highways that lead to that parking lot. However, applications that need to use GPS data in real-time might have issues with data acquisition. Some might find that privacy is an issue and, considering that a certain percentage of the trucks in operation would need to participate in order to get reliable data, drivers/companies engagement might be a concern.

In the context of urban parking, [58] utilized GPS, along with the accelerometer and Bluetooth sensors of phones in order to detect when and where the device's owner parked his/her car and when the parking space was released. Same as with the case of the trucks, system penetration would also be an issue. So, the availability estimation algorithm would need to take this and other possible errors (GPS measurements, Bluetooth pairing, transportation mode detection) into account before providing a final estimate.

Crowdsourcing: This method consists of enlisting a large number of people, usually the service users themselves, to obtain the information needed instead of setting up specific infrastructure to acquire the information directly. This approach has been tried for urban parking [59]–[62] where the system is used to collect information about parking availability and distribute it to its users. In [59] users and parking operators can publish information and the system uses driver behavior to improve the parking availability estimation. In [61] ultrasonic rangefinders were installed in cars and used to measure individual on-street parking space occupancy along their commute. The GPS accuracy was an issue, so, by relying on the spatial correlation of the GPS error, environmental fingerprinting was used to correct the error and improve accuracy. [60] estimates the circling time until a vehicle parks. This system gathers user data implicitly and its results are highly dependent on having a large pool of users. This is a big problem during the early stages of implementation, so they also propose a model to estimate the circling time based on the current parking utilization.

In order to address the user participation problem, and also to improve information quality in systems that require users to contribute explicitly, [62] suggests a platform that gives monetary incentives for the participation and bonuses when the information is used successfully by other users.

On the commercial vehicles side, apps like Truckerpath [63] and ParkMyTruck [64] are already operational, providing information to truck drivers. TruckerPath gets information from both drivers and parking providers, while ParkMyTruck only allows registered providers to update their availability.

VANETs: Vehicles can also acquire extra information communicating with each other through vehicle ad hoc networks. If a vehicle gets important information through its own sensors or the infrastructure, this information can be distributed to other vehicles in the network for which this data might be useful. An implementation of this concept was presented in [65] by Caliskan, who also used it in [25] as a means to acquire data for a parking prediction algorithm. As in other forms that rely on independent agent to acquire information, the availability and quality of the data is dependent on the penetration rate of the technology. [66] proposed a VANET-based system to provide convenient services in large parking lots, like parking navigation, occupancy information and anti-theft protection. Similarly, [67] used VANETs as the medium for its parking information dissemination.

Problem Description

There are 2 approaches to deal with truck parking availability issues:

- invest in infrastructure by increasing the number of parking spaces which involves expanding current rest areas and/or building new ones;
- improve usage efficiency of existing parking lots: reorganize the demand over time and space aiming to decrease the peak demand.

This project focuses on the second approach, by studying tools that enable truck drivers to make better decisions about when and where to find parking.

Objective

The main objective is to develop a method that provides reliable information to truck drivers with respect to future parking availability. Such information will assist truck drivers and planners to decide when is more convenient for drivers to take a break by considering the current driving time and expected parking availability and personal preferences.

Approach

The approach taken is the use of a prediction model to estimate the future state of the parking lots of a certain region. This result is then fed into a decision model, which decides on whether to recommend that parking lot at that specific time or not. Basically, it categorizes them as “likely to be full” or “unlikely to be full”. Although the prediction model already generates an estimate of the parking availability it is important to remember that predictions are not perfect, so the decision model considers the prediction model’s error analysis when classifying the parking lots, not only if they predicted zero or not. A diagram with the general structure of this approach is shown in Figure 3. The prediction and decision models used will be described in a different section.

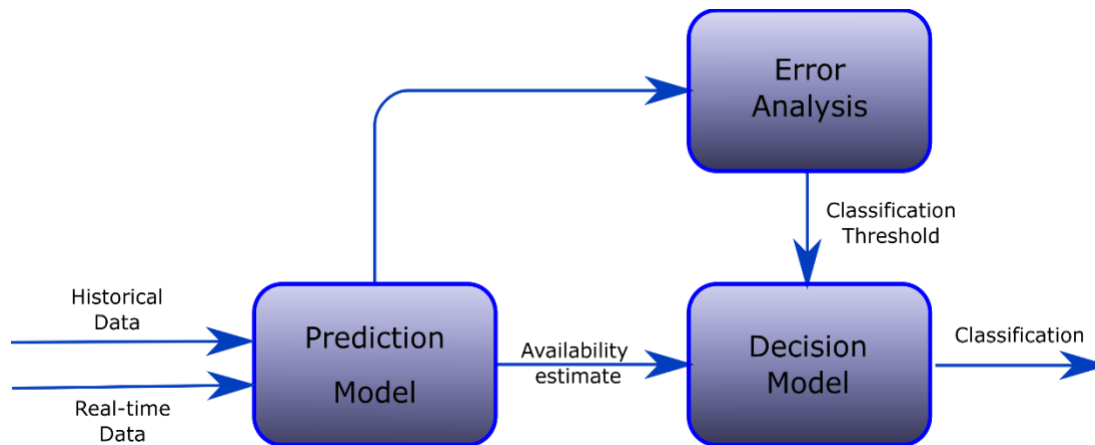


Figure 3. General structure of the system

Evaluation

Depending on the prediction model used, 2 different kinds of information can be provided to the drivers:

- Estimated parking availability (relative or absolute);
- Classification as “Full” or “Available.”

The performance for these 2 items will be treated separately as each model can perform differently in each one and the drivers might prefer to use different ones according to the situation. Also, as different applications might require estimates for different prediction horizons, the impact of the prediction horizon on the performance will also be studied.

The estimated parking availability will be evaluated according to the mean squared error (MSE) calculated using the relative availability values (availability/capacity) of each parking lot, instead of their actual availability. The classification error will be evaluated through an analysis of the sensitivity and specificity values calculated based on the number of true/false positives and true/false negatives.

Data Description

The data consists of available parking spots for the reservation service of the company Pilot Flying J[68]. This service reserves a spot at a given rest area from 4pm on the day of the reservation until 3pm on the next calendar day. Hence, it is reasonable to assume that, after 4pm, the truck stop has at least as many available spots as shown on the data for that day, even though before that time some customers from the previous day might still be present.

It is important to note that the data is for the reservation service and not the real-time occupancy of each truck stop. The reserved spots can be considered occupied even before the driver arrives as the reservation guarantees a spot, but the percentage of spots that are part of the reservation system can be different for each establishment. Also, the data reflects the reservation time, not the actual arrival time of the trucks.

The data, relative to more than 300 private truck stops throughout the USA, was collected from the company website every 15 minutes using a Python script for the period April 28th, 2017 to October 14th, 2017. Figure 4 shows the location of every truck stop in the USA according to data by the Federal Highway Administration (FHWA) and Figure 5 shows the location of every Pilot Flying J truck stop for which data was gathered in this study. Although data from the whole country was gathered, the presented experiments used only data from the state of California (CA).

The information collected was: available parking spaces at each time, price, parking lot address. The website allows reservations for a certain day to be made until 9pm of that same day, when they start accepting reservation only for the following days. We collected data from 12am until 9pm of each day for the available spots for that same day. There are no availability data on any day from 9pm to 11:59pm. Even though there is a possibility of making reservations more than a day in advance that data was not collected. The data was stored in a database built using MySQL 5.7 and processed using Python 3.4.4 with the library numpy 1.11.3.

Figure 6 shows the behavior of the 14 CA truck stops from October 5, 2017 to October 14, 2017. It plots the relative availability of each of these parking lots over the mentioned time period, where relative availability is defined as:

$$relative\ availability = \frac{availability}{capacity}$$

Figure 7 shows the parking lots' average daily behavior over the whole collected data. In this plot, the line is the average relative availability for each time of the day. This average is calculated over all the days in for which measurements were taken.

$$avg\ relative\ availability\ (time) = \sum_{day \in dataset} \frac{relative\ availability\ (time, day)}{total\ number\ of\ days}$$

The red regions of the plots indicate the regions for which there was no data and they can be ignored. It can be seen from the sample data that there is a large variance among different parking lots' availabilities, specially by the end of the day around 9pm when the parking lots reach peak occupancy. While some truck stops tend to have, on average, a very low availability by the end of the day, others seem to have a lot of spare capacity.

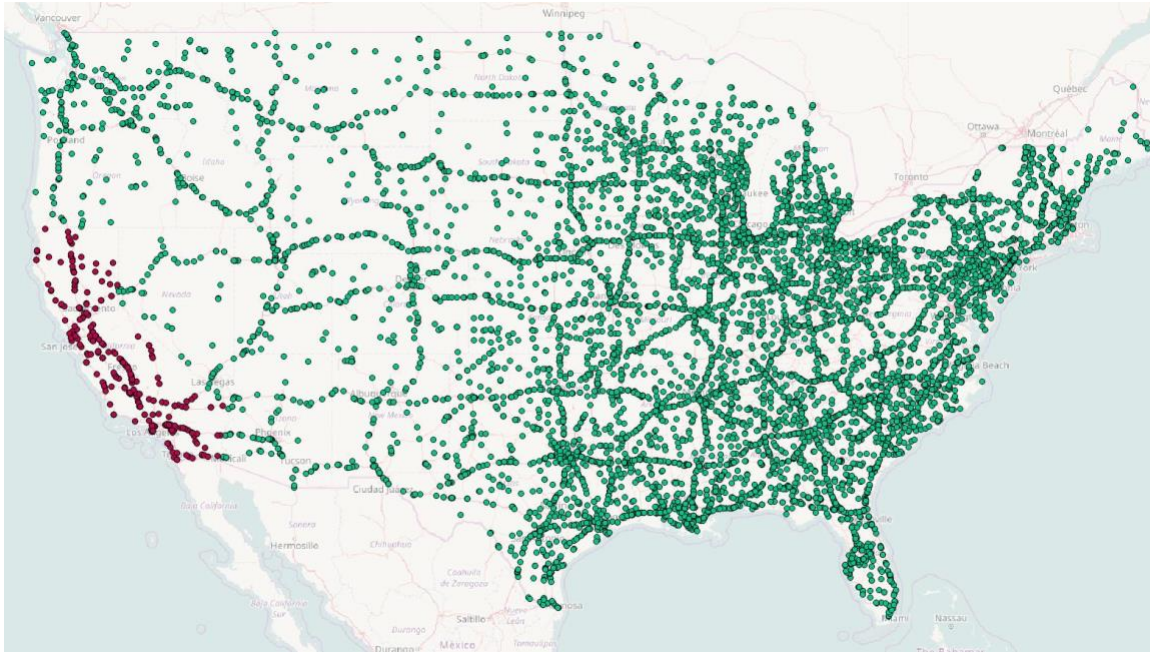


Figure 4. All USA Truck Stops, California stops in red

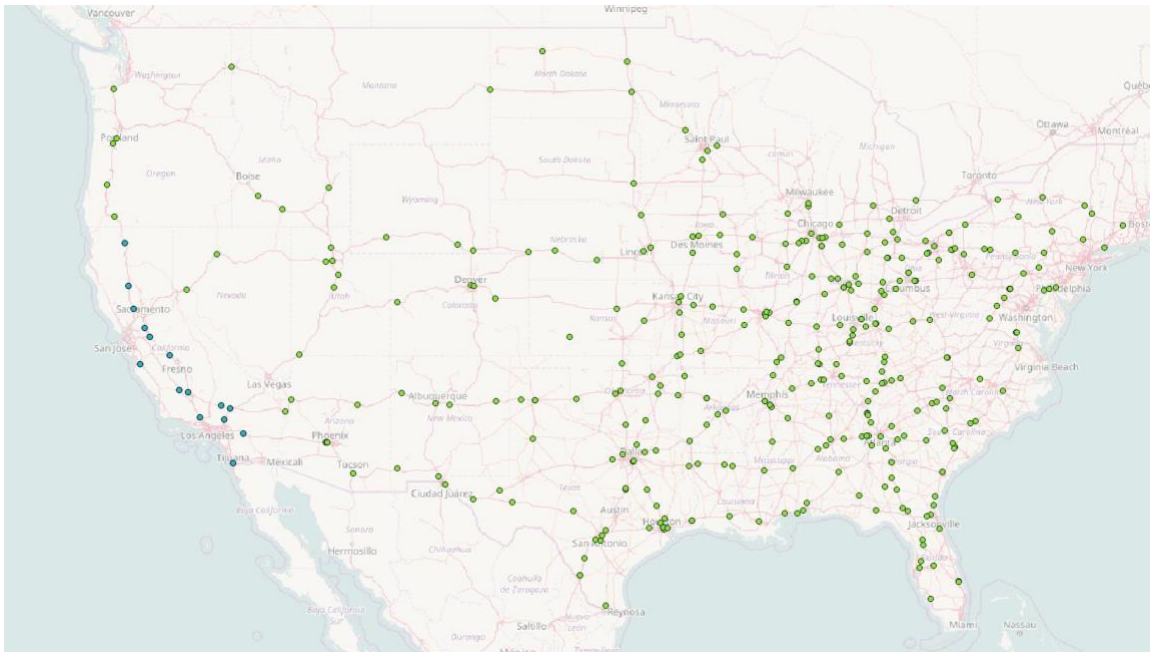


Figure 5. Pilot Flying J Truck Stops, California stops in blue [68].

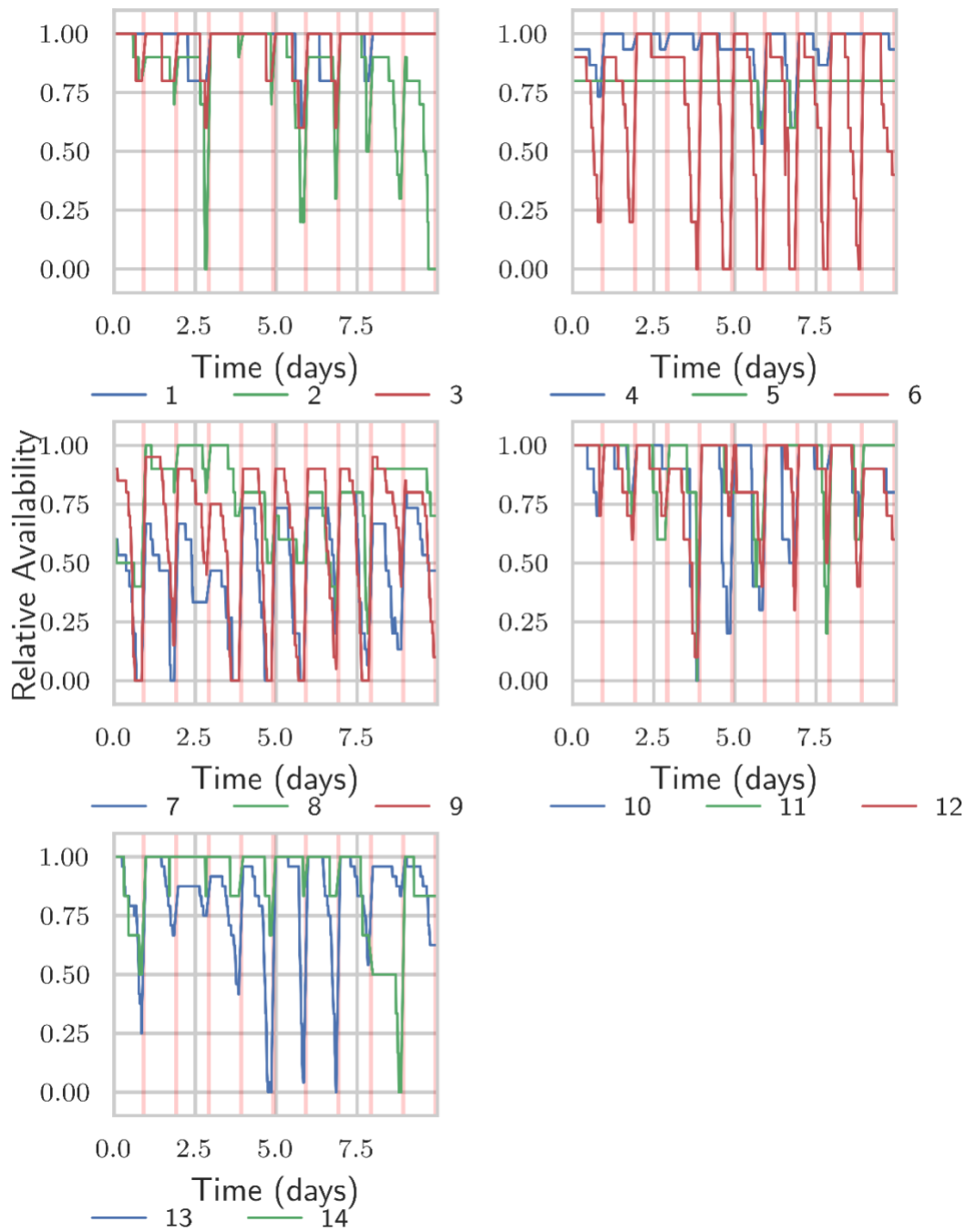


Figure 6. Relative availability of each truck stop (1 to 14) in CA from October 5, 2017 to October 14, 2017

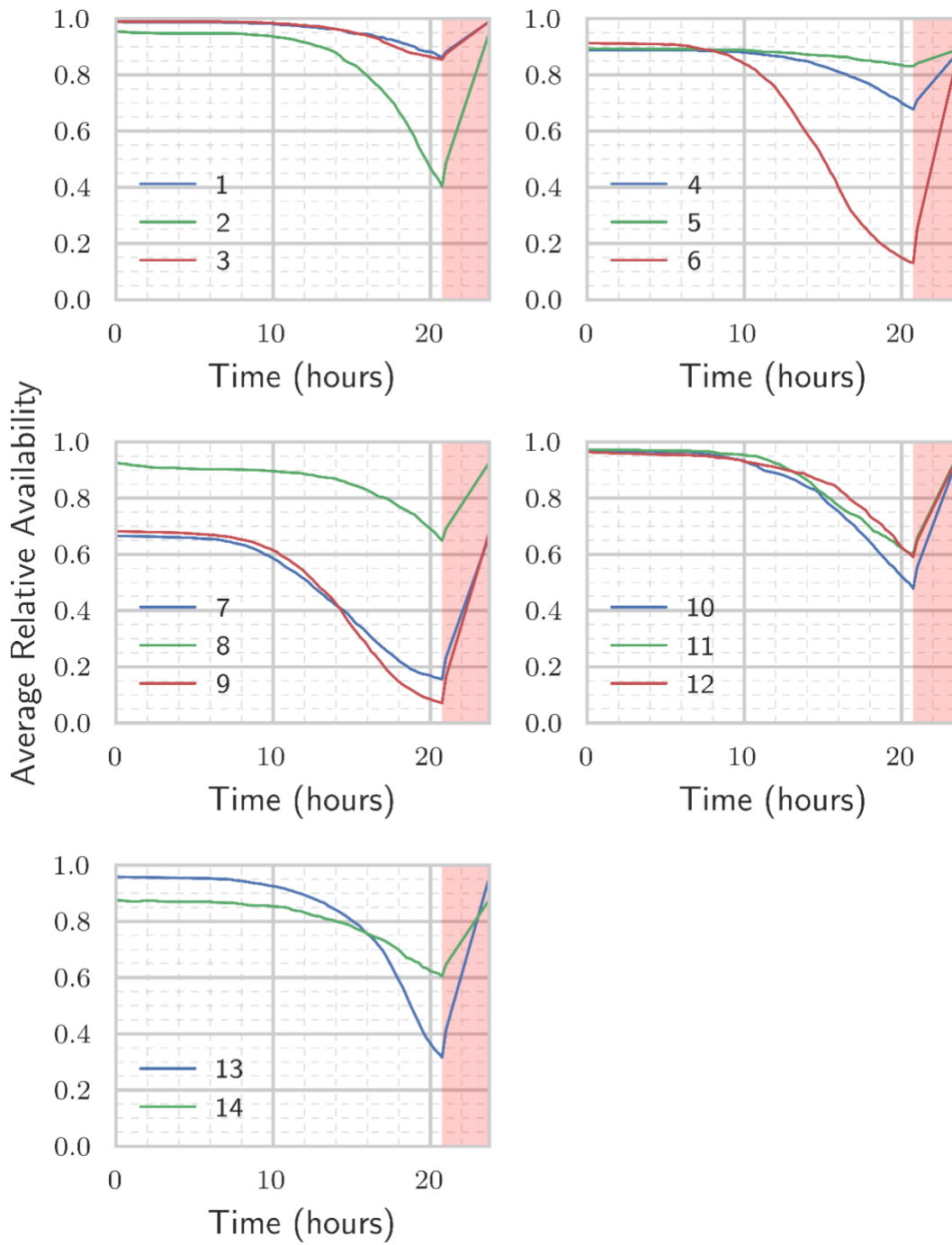


Figure 7. Average Relative availability during the day of each CA truck stop (1 to 14) over the whole collected data.

Models Description

Nonhomogeneous Poisson (NHP)

Similar to [25], [69], this model treats the number of vehicle arrivals/reservations ($v_{l,t}$) at a location l during each time interval $[t - 1, t]$ as a Poisson random variable with a different parameter $\lambda_{l,t} = c_l \bar{v}_{l,t}$ for each time interval. However, in this case we are not dealing with the outgoing flow as the data being used is of a daily reservation system. In this case $v_{l,t}$ is defined as:

$$v_{l,t} = \frac{a_{l,t} - a_{l,t-1}}{c_l} \quad (1)$$

Where c_l is the total capacity of parking lot l and $a_{l,t}$ is the availability of parking lot l at time t . The parameter $\bar{v}_{l,t}$ is estimated through a weighted average, over the last W weeks (current one not included), of windows of width H starting at time t of the same weekday as the one for which the estimate is needed. e.g., Mondays will only use data from the last W Mondays and so on. It is defined as:

$$\bar{v}_{l,t} = \frac{1}{H} \sum_{j=1}^W w_j \sum_{i=0}^{H-1} v_{l,t-i-jT_w} \quad (2)$$

where T_w is the number of intervals in a week and w_i are the weights, summing to 1, for the data from each week. In this project, the weights were all set to $1/W$. Assuming the occupancy at time t is known, the estimated occupancy $\hat{a}_{l,t+k}$ at a certain time $t+k$ and location l can be calculated by adding the expected number of arrivals for each time interval in $[t, t+k]$ to the known occupancy $a_{l,t}$:

$$\hat{a}_{l,t+k} = a_{l,t} + E \left[\text{Poisson} \left(\sum_{i=1}^k \lambda_{l,t+i} \right) \right] = a_{l,t} + c_l \sum_{i=1}^k \bar{v}_{l,t+i}$$

Multivariate Spatiotemporal Model (MSM)

Proposed on [20], this approach detrends the variation in availability at each time interval and then uses an autoregressive model with information from all parking lots in the area to estimate the future detrended variation. The variation and average variation are defined as in the Nonhomogeneous Poisson model presented before. The detrended availability variation $\tilde{v}_{l,t}$ is defined as:

$$\tilde{v}_{l,t} \triangleq v_{l,t} - \bar{v}_{l,t}$$

Where $v_{l,t}$ is defined as in (1) and $\bar{v}_{l,t}$ as in (2).

Autoregressive Model: The autoregressive model is used to estimate each parking lot's $\tilde{v}_{l,t}$ with the past $\tilde{v}_{l,t}$ from L parking lots in the region. With this the spatial correlation of the parking lots availabilities can also be taken into account. Let

$$\tilde{V}_t \triangleq \begin{bmatrix} \tilde{v}_{1,t} \\ \tilde{v}_{2,t} \\ \vdots \\ \tilde{v}_{L,t} \end{bmatrix}, E_t \triangleq \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ \vdots \\ e_{L,t} \end{bmatrix}$$

$$\tilde{V}_t = \sum_{m=1}^M A_m \tilde{V}_{t-m} + E_t \quad (3)$$

Where the coefficients A_m are $L \times L$ matrices that need to be estimated and E_t is a white random vector.

The estimator \hat{V}_t of \tilde{V}_t is defined as:

$$\begin{bmatrix} \hat{v}_{1,t} \\ \hat{v}_{2,t} \\ \vdots \\ \hat{v}_{L,t} \end{bmatrix} = \hat{V}_t = \sum_{m=1}^M \hat{A}_m \tilde{V}_{t-m} \quad (4)$$

Where \hat{A}_m are the estimated values of A_m .

After finding \hat{V}_t the estimated availabilities $\hat{a}_{l,t+1}$ can be found as follows:

$$\hat{a}_{l,t+1} = a_{l,t} + c_l(\tilde{v}_{l,t+1} + \hat{v}_{l,t+1})$$

Training the model: Let $R(A)$ be the mean squared residual prediction error for the last T time steps for a given set of coefficients $A = (\hat{A}_1, \hat{A}_2, \dots, \hat{A}_m)$.

$$R(A) = \sum_{t=k-T}^k (\tilde{V}_t - \hat{V}_t)' (\tilde{V}_t - \hat{V}_t)$$

Let

$$\Psi(k) = [\tilde{V}_{k-T+1} \quad \tilde{V}_{k-T+2} \quad \dots \quad \tilde{V}_k]$$

$$\Theta = [A_1 \quad A_2 \quad \dots \quad A_m]$$

$$\Phi(k-1) = \begin{bmatrix} \tilde{V}_{k-T} & \tilde{V}_{k-T+1} & \dots & \tilde{V}_{k-1} \\ \tilde{V}_{k-T-1} & \tilde{V}_{k-T} & \dots & \tilde{V}_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{V}_{k-T-M+1} & \tilde{V}_{k-T-M+2} & \dots & \tilde{V}_{k-M} \end{bmatrix}$$

$$E(k) = [E_{k-T+1} \quad E_{k-T+2} \quad \dots \quad E_k]$$

Equation (3) can be rewritten in matrix form as:

$$\tilde{V}_t = [A_1 \quad A_2 \quad \cdots \quad A_m] \begin{bmatrix} \tilde{V}_{t-1} \\ \tilde{V}_{t-2} \\ \vdots \\ \tilde{V}_{t-M} \end{bmatrix} + E_t \quad (5)$$

Therefore, we can write that:

$$\Psi(k) = \Theta \Phi(k-1) + E(k)$$

A consistent and unbiased least square estimator of Θ , that minimizes $R(A)$, is found to be:

$$\hat{\Theta}(k) = \Psi(k)\Phi'(k-1)[\Phi(k-1)\Phi'(k-1)]^{-1}$$

This method uses data from $T+M$ time steps to calculate an estimate of the value of Θ for the current time step. This $\hat{\Theta}(k)$ can then be used to calculate the value \hat{V}_t for the next time step by using (4) after writing it in matrix form as (5).

Predicting multiple steps in advance: The presented model estimates \tilde{V}_t for 1 step ahead of the current time. However, it might be necessary to estimate the availability with different prediction horizons. In this case the predictions can be fed to the model as if were real data and used to increase the prediction horizon, with the drawback of an amplification of the prediction error. Rewriting (5) so that the output and input have the same form:

$$\begin{bmatrix} \tilde{V}_t \\ \tilde{V}_{t-1} \\ \vdots \\ \tilde{V}_{t-M+1} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & A_3 & \cdots & A_M \\ I_L & 0 & 0 & \cdots & 0 \\ 0 & I_L & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_L & 0 \end{bmatrix} \begin{bmatrix} \tilde{V}_{t-1} \\ \tilde{V}_{t-2} \\ \vdots \\ \tilde{V}_{t-M} \end{bmatrix} + \begin{bmatrix} E_t \\ E_{t-1} \\ \vdots \\ E_{t-M+1} \end{bmatrix}$$

Or

$$\tilde{V}_t = \mathbb{A}\tilde{V}_{t-1} + \mathbb{E}_t \quad (6)$$

Similarly to (4), the estimator \hat{V}_t is defined as:

$$\hat{V}_t = E[\tilde{V}_t | \mathbb{I}_{t-1}] = \hat{\mathbb{A}}\tilde{V}_{t-1}$$

Where \mathbb{I}_{t-1} is all available information up until time $t-1$.

By using (6) we can iteratively get to the estimator for k -steps ahead predictions, which is defined as:

$$\hat{V}_{t+k} = E[\tilde{V}_{t+k} | \mathbb{I}_t] = \hat{\mathbb{A}}^k \tilde{V}_t$$

After finding all \hat{V}_t 's in the interval $[t + 1, t + k]$ the estimated availabilities $\hat{a}_{l,t+k}$ can be found as follows:

$$\hat{a}_{l,t+k} = a_{l,t} + c_l \sum_{i=1}^k (\bar{v}_{l,t+i} + \hat{v}_{l,t+i})$$

Curves Similarity (CS)

This forecasting model, presented in [70], was developed with hotel reservations in mind. As the gathered data come from a reservation system this method was selected to be tested and compared with the other methods that were developed for real-time occupancy data. The general idea of this approach is comparing the incomplete daily availability curve for each parking lot with the daily availability curves from previous days up to the current time and saying that the current day's reservations will behave the same way as whichever days had the most similar curves.

Let $a_{l,d,t}$ be the availability at location l , day d and time t , where $t \in \mathbb{N}$ and $t = 0$ means the first data point of the day. This model compares the incomplete curve of the forecasted day to the complete curves in historical data. The similarity measure S_i is obtained by comparing several points along the curves:

$$S_i = \sqrt{(a_{l,d,t} - a_{l,i,t})^2 + (a_{l,d,t-k} - a_{l,i,t-k})^2 + (a_{l,d,t-2k} - a_{l,i,t-2k})^2}$$

where d is the forecasted day, i is the day being compared and k is a parameter, chosen by the user, defining how disperse the measurements should be. However, in this project instead of choosing only 3 points, the whole incomplete curve was compared. Therefore, the used expression for S_i was:

$$S_i = \sqrt{\sum_{k=0}^t (a_{l,d,t-k} - a_{l,i,t-k})^2}$$

Let Z be the set of days for which the S_i is smaller than a threshold Ψ .

$$Z = \{i \mid S_i < \Psi\}$$

The estimate $\hat{a}_{l,d,t+n}$ of $a_{l,d,t+n}$ is taken as:

$$\hat{a}_{l,d,t+n} = \frac{1}{|Z|} \sum_{i \in Z} a_{l,i,t+n}$$

Forecasting (F / wF)

This is a simple forecasting method that estimates the availability on the forecasted day at a given time as the average availability at that time from historical data. This method can consider either only the time (F) or the time and the weekday (wF), both options were tested in the project. It would also be reasonable to consider the month, but the gathered data is not enough for this consideration.

Let $a_{l,d,t}$ be the availability at location l , day d and time t , where $t \in \mathbb{N}$ and $t = 0$ means the first data point of the day. The estimate $\hat{a}_{l,d,t}$ when considering only the time of the day is defined as:

$$\hat{a}_{l,d,t+n} = \frac{1}{|S|} \sum_{i \in S} a_{l,i,t}$$

where S is the set of days for which data for time t is available.

The estimate $\hat{a}_{l,d,t}$ when considering the time and the weekday is defined as:

$$\hat{a}_{l,d,t+n} = \frac{1}{|S_d|} \sum_{i \in S_d} a_{l,i,t}$$

Where S_d is the set of days that are the same weekday as d and for which data for time t is available.

Decision Model

The classification was done based on the predicted relative availability for each time. A threshold H was defined for each method and whenever the predicted availability is below that threshold, the parking lot is classified as full. So, the classifier $C(\cdot)$ for a location l on day d at time t is defined as:

$$C(l, d, t) = \begin{cases} 1, & \text{if } \hat{a}_{l,d,t} < H \\ 0, & \text{if } \hat{a}_{l,d,t} \geq H \end{cases}$$

In order to choose the threshold, the experiments were run for a different set of 40 target days. The receiver operating characteristic (ROC) curve was generated for each experiment and each prediction model and the threshold was selected for each curve by maximizing the Youden index, which is defined as:

$$\text{Youden Index} = \text{sensitivity} + \text{specificity} - 1$$

$$\text{sensitivity} = \frac{TP}{TP + FN}$$

$$\text{specificity} = \frac{TN}{TN + FP} \quad (7)$$

where TP is the number of true positives (parking lot is full and was classified as full) in the experiment, FN is the number of false negatives (parking lot is full and was classified as available), TN is the number of true negatives (parking lot is available and is classified as available) and FP is the number of false positives (parking lot is available and is classified as full).

The true positive rate and false positive rate used to plot the ROC curves are defined as:

$$\text{true positive rate} = \frac{TP}{TP + FN}$$

$$\text{false positive rate} = \frac{FP}{TN + FP}$$

The thresholds for each experiment are shown on Table 1 and the ROC curves used to find them can be seen on Figures 8 through 16 with markers at the points generated by the chosen thresholds.

Table 1. Classification thresholds found for each experiment.

Experiment	MSM	NHP	CS	Forecast	W Forecast
1.1	0.12	0.11	0.41	0.56	0.56
1.2	0.3	0.29	0.41	0.56	0.56
1.3	0.41	0.43	0.41	0.56	0.56
2.1	0.31	0.3	0.56	0.505	0.28
2.2	0.4	0.36	0.41	0.21	0.43
2.3	0.505	0.505	0.45	0.505	0.56
3.1	0.1	0.4	0.31	0.505	0.34
3.2	0.3	0.26	0.31	0.505	0.39
3.3	0.37	0.32	0.41	0.56	0.56

ROC of Experiment 1 with a prediction horizon of 15min

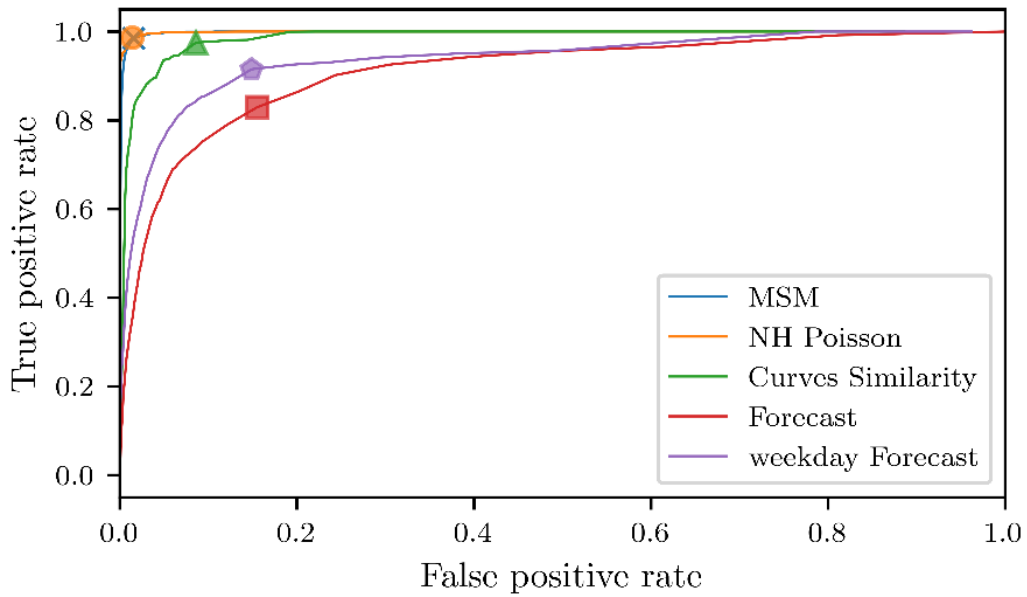


Figure 8. ROC of experiment 1.1

ROC of Experiment 1 with a prediction horizon of 1 hour

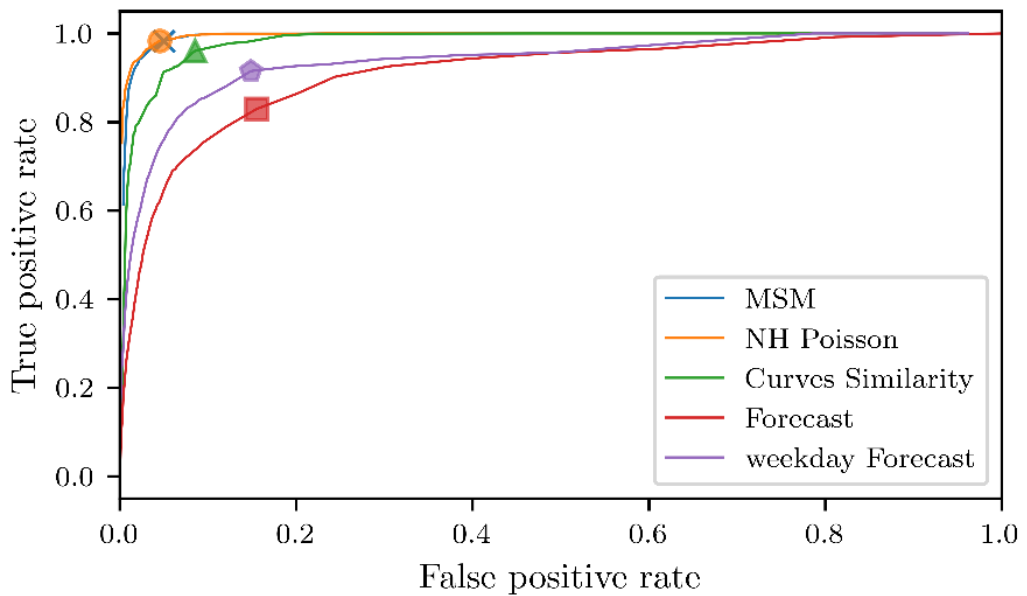


Figure 9. ROC of experiment 1.2

ROC of Experiment 1 with a prediction horizon of 3 hours

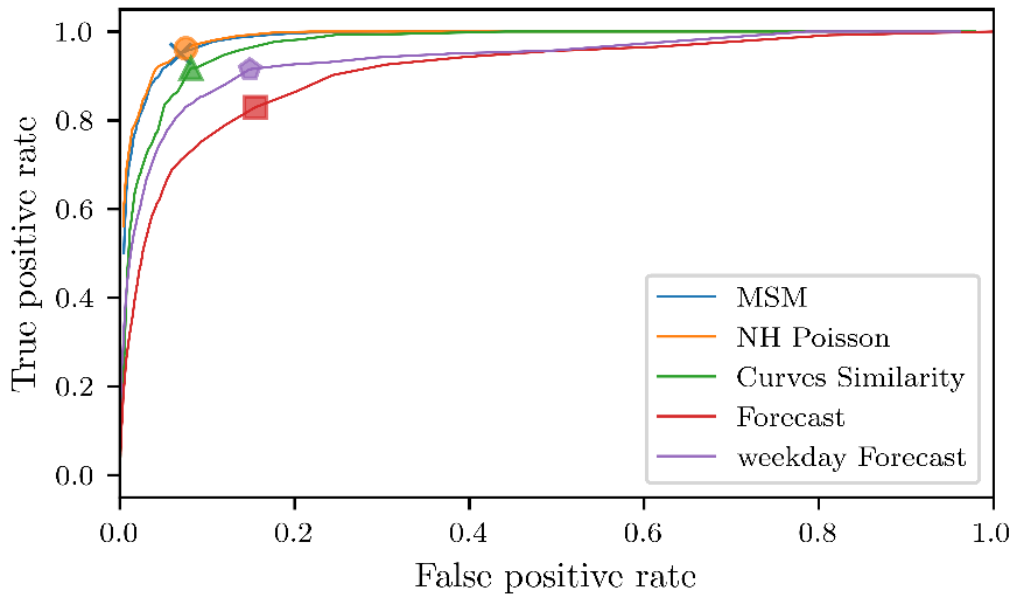


Figure 10. ROC of experiment 1.3

ROC of Experiment 2 with a target time of 20:45

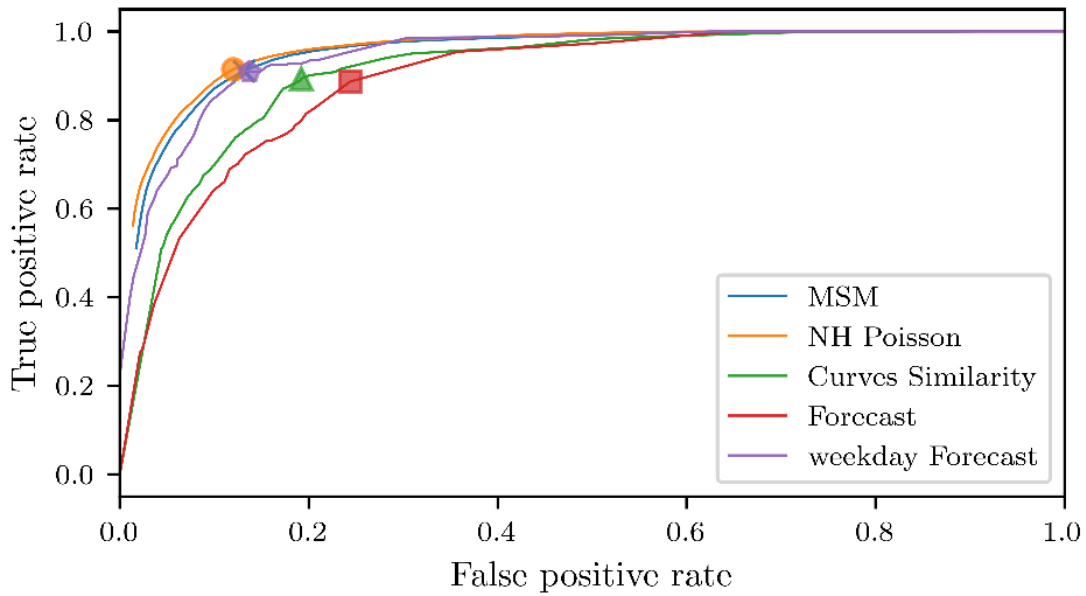


Figure 11. ROC of experiment 2.1

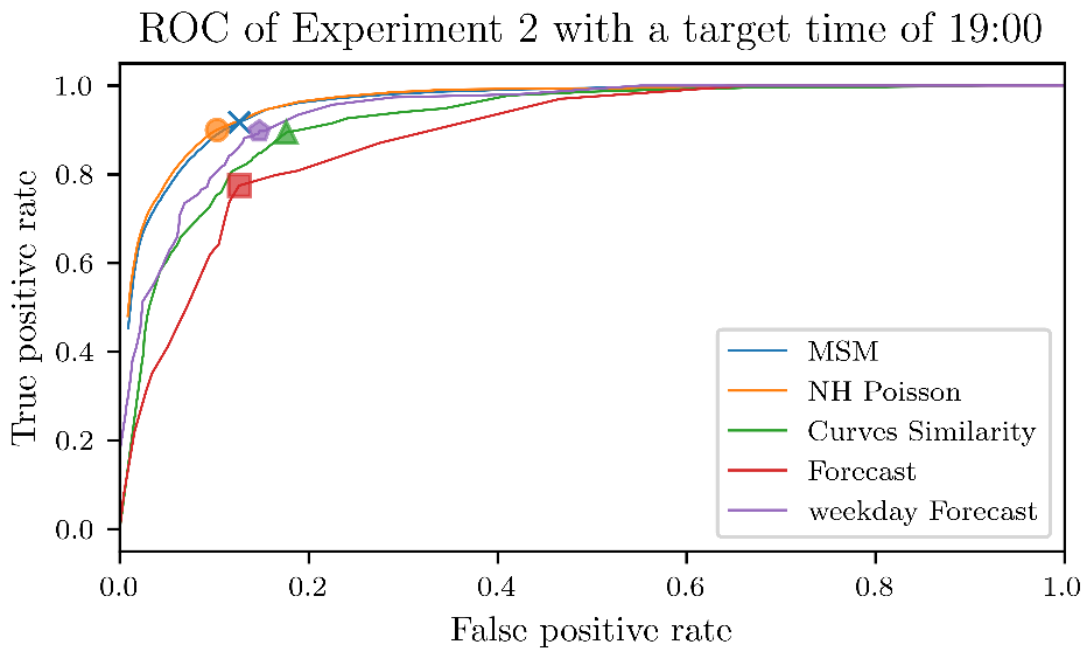


Figure 12. ROC of experiment 2.2

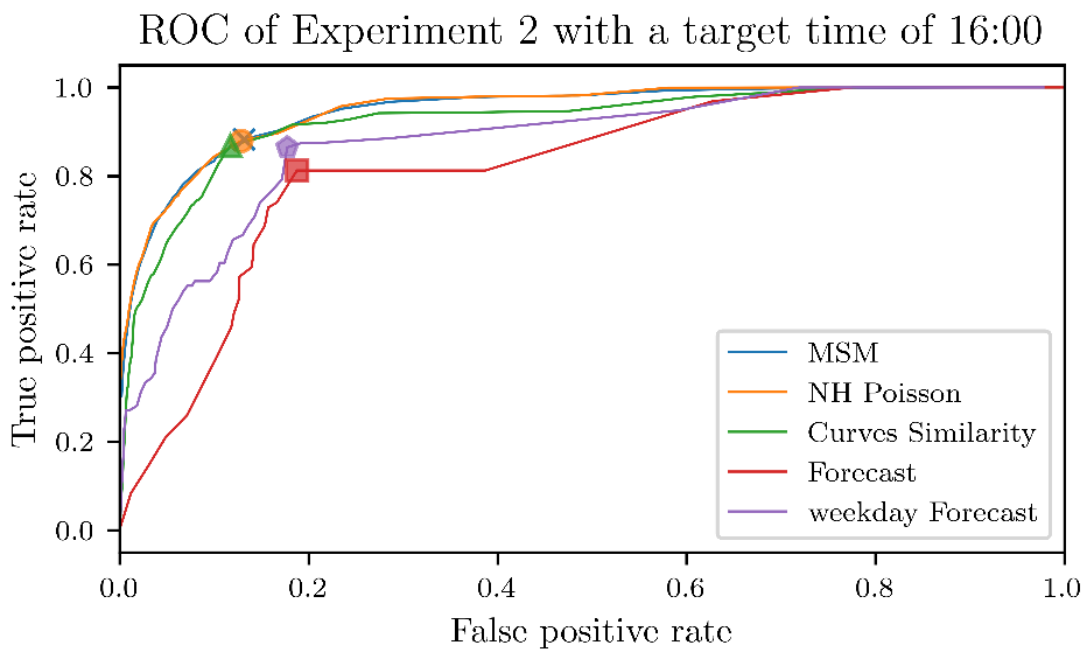


Figure 13. ROC of experiment 2.3

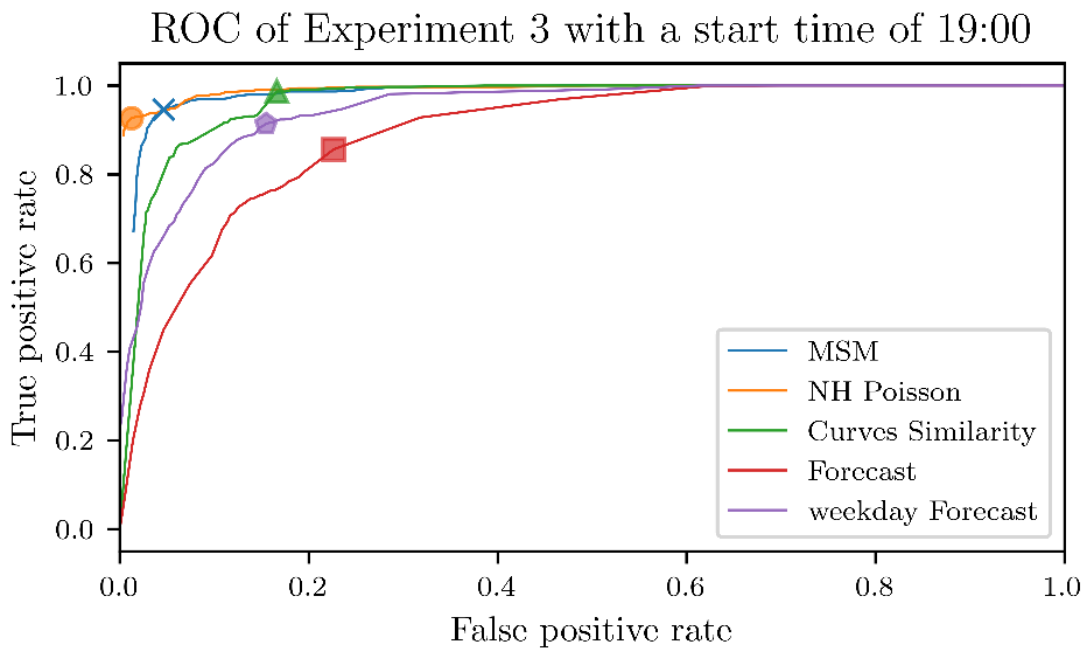


Figure 14. ROC of experiment 3.1

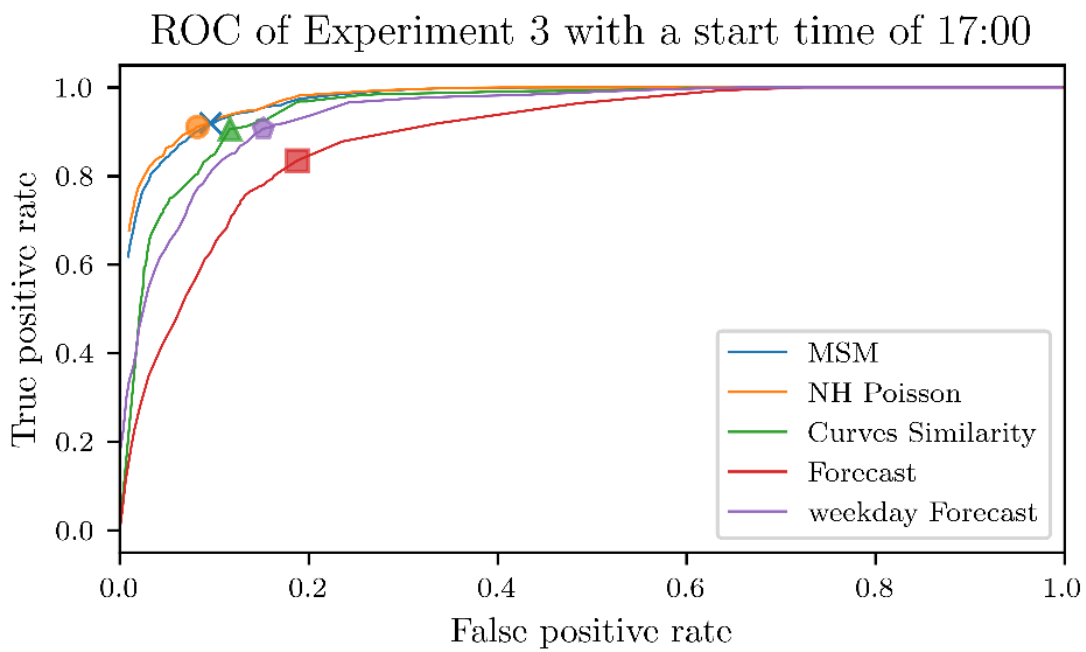


Figure 15. ROC of experiment 3.2

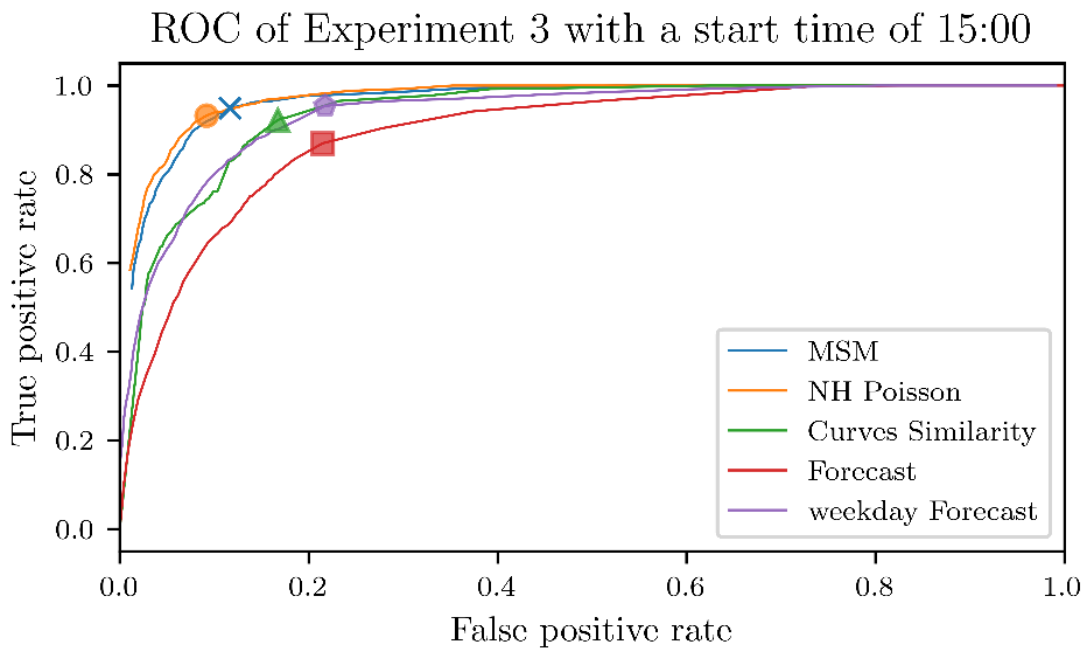


Figure 16. ROC of experiment 3.3

Evaluation Results

Experiments

For the evaluation of the results, 3 types of experiments were set up. They aim to study how different factors affect the performance of each algorithm, so that a decision can be made as to which method is more reliable in different situations.

Due to the nature of the data, aside from internal parameters specific to each algorithm, the prediction system can be affected by 4 factors:

- target time: the time for which a prediction is needed;
- start time: the current time, or the latest time for which data is available;
- prediction horizon: the difference between the target and start times;
- weekday: in this case only predictions within the same day are being considered, so both the start and target time are on the same weekday.

With these factors in mind, 3 experiments were chosen to analyze their influence on the performance of the algorithms.

1. Fixed prediction horizon / varying target time
2. Fixed target time / varying prediction horizon
3. Fixed start time / varying prediction horizon

Experiment 1 - Fixed prediction horizon / varying target time

For this experiment 3 different prediction horizons were chosen: 15 min, 1h and 3h. Each algorithm was run with target times varying from 7:00 to 20:45 and these 3 prediction horizons for n consecutive days from the data sample. In this project's experiments n was set to 40.

For this experiment the MSE is calculated as:

$$e(t) = \frac{1}{D \cdot L} \sum_d \sum_l \left(\frac{a_{l,d,t} - \hat{a}_{l,d,t}}{c_l} \right)^2$$

Where D is the number of days that were evaluated, L is the number of parking lots, c_l is the total capacity of parking lot l , $a_{l,d,t}$ and $\hat{a}_{l,d,t}$ are, respectively, the availability and predicted availability of parking lot l at time t on day d . The variables d and l vary over all evaluated days and parking lots, respectively.

The number of false(true) positives(negatives) is calculated by summing the false(true) positives(negatives) of a specific target time for all parking lots and all evaluated days.

$$TP(t) = \sum_d \sum_l TP(l, d, t)$$

where $TP(l, d, t)$ is 1 if the classification given to parking lot l , target time t on day d is a true positive and 0 otherwise. The variables d and l vary over all evaluated days and parking lots, respectively. The false positives, true negatives and false negatives are calculated in the same way.

The specificity and sensitivity values plotted are calculated according to (7).

Experiment 2 - Fixed target time / varying prediction horizon

For this experiment 3 different target times were chosen: 20:45, 19:00 and 16:00. Each algorithm was run with prediction horizons varying from 15min to 10h and these 3 target times and 1h intervals preceding them for n consecutive days from the data sample. In this project's experiments n was set to 40.

For this experiment the MSE is calculated as:

$$e(h) = \frac{1}{4D \cdot L} \sum_d \sum_l \sum_{i=1}^4 \left(\frac{a_{l,d,t+i,h} - \hat{a}_{l,d,t+i,h}}{c_l} \right)^2$$

Where D is the number of days that were evaluated, L is the number of parking lots, c_l is the total capacity of parking lot l , $a_{l,d,t+i,h}$ and $\hat{a}_{l,d,t+i,h}$ are, respectively, the availability and predicted availability of parking lot l at time $t + i$ on day d considering a prediction horizon of h . The variables d and l vary over all evaluated days and parking lots, respectively.

The number of false(true) positives(negatives) is calculated by summing the false(true) positives(negatives) of a specific target time for all parking lots and all evaluated days.

$$TP(h) = \sum_d \sum_l \sum_{i=1}^4 TP(l, d, t + i, h)$$

Where $TP(l, d, t + i, h)$ is 1 if the classification given to parking lot l , target time $t + i$ on day d with a prediction horizon of h is a true positive and 0 otherwise. The variables d and l vary over all evaluated days and parking lots, respectively. The false positives, true negatives and false negatives are calculated in the same way.

The specificity and sensitivity values plotted are calculated according to (7).

Experiment 3 - Fixed start time / varying prediction horizon

For this experiment 3 different start times were chosen: 19:00, 17:00 and 15:00. Each algorithm was run with these 3 start times and target times from $start\ time + 15min$ up to 20:45 for n consecutive days from the data sample. In this project's experiments n was set to 40.

For this experiment the MSE is calculated as:

$$e(h) = \frac{1}{D \cdot L} \sum_d \sum_l \left(\frac{a_{l,d,st+h,h} - \hat{a}_{l,d,st+h,h}}{c_l} \right)^2$$

Where D is the number of days that were evaluated, L is the number of parking lots, c_l is the total capacity of parking lot l and st is the start time set in the experiment. $a_{l,d,st+h,h}$ and $\hat{a}_{l,d,st+h,h}$ are, respectively, the availability and predicted availability of parking lot l at time $st + h$ on day d considering a prediction horizon of h . The variables d and l vary over all evaluated days and parking lots, respectively.

The number of false(true) positives(negatives) is calculated by summing the false(true) positives(negatives) of a specific target time for all parking lots and all evaluated days.

$$TP(h) = \sum_d \sum_l TP(l, d, st + h, h)$$

Where st is the start time set in the experiment and $TP(l, d, st + h)$ is 1 if the classification given to parking lot l , target time $st + h$ on day d with a prediction horizon of h is a true positive and 0 otherwise. The variables d and l vary over all evaluated days and parking lots, respectively. The false positives, true negatives and false negatives are calculated in the same way.

The specificity and sensitivity values plotted are calculated according to (7).

Results

Experiment 1

As described in the previous subsection, this experiment was realized with 3 different prediction horizons (15min, 1h and 3h), varying the target time for prediction from 8:00 to 20:45.

The MSE of the availability estimate is shown in Figures 17, 18 and 19. It can be seen that for this measure the NHP behaves consistently better than the other models, with the MSM having similar results for lower prediction horizons. As the prediction horizon is increased the differences between the models is reduced, but the NHP keeps performing better.

Experiment 1 with a prediction horizon of 0.25 hours

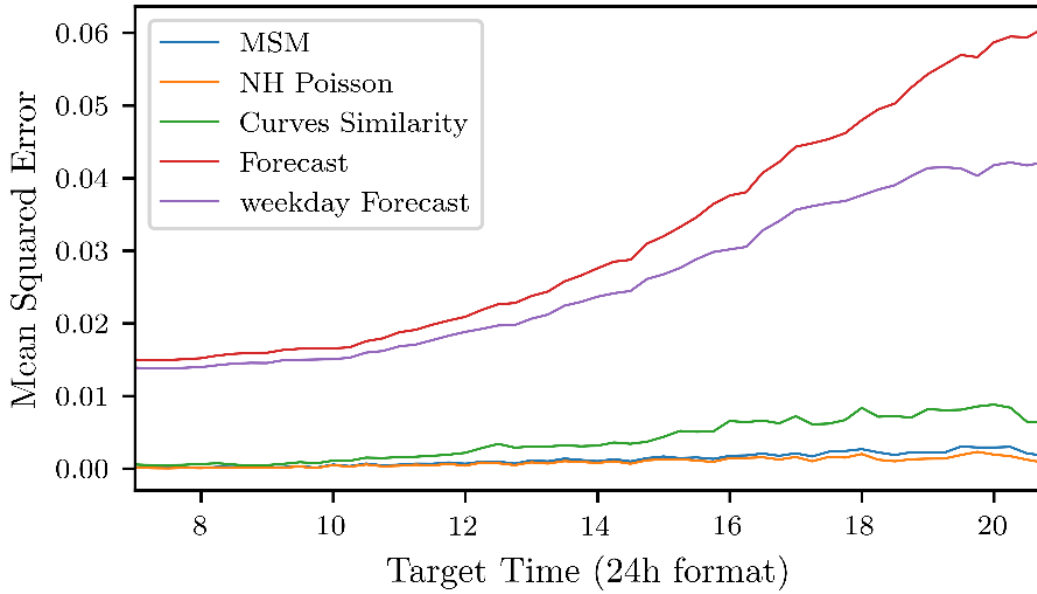


Figure 17. MSE of the availability estimate of experiment 1.1

Experiment 1 with a prediction horizon of 1.0 hours

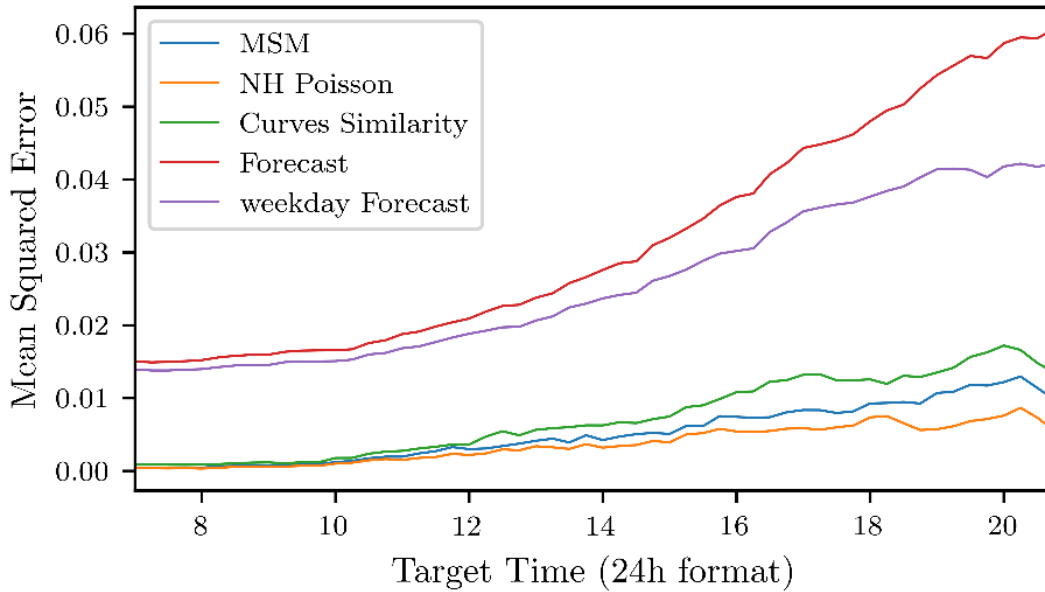


Figure 18. MSE of the availability estimate of experiment 1.2

Experiment 1 with a prediction horizon of 3.0 hours

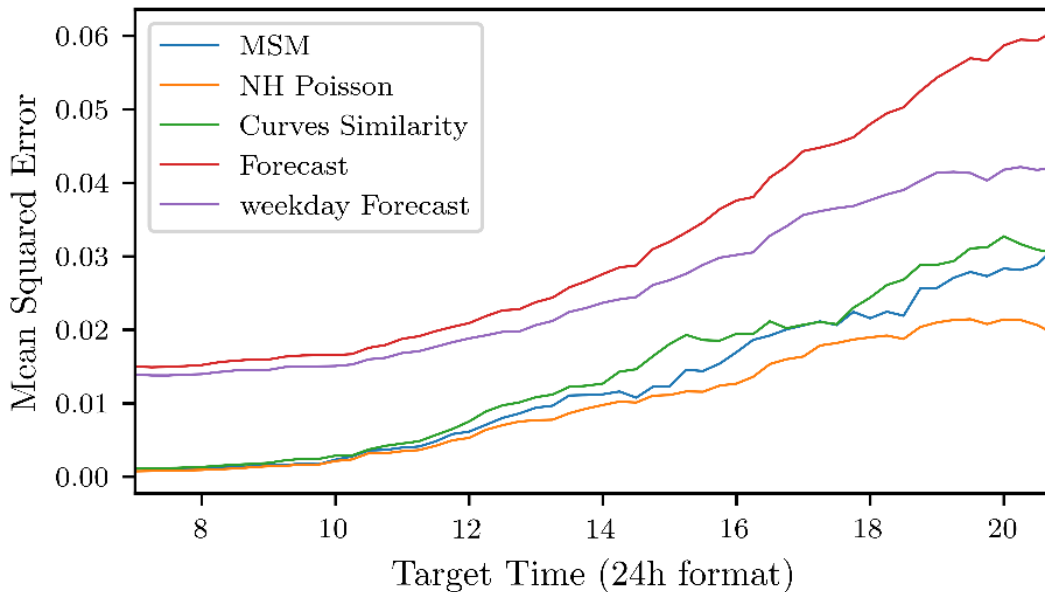


Figure 19. MSE of the availability estimate of experiment 1.3

In Figures 20-22 we can see the performance of the classifier. In this case the classifier is evaluating whether each rest area is full. So, a positive output means that the parking lot is full. The sensitivity data for the beginning of the day is missing as the classifier did not find any true positives or false negatives at those times. It can be seen that the CS, MSM and NHP have very similar sensitivity for all 3 prediction horizons tested. However, the specificity of the MSM and NHP is significantly better for small prediction horizons making them a better choice in this case. For larger prediction horizons all 3 of them seem reasonable choices.

Experiment 1 with a prediction horizon of 0.25 hours

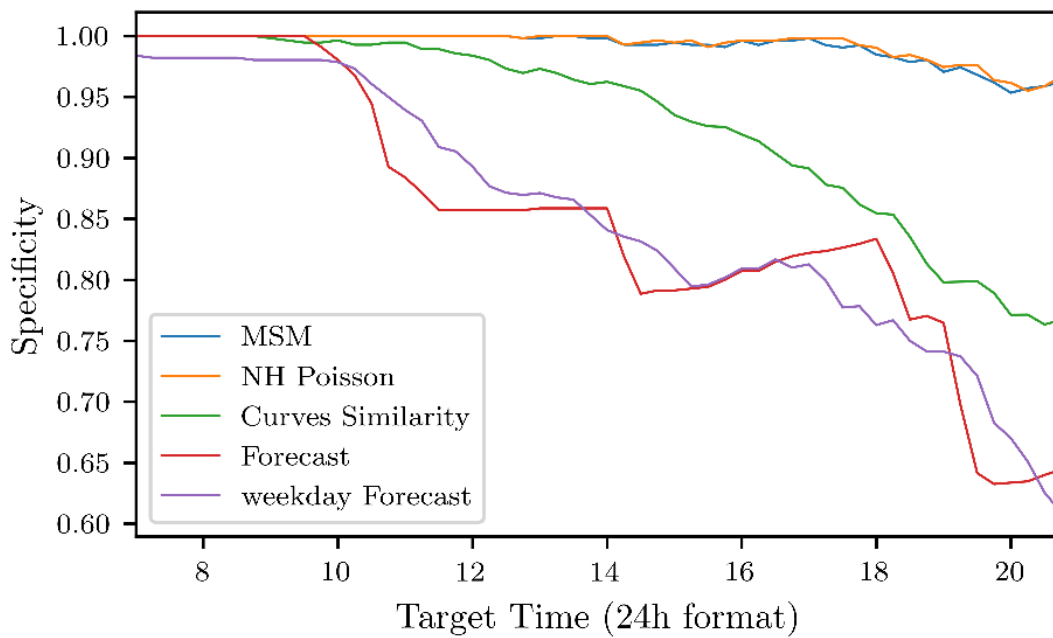
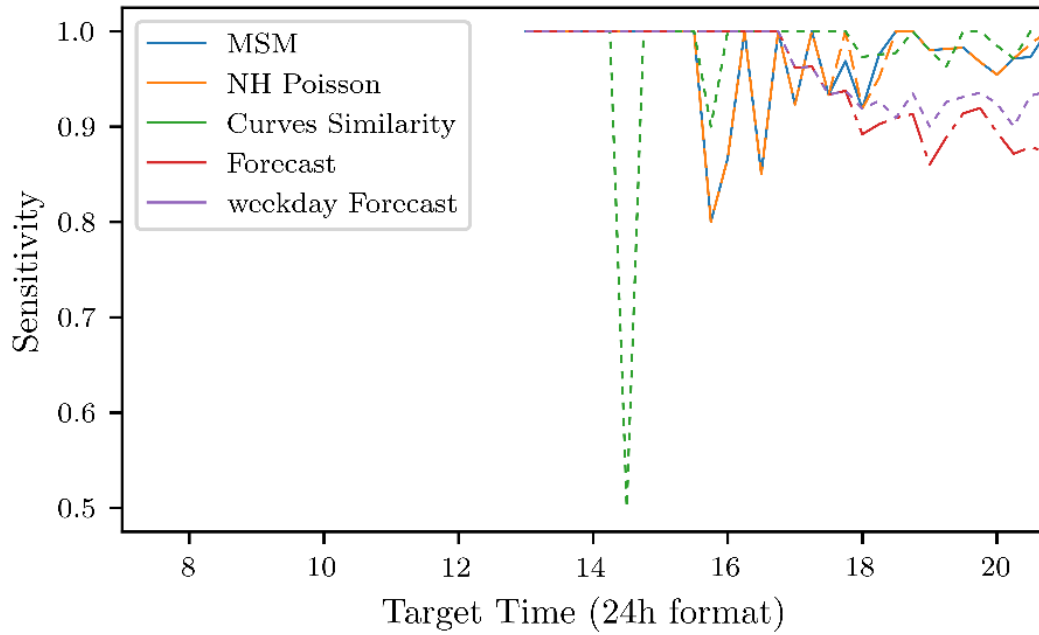


Figure 20. Classification performance of experiment 1.1

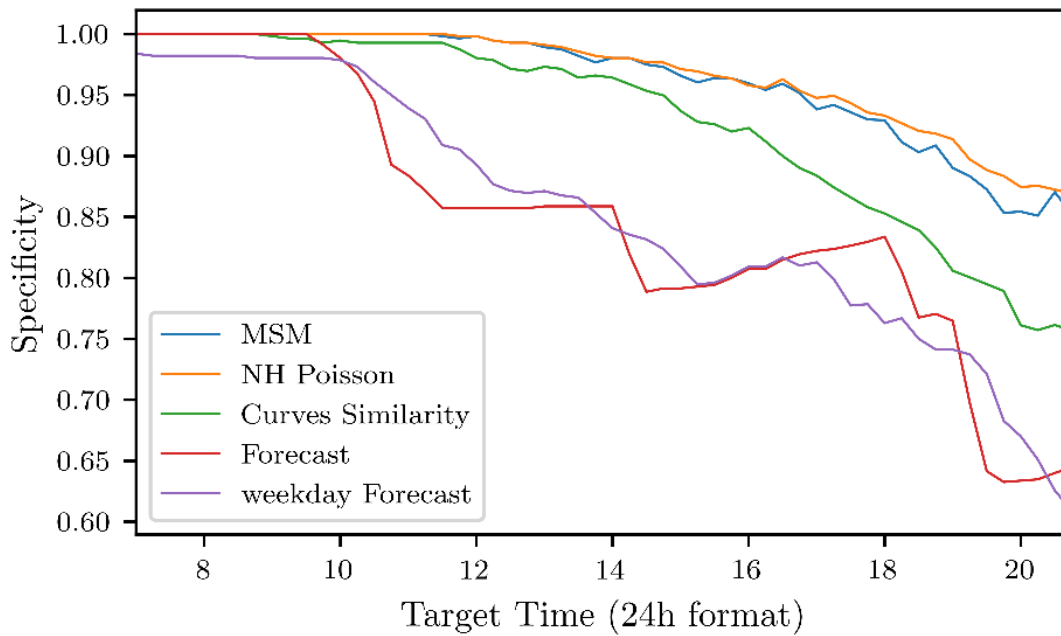
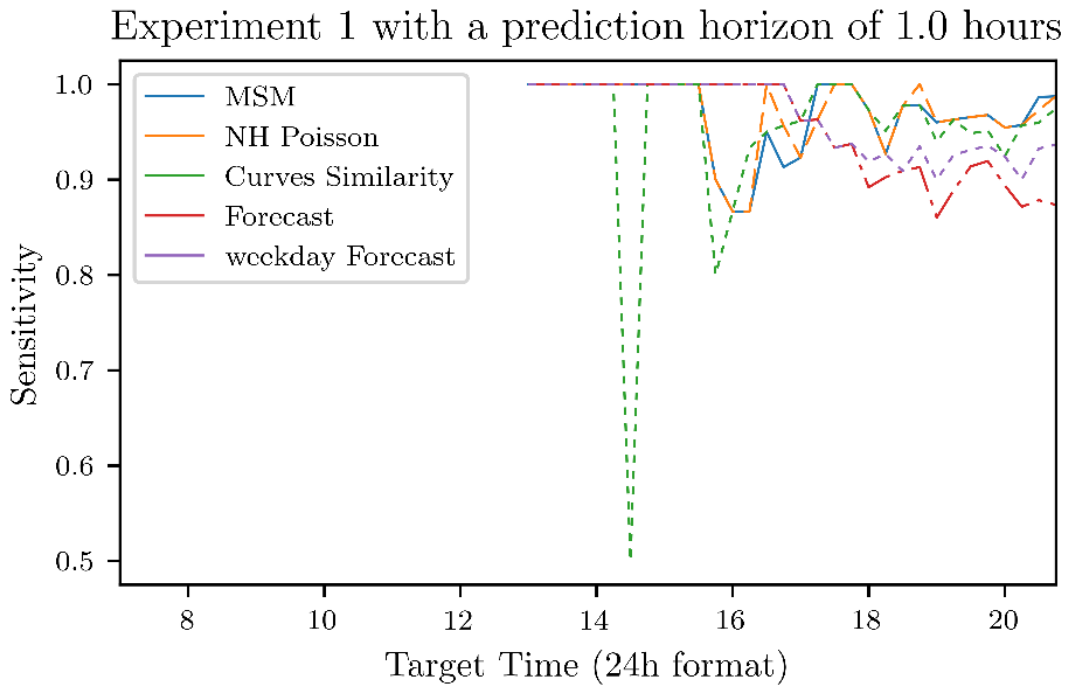


Figure 21. Classification performance of experiment 1.2

Experiment 1 with a prediction horizon of 3.0 hours

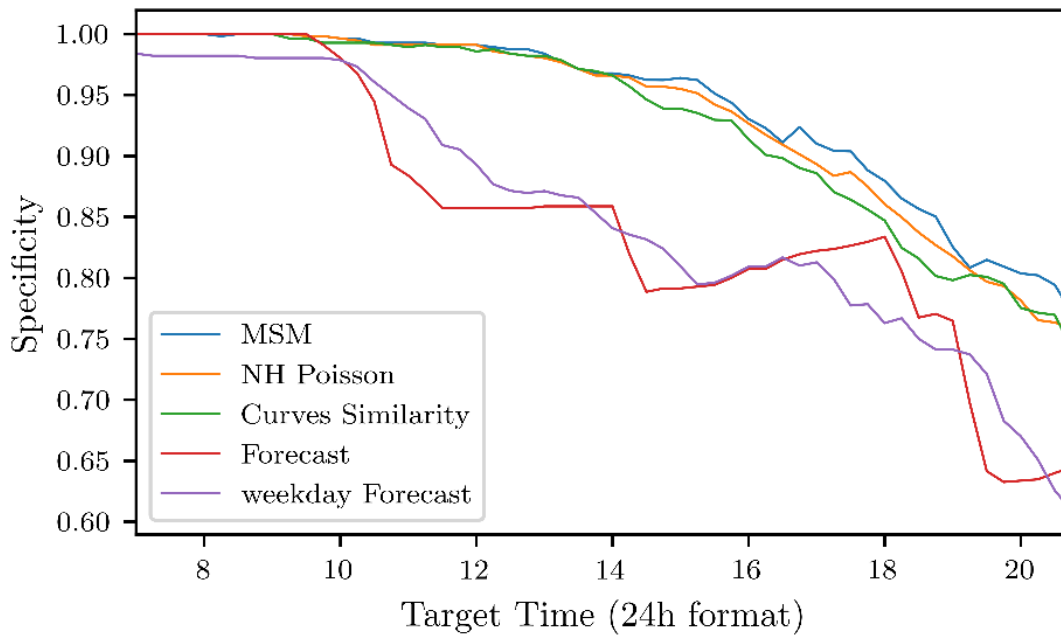
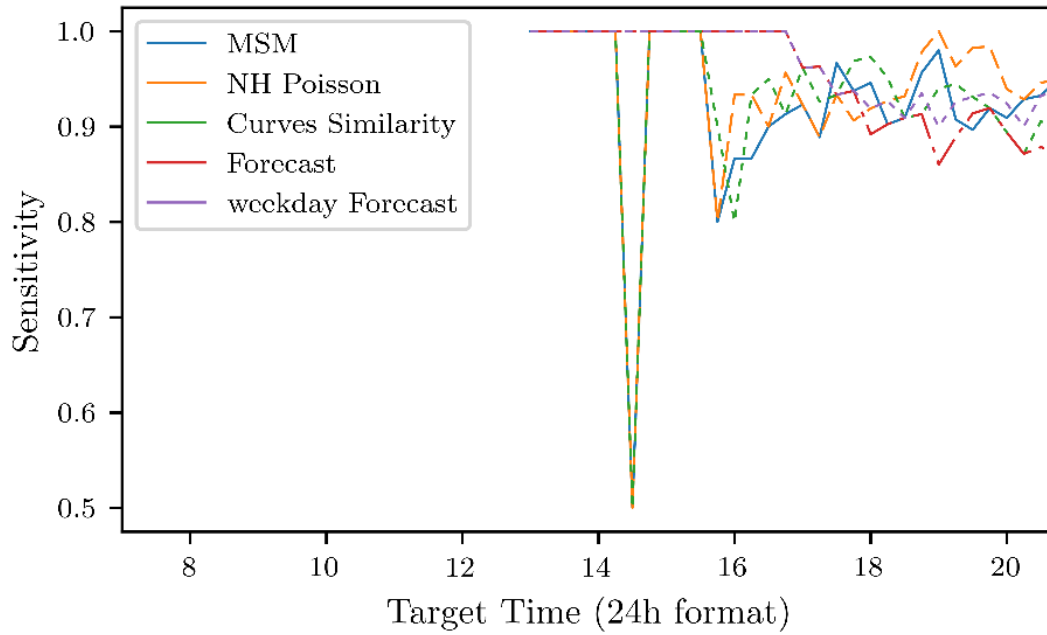


Figure 22. Classification performance of experiment 1.3

Experiment 2

As described in the previous subsection, this experiment was realized with 3 different target times (20:45, 19:00 and 16:00), varying the prediction horizon from 15min to 10h.

The MSE of the availability estimate can be seen in Figures 23, 24 and 25. It can be seen that the NHP still performs better most of the time, but the greater the prediction horizon used the closer it gets to the performance of the forecasting methods. For the earlier times of the day, when the variance of the data is smaller, the NHP, MSM and CS methods still perform better than the forecasting methods, but for later times their performance is very similar to or worse than the forecasting methods' performance when prediction horizons of more than 8 hours are used.

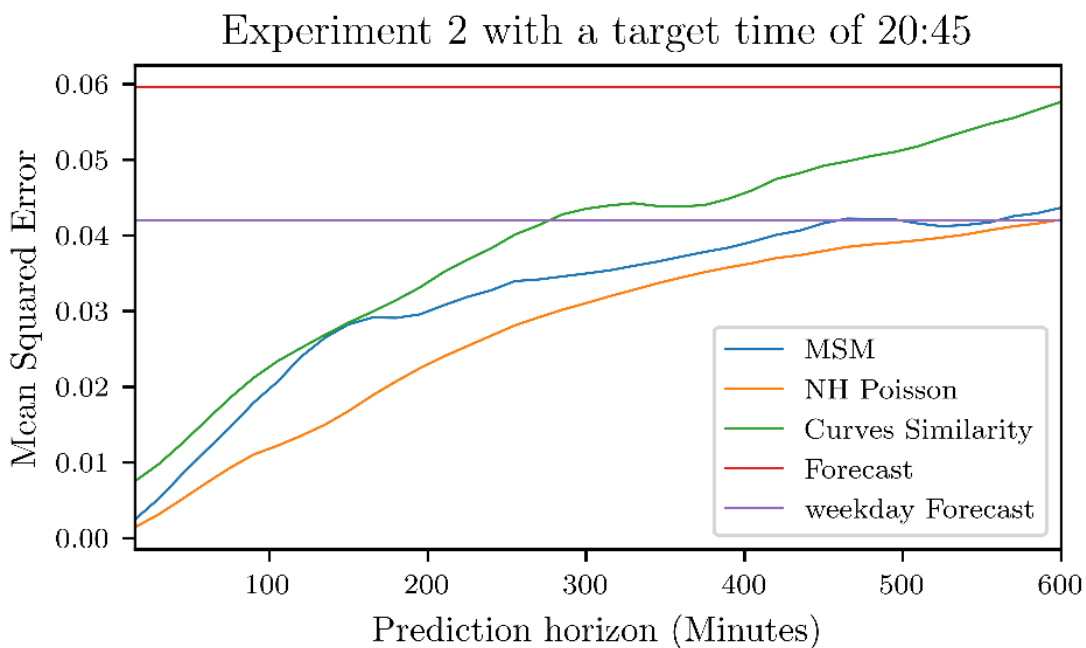


Figure 23. MSE of the availability estimate of experiment 2.1

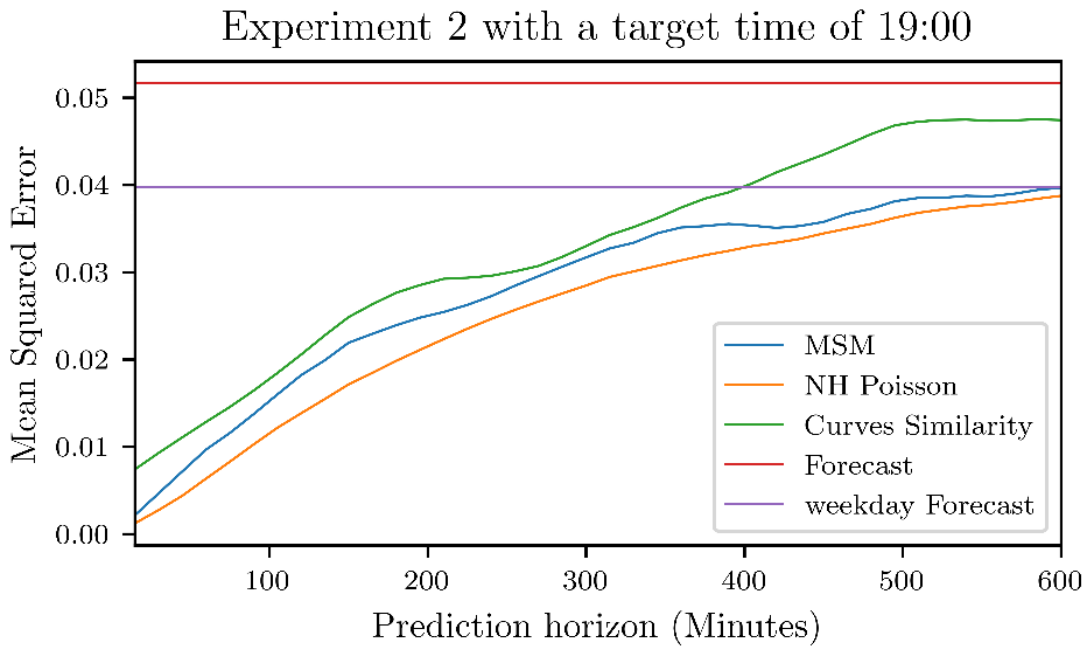


Figure 24. MSE of the availability estimate of experiment 2.2

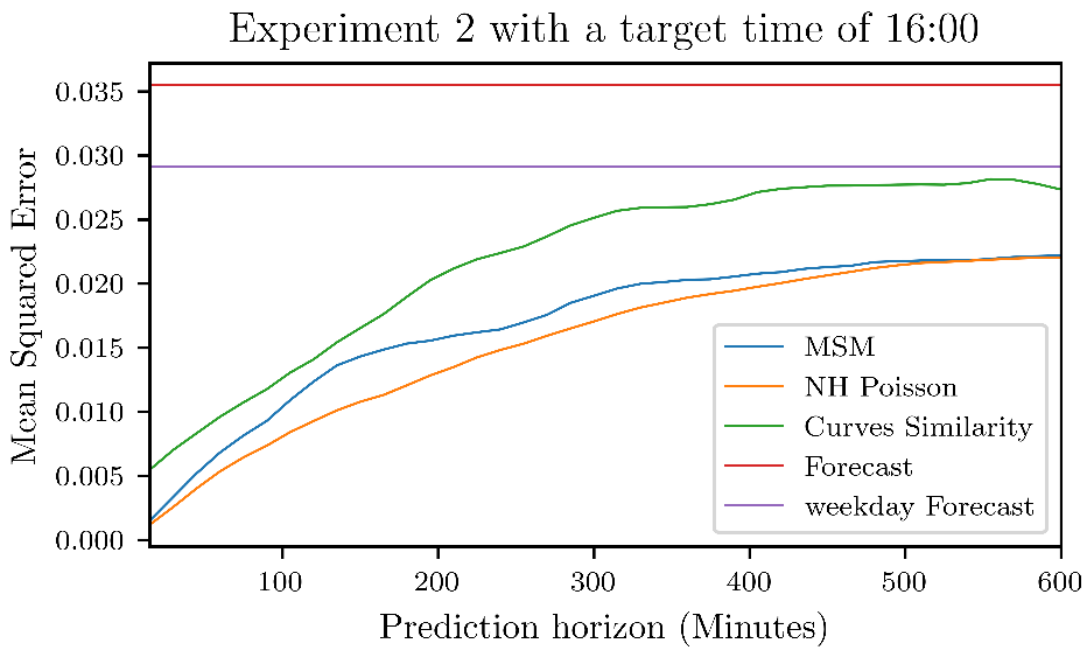


Figure 25. MSE of the availability estimate of experiment 2.3

In Figures 26-28 we can see the performance of the classifier. In this case, the classifier is evaluating whether each rest area is full. So, a positive output means that the parking is full.

The results of experiment 2 are not as homogeneous as for experiment 1. In experiment 2.1 NHP and MSM have a significantly better sensitivity for smaller prediction horizons, and a better specificity in general. Therefore, MSM and NHP would be good choices for small prediction horizons for target times later in the day. However, their sensitivity drops sharply as the prediction horizon increases making the wF method become a better option in these cases as its sensitivity is constant at a decent value and its specificity is not much smaller than the NHP and MSM ones.

The situation with experiment 2.2 is very similar. In this case, NHP and MSM perform similarly to experiment 2.1, with NHP having a slightly better sensitivity and MSM a slightly better specificity. Once more, as the prediction horizon increases wF becomes a better option due to the drop in the MSM and NHP's sensitivities.

In experiment 2.3 the wF has always a better sensitivity while its specificity is lower but not bad. The choice in this case would depend on how much the user prioritizes decreasing the number of false negatives. A decrease in the number of false negatives could be achieved by choosing the method with higher sensitivity, wF in this case. As the NHP and MSM still have a good sensitivity with a better specificity they could be used when a more balanced system is preferred.

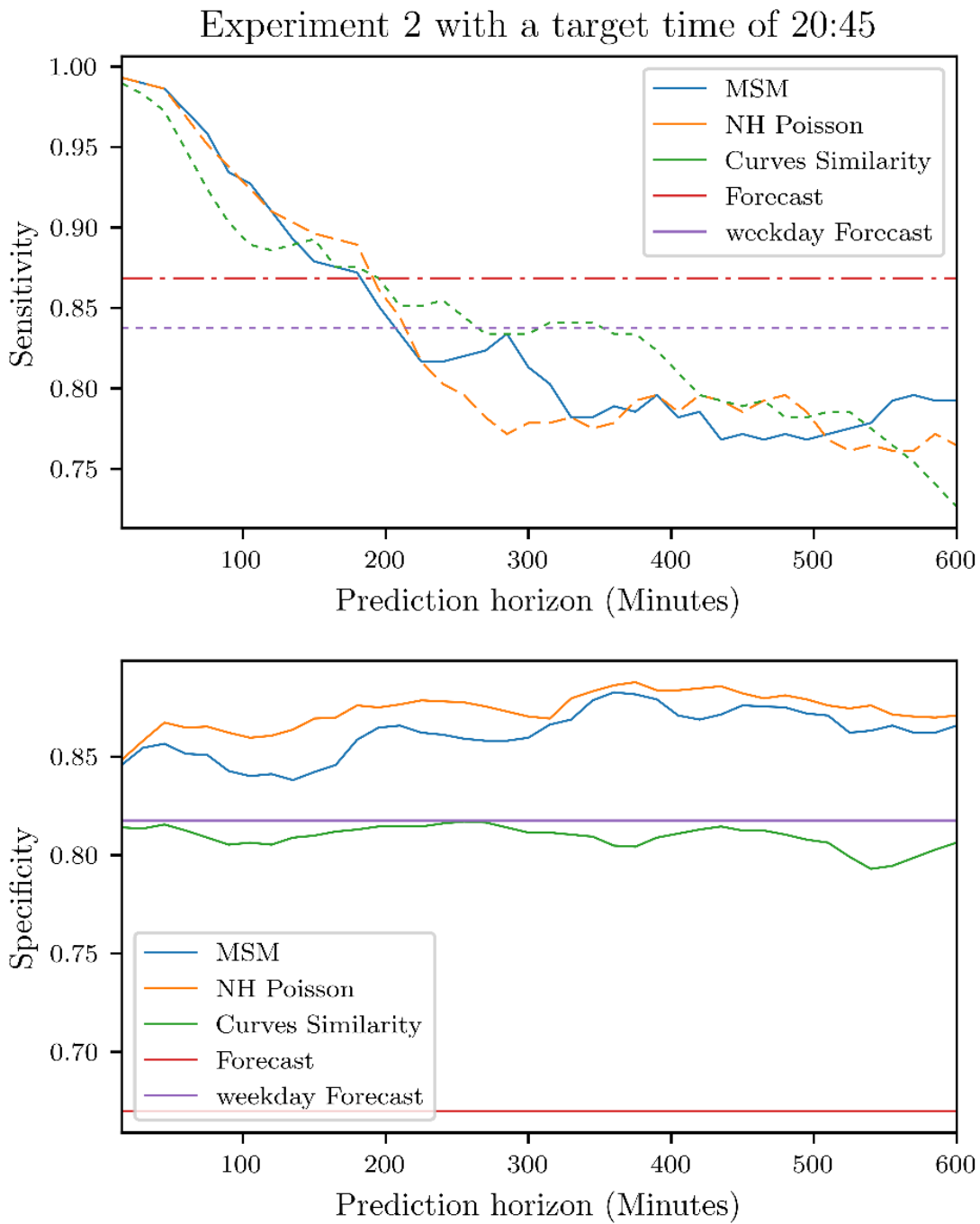


Figure 26. Classification performance of experiment 2.1

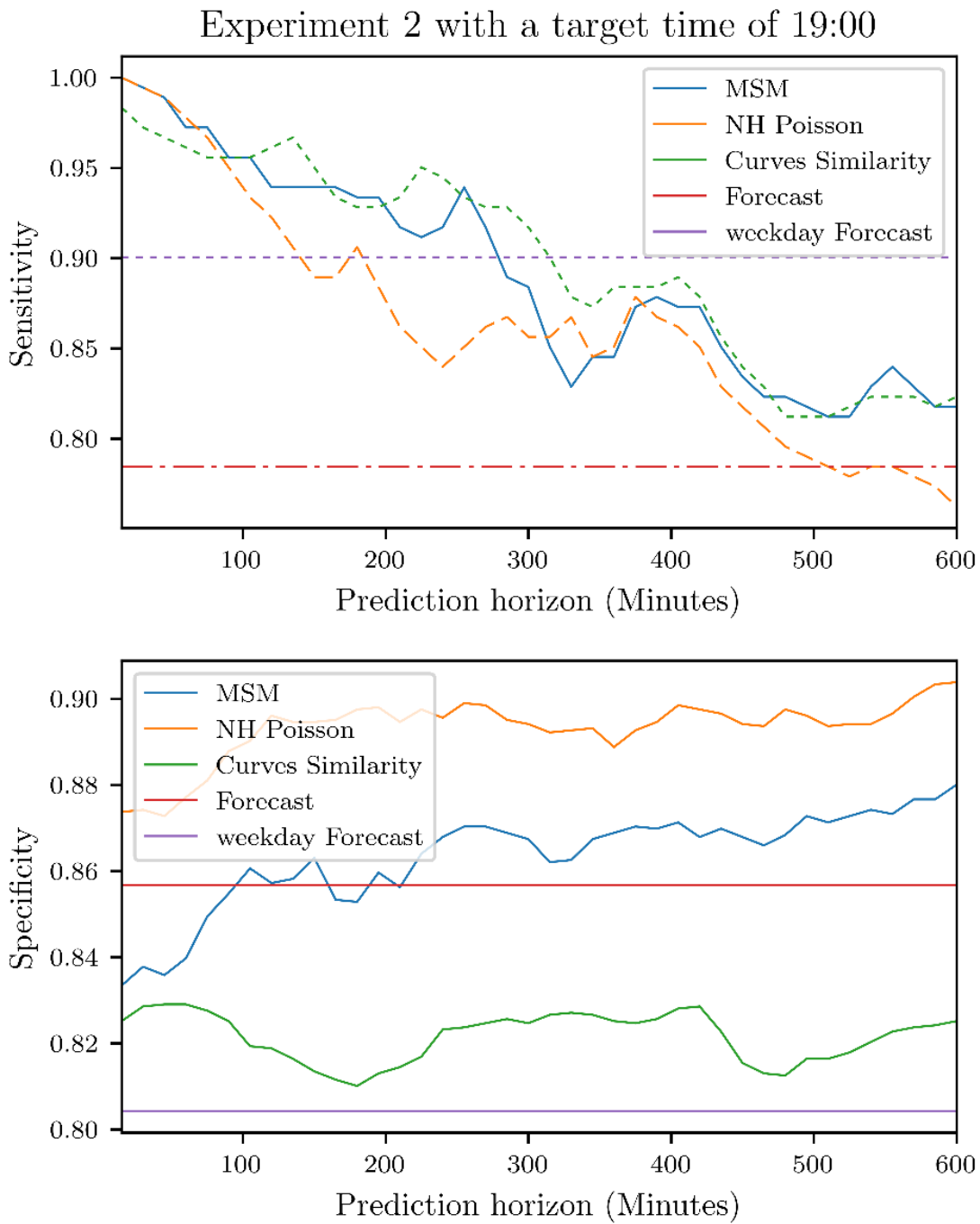


Figure 27. Classification performance of experiment 2.2

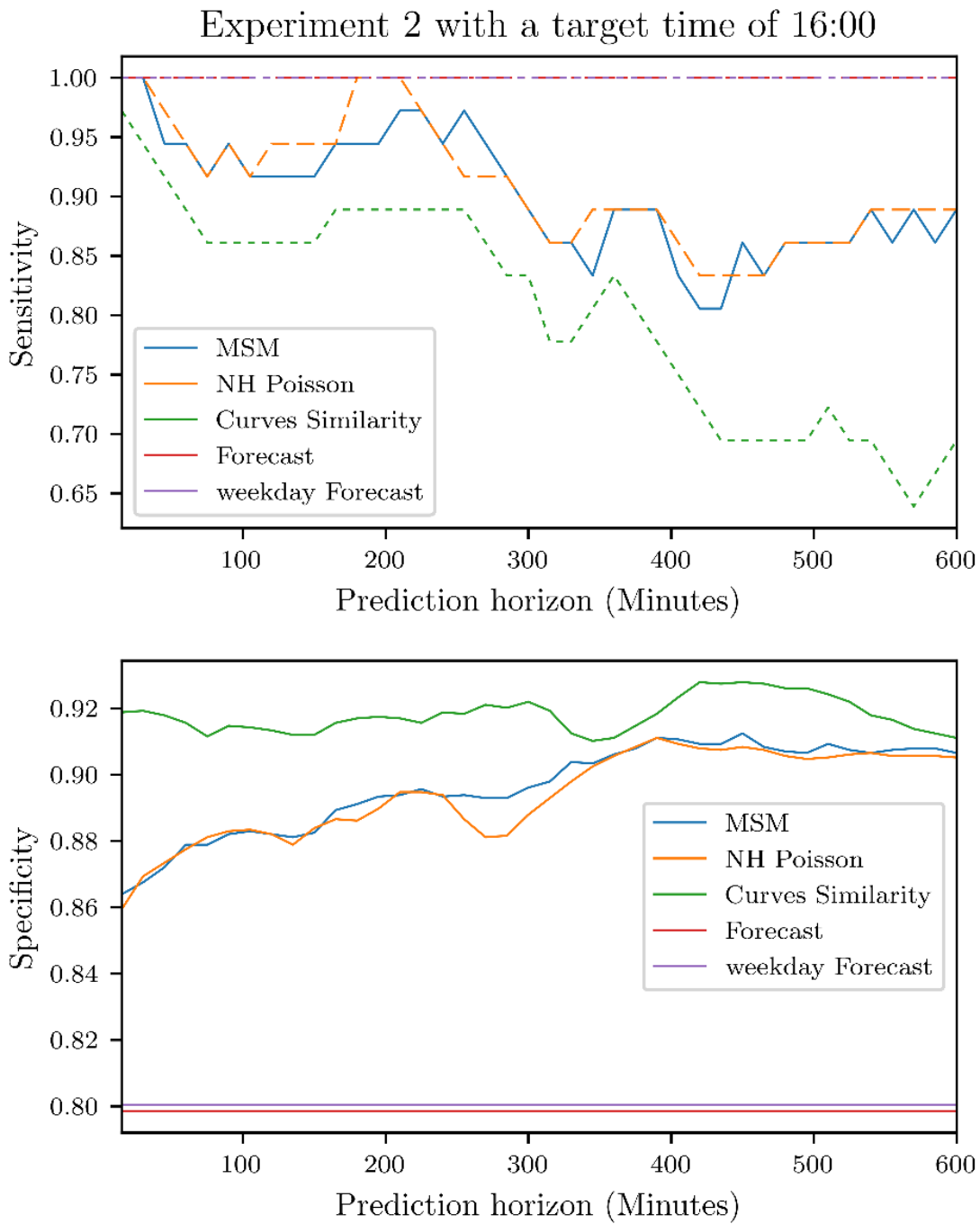


Figure 28. Classification performance of experiment 2.3

Experiment 3

As described in the previous subsection, this experiment was realized with 3 different start times (19:00, 17:00 and 15:00), varying the prediction horizon from 15min to the needed value to get a target time of 20:45.

The MSE of the availability estimate can be seen in Figures 29, 30 and 31. These plots show results consistent with the other experiments. The NHP model still performs better than the other models, with a smaller difference for large prediction horizons. We can also note with this experiment that, with a fixed start time, the MSE of the prediction increases almost linearly with the prediction horizon. Knowing how fast the predictions worsen with the prediction horizon can be useful for the user.

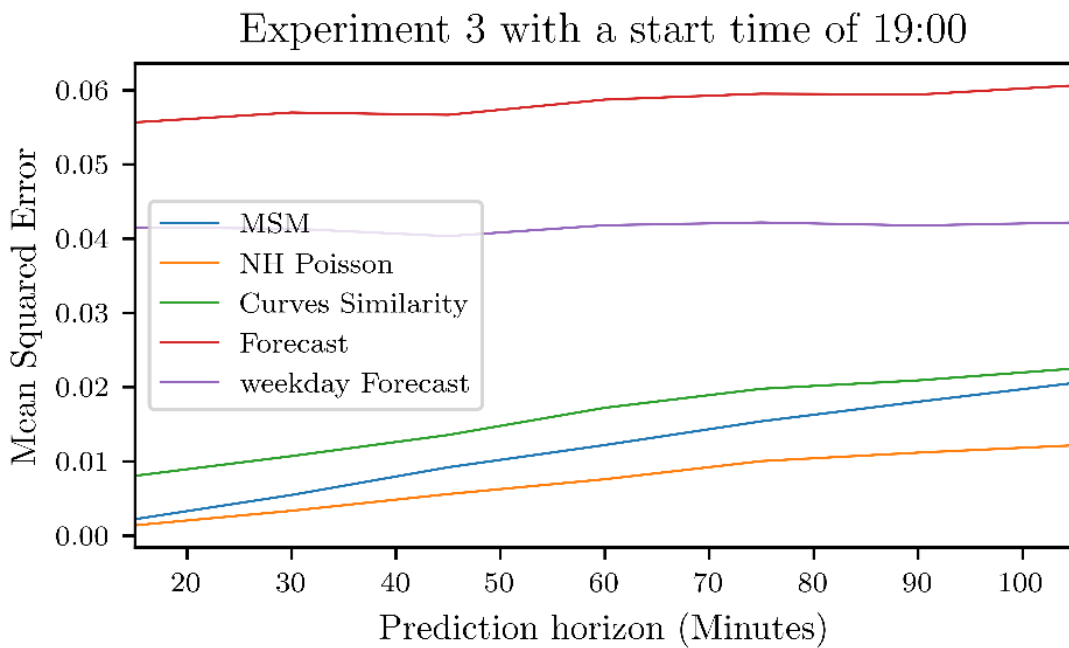


Figure 29. MSE of the availability estimate of experiment 3.1

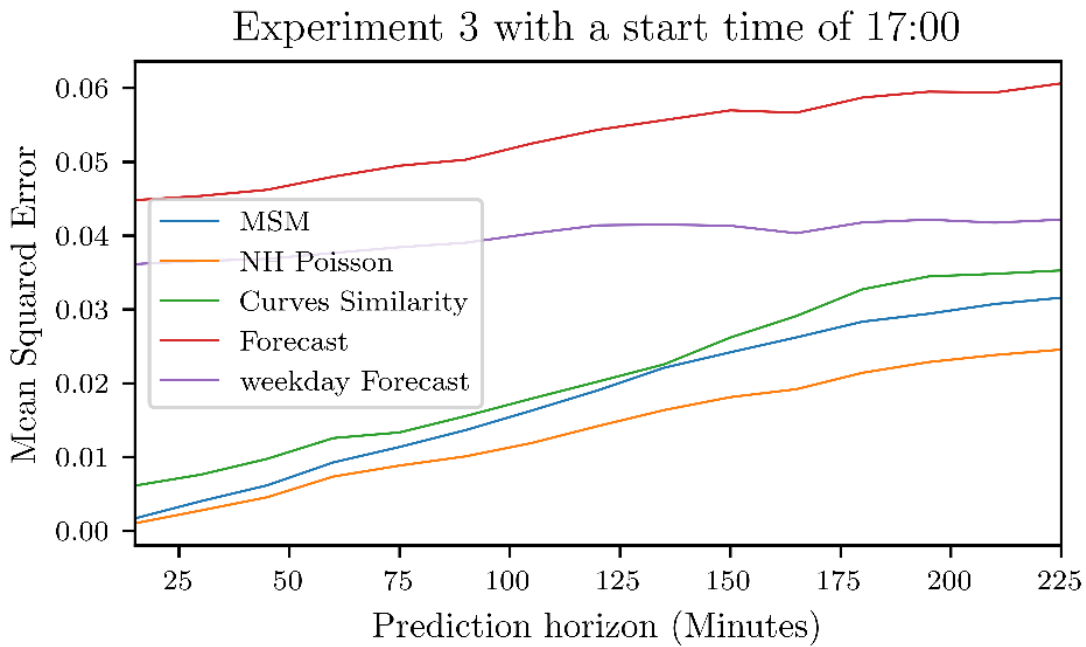


Figure 30. MSE of the availability estimate of experiment 3.2

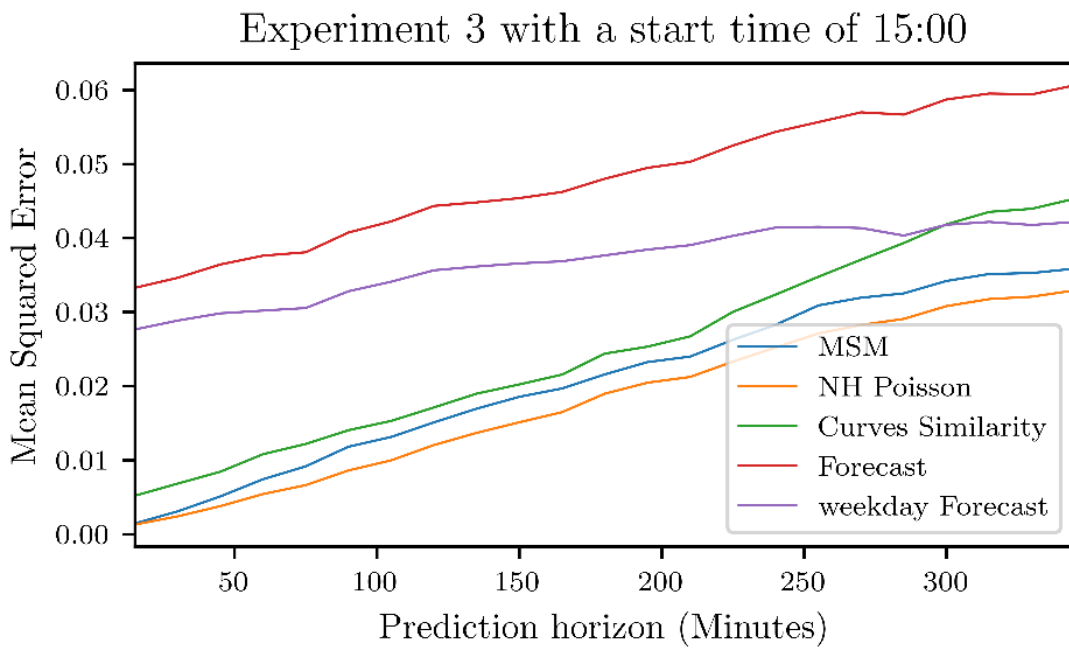


Figure 31. MSE of the availability estimate of experiment 3.3

In Figures 32-34 we can see the performance of the classifier. In this case the classifier is evaluating whether each rest area is full. So, a positive output means that the parking is full.

In experiment 3.1, the CS model had a more balanced performance, with sensitivity ranging from 0.9 to 0.95 and specificity around 0.86. MSM and NHP had better specificity but their sensitivity drops a lot for higher prediction horizons.

In experiment 3.2, the MSM and CS perform similarly overall, with MSM performing slightly better. They perform better for smaller prediction horizons, but their sensitivity falls below the wF one for larger ones, making wF a possible choice for large prediction horizons when sensitivity is being prioritized.

In experiment 3.3, the best alternative for reducing false negatives would be using wF or F for small prediction horizons, with NHP and MSM being more balanced options. For larger prediction horizons both the CS and the MSM methods perform well, with CS being better in sensitivity and MSM better in specificity. As in this problem false negatives have a worse impact than false positives, the preference would go to using the CS method.

Experiment 3 with a start time of 19:00

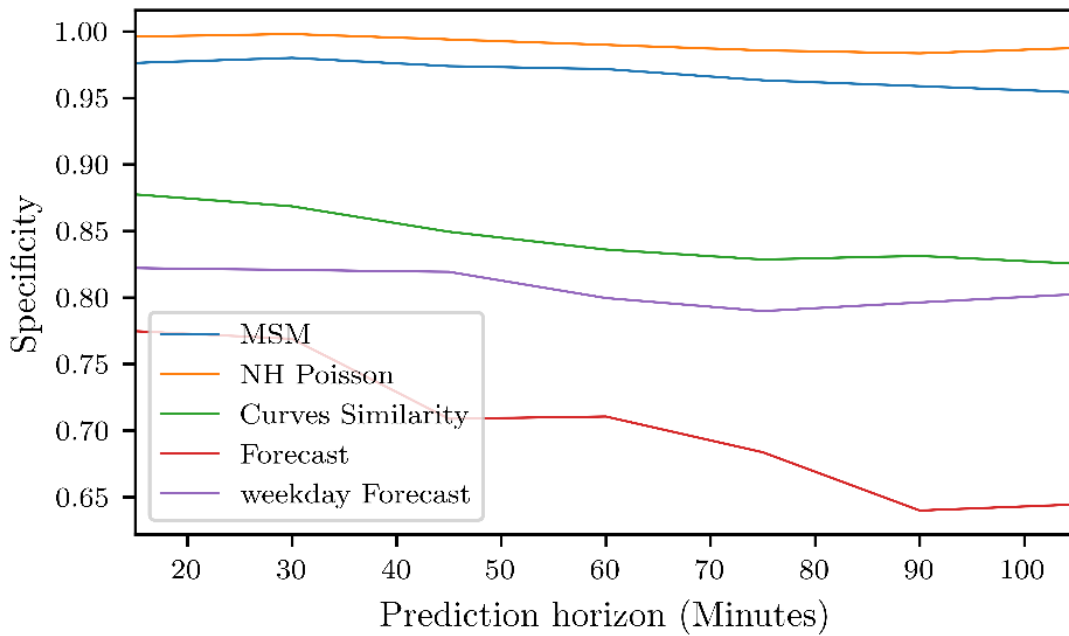
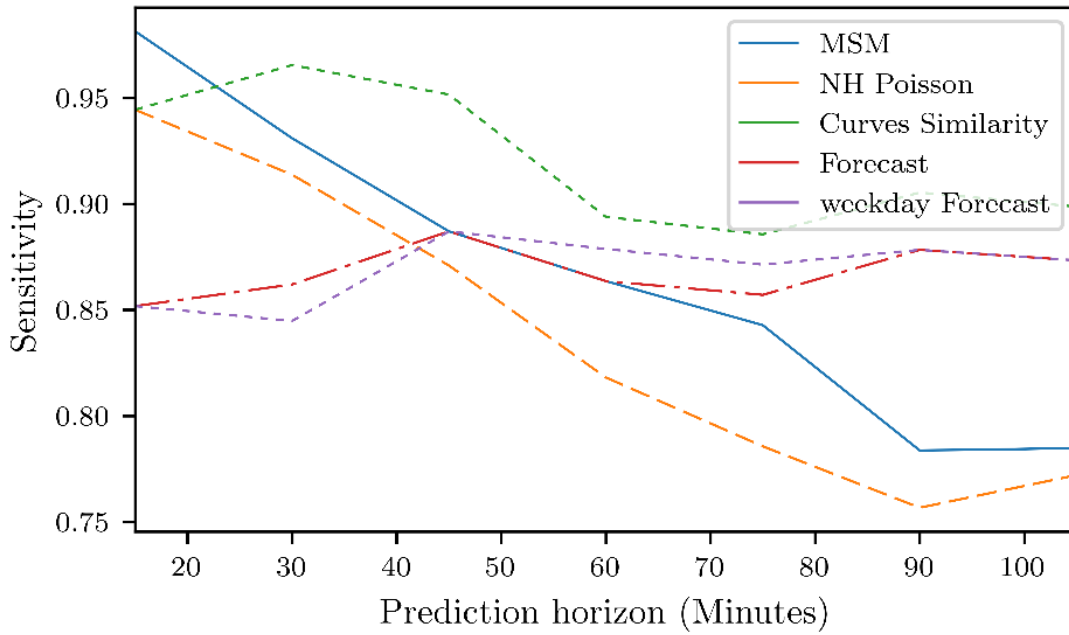


Figure 32. Classification performance of experiment 3.1

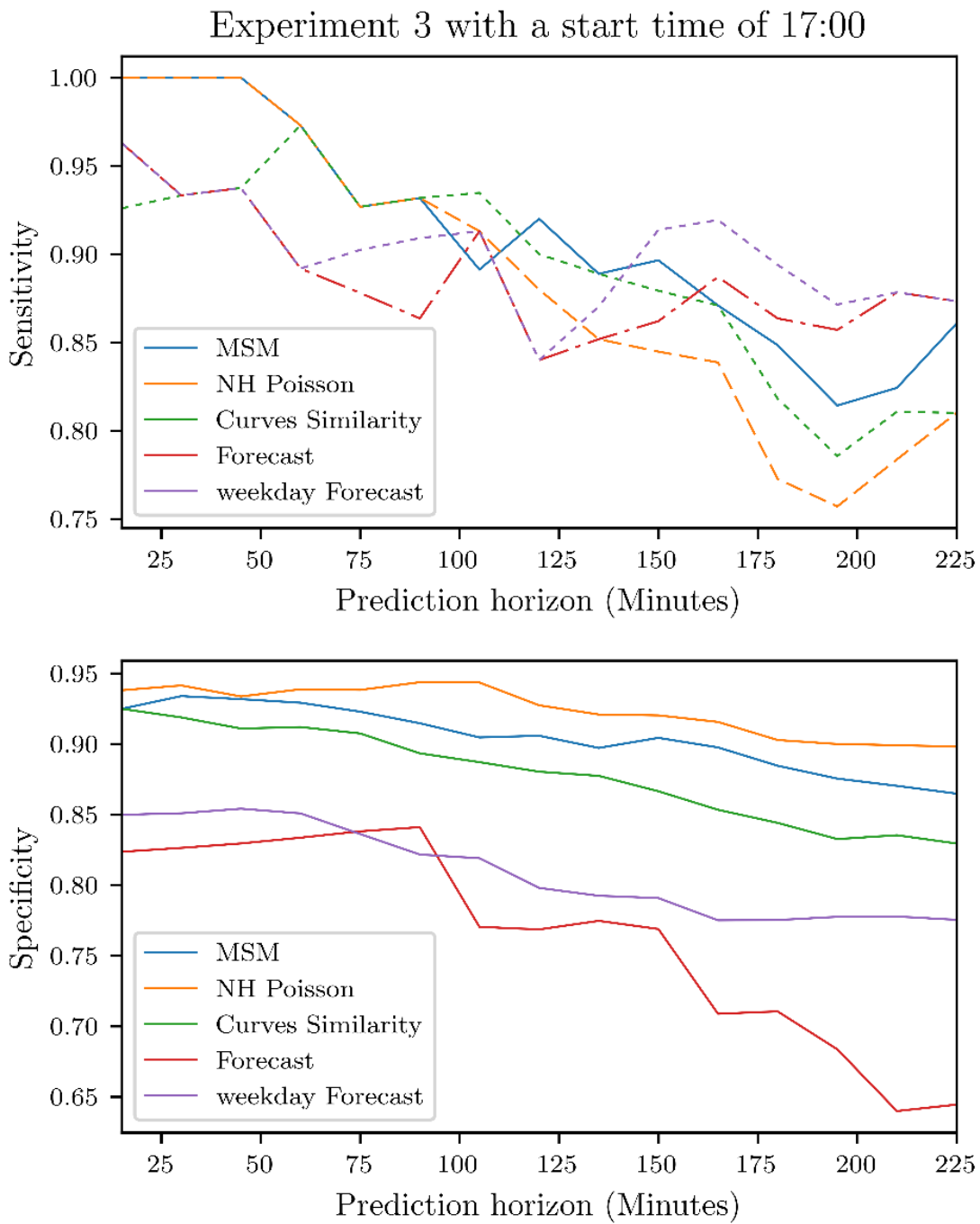


Figure 33. Classification performance of experiment 3.2

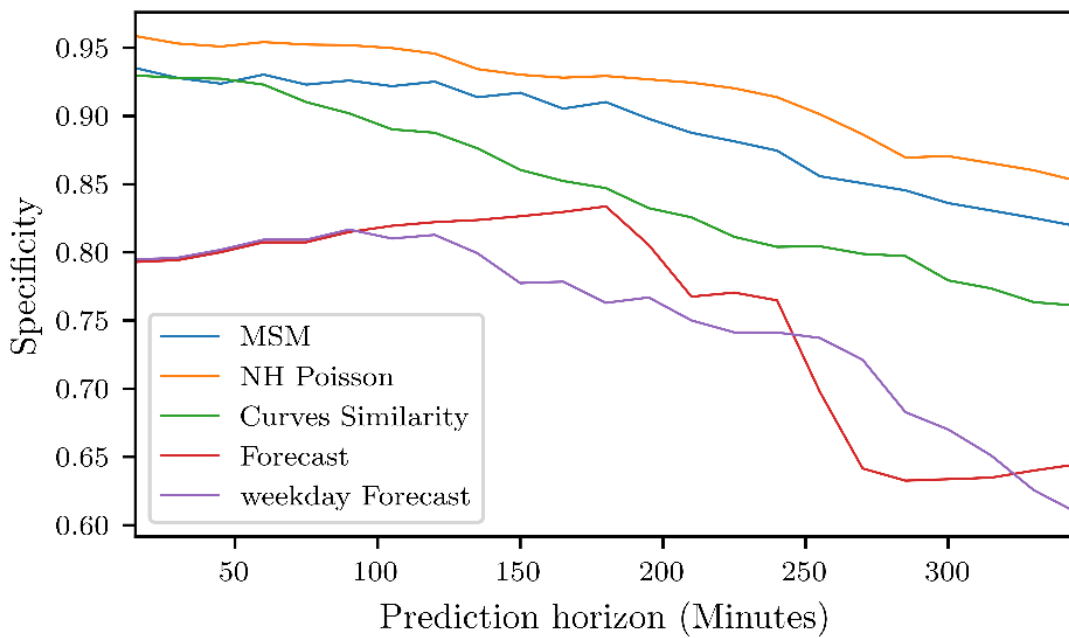
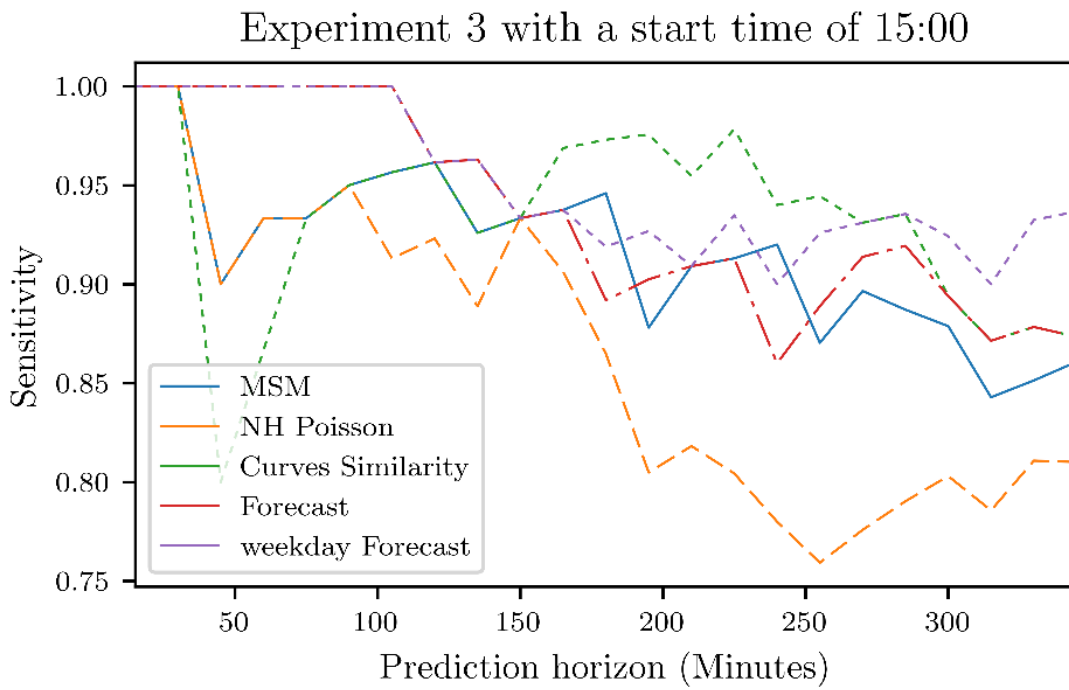


Figure 34. Classification performance of experiment 3.3

Conclusion and Recommendations

The project's objective was to study the truck parking problem and develop a way to generate useful information that could assist truck drivers in better planning their trips. By providing information about parking availability to truck drivers we expect to induce them to better distribute themselves among existing rest areas. This would decrease the peak demand in the most popular truck stops and attenuate the problems caused by the parking shortage.

In this project, several algorithms were tested using data from a company's private truck stops reservation system. It was shown that none of the tested algorithms is absolutely better than the others by having a superior performance in all situations. The results presented show that a more efficient way would be to combine them and use the most appropriate one according to the situation. It was also shown that the estimator with the smallest MSE is not necessarily the one that gives the most useful information to the decision model.

The different experiments realized show the performance of each method in different situations and allow us to suggest a combination of them that would perform better than any single one of them. Combinations that prioritize more sensitivity and a more balanced one were included. In a real application both could be implemented, and the user could check whether this change of priorities would impact the classification results and, consequently, his planning. These suggestions can be found on Table 2.

Table 2. Suggested model usage

Current Time	Target Time	Method (balanced)	Method (sensitivity)
<17:00	<17:00	NHP	wF
<17:00	>17:00	CS	CS
17:00 – 19:00	17:00 – 21:00	MSM/NHP	MSM/NHP
19:00 – 21:00	19:00 – 21:00	CS	CS

It follows from the experimental results that with the exception of the situation of predictions for early in the day, the more complex models usually behave better than the simple forecasting ones. This indicates that, even if the drivers have some sense of the expected status of the parking lots due to experience, this system could still provide them with valuable information.

Future research directions would be:

- Increase the choice of models and time resolution of the model assignment;
- Use this parking prediction in higher levels of planning, not only for micro adjustments, as where to stop in a given route, but also to decide the routing and departure time of the trucks.

Implementation

This study could be implemented by means of a smartphone application. The use of this application by truck drivers would help reduce the peak demand for the most popular rest areas and the number of drivers parking illegally in their vicinities. The application would tell the drivers in advance that those rest areas will be full by the time they get there, and which truck stops in the region are likely to have available spaces. This will help better distribute the trucks among existing rest areas. The extent of these benefits will be dependent on the market penetration of the application as well as on the level of user compliance to the application suggestions and on the existing parking infrastructure.

In order for it to be implemented, truck parking availability data for the regions of operation would be needed. Private truck stops might already have some infrastructure to provide data as they have to control the payments, especially the ones that have reservation systems. Public stops however, would need an investment in infrastructure in order to collect that data. Investment would also be needed to develop, implement and maintain the application, in both hardware and software.

There are several ways to collect occupancy data, from installing gates and registering how many vehicles enter and exit, to installing magnetic sensors in every single parking space in the parking lot. Some pilot projects were already implemented both in Europe and in the USA testing different kinds of systems [19], [34]–[37], [39]–[43]. The choice of method will depend on the budget available and specific restrictions for each parking lot, e.g., some sensors might not work very well in snowy or rainy weather, others might not perform well under hot temperatures.

References

- [1] American Transportation Research Institute, “Critical Issues in the Trucking Industry – 2015,” Arlington, VA, 2015.
- [2] J. Sochor and G. Mbiydzennyuy, “Assessing the Benefits of Intelligent Truck Parking,” *Int. J. Intell. Transp. Syst. Res.*, vol. 11, no. 2, pp. 43–53, 2013.
- [3] U.S. Department of Transportation and Office of Freight Management and Operations, “Jason’s Law Truck Parking Survey Results and Comparative Analysis,” 2015.
- [4] J. G. Pigman, G. Winchester, and N. Swallom, “Commercial Truck Parking and Other Safety Issues,” 2015.
- [5] SETPOS Consortium, “Background Information and Considerations for Secure Truck Parking,” 2009.
- [6] U.S. Department Of Transportation, “Freight Facts and Figures 2015,” 2015.
- [7] S. Shaheen and C. J. Rodier, “Commercial Vehicle Parking in California: Exploratory Evaluation of the Problem and Possible Technology-Based Solutions,” 2007.
- [8] Kimley Horn, “Virginia Truck Parking Study,” 2015.
- [9] B. C. Tefft, “Prevalence of Motor Vehicle Crashes Involving Drowsy, United States, 2009-2013,” Washington, DC, 2014.
- [10] I. Banerjee, J. Lee, K. Jang, S. Pande, and D. Ragland, *Rest Areas – Reducing Accidents Involving Driver Fatigue*. California PATH Program, Institute of Transportation Studies, University of California at Berkeley, 2010.
- [11] American Transportation Research Institute, “Compendium of idling regulations,” 2016.
- [12] M. Palaniappan, D. Wu, and J. Kohleriter, “Clearing the Air: Reducing Diesel Pollution in West Oakland,” Oakland, CA, 2003.
- [13] N. Lutsey, C.-J. Brodrick, D. SPERLING, and C. Oglesby, “Heavy-Duty Truck Idling Characteristics: Results From a Nationwide Truck Survey,” *Transp. Res. Rec.*, no. 1880, pp. 29–38, 2004.
- [14] I. W. Ford Torrey and D. Murray, “An Analysis of the Operational Costs of Trucking: 2016 Update,” Arlington, VA, 2016.
- [15] United States Environmental Protection Agency, “What You Should Know About Truck Engine Idling, April 2002,” no. April. 2002.
- [16] S. A. Fleger, R. P. Haas, J. W. Trombly, R. H. Cross III, J. E. Noltenius, K. K. Pécheux, and K. J. Chen, “STUDY OF ADEQUACY OF COMMERCIAL TRUCK PARKING FACILITIES - TECHNICAL REPORT,” 2002.

- [17] F. M. Heinitz and N. Hesse, "Estimating Time-Dependent Demand for Truck Parking Facilities Along a Federal Highway," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2097, no. 1, pp. 26–34, 2009.
- [18] M. E. Bayraktar, F. Arif, H. Ozen, and G. Tuxen, "Smart Parking-Management System for Commercial Vehicle Parking at Public Rest Areas," *J. Transp. Eng.*, vol. 141, no. 5, p. 4014094, May 2015.
- [19] F. Bayraktar, Mehmet Emre;Zhu, Yimin; Arif, "Commercial Motor Vehicle Parking Trends At Rest Areas and Weigh Stations," no. December, p. 140, 2012.
- [20] T. Rajabioun and P. Ioannou, "On-Street and off-street parking availability prediction using multivariate spatiotemporal models," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2913–2924, 2015.
- [21] E. I. Vlahogianni, K. Kepaptsoglou, V. Tsetos, and M. G. Karlaftis, "A Real-Time Parking Prediction System for Smart Cities," *J. Intell. Transp. Syst.*, vol. 20(2), pp. 192–204, 2016.
- [22] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach," *Transp. Res. Part C Emerg. Technol.*, vol. 13, no. 3, pp. 211–234, 2005.
- [23] T. Fabusuyi, R. C. Hampshire, V. A. Hill, and K. Sasanuma, "Decision analytics for parking availability in downtown pittsburgh," *Interfaces (Providence)*, vol. 44, no. 3, pp. 286–299, 2014.
- [24] F. Caicedo, C. Blazquez, and P. Miranda, "Prediction of parking space availability in real time," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7281–7290, 2012.
- [25] M. Caliskan, A. Barthels, B. Scheuermann, and M. Mauve, "Predicting Parking Lot Occupancy in Vehicular Ad Hoc Networks," *2007 IEEE 65th Veh. Technol. Conf. VTC2007Spring*, pp. 277–281, 2007.
- [26] A. Klappenecker, H. Lee, and J. L. Welch, "Finding available parking spaces made easy," *Ad Hoc Networks*, vol. 12, no. 1, pp. 243–249, 2014.
- [27] S. Pullola, P. K. Atrey, and A. El Saddik, "Towards an intelligent GPS-based vehicle navigation system for finding street parking lots," *ICSPC 2007 Proc. - 2007 IEEE Int. Conf. Signal Process. Commun.*, no. November, pp. 1251–1254, 2007.
- [28] R. Beheshti and G. Sukthankar, "A hybrid modeling approach for parking and traffic prediction in urban simulations," *AI Soc.*, vol. 30, no. 3, pp. 333–344, 2014.
- [29] J. F. Garc, V. R. Tom, L. A. Garc, and J. J. Mart, "An Autonomic System for Intelligent Truck Parking," 2014.
- [30] M. Capdevila, V. R. Tomas, L. A. Garca, and M. P. Farron, "Dynamic management of parking spaces in road rest areas with automatic negotiation," *Proc. - 2013 IEEE Int. Conf. Syst. Man, Cybern. SMC 2013*, pp. 3609–3614, 2013.

- [31] Y. Geng and C. G. Cassandras, "Dynamic resource allocation in urban settings: A 'smart parking' approach," *Proc. IEEE Int. Symp. Comput. Control Syst. Des.*, pp. 1–6, 2011.
- [32] N. Doulamis, E. Protopapadakis, and L. Lambrinos, "Improving service quality for parking lot users using intelligent parking reservation policies," *Proc. - 27th Int. Conf. Adv. Inf. Netw. Appl. Work. WAINA 2013*, pp. 1392–1397, 2013.
- [33] SETPOS Partners, "Secured European Truck Parking Best Practice Handbook," 2010.
- [34] EasyWay Project, "EasyWay Highlights." [Online]. Available: <https://easyway12.its-platform.eu/news/publications-within-ewi-and-ewii-projects>. [Accessed: 28-Nov-2016].
- [35] EasyWay Project, "Freight \& Logistics Services INTELLIGENT AND SECURE TRUCK PARKING Deployment Guideline," 2015.
- [36] J. Dierke, J. Kleine, and R. Lehmann, "Intelligent Controlled Compact Parking for Modern Parking Management on German Motorways," *Transp. Res. Procedia*, vol. 15, pp. 620–627, 2016.
- [37] Siemens AG, "Help for road-weary truckers: Siemens installs truck parking information system," 2015. [Online]. Available: www.siemens.com/press/hgv-guidance-system. [Accessed: 05-Oct-2016].
- [38] Nedap, "Guiding truckers in port of Hamburg." [Online]. Available: <http://www.nedapidentification.com/solutions/cases/guiding-truckers-in-port-of-hamburg.html> (Accessed 2017-01-30). [Accessed: 30-Jan-2017].
- [39] T. Morris, N. Papanikolopolous, D. Murray, K. Fender, A. Weber, V. Morellas, and D. Cook, "A Comprehensive System for Assessing Truck Parking Availability," Minneapolis, MN.
- [40] D. J. Cook, T. Morris, V. Morellas, and N. Papanikolopoulos, "An automated system for persistent real-time truck parking detection and information dissemination," *Proc. - IEEE Int. Conf. Robot. Autom.*, pp. 3989–3994, 2014.
- [41] J. Fallon and K. Howard, "SmartPark Truck Parking Availability System: Magnetometer Technology Field Operational Test Results," 2011.
- [42] J. Gertler and J. Murray, "SmartPark Truck Parking Availability System: Video Technology Field Operational Test Results," 2011.
- [43] V. López-Jacobs, J. Ellerbee, and M. Hoover, "SmartPark Technology Demonstration Project," Camp Hill, PA, 2013.
- [44] S. Yoo, P. K. Chong, T. Kim, J. Kang, and D. Kim, "PGS : Parking Guidance System based on wireless sensor network PGS : Parking Guidance System based on Wireless Sensor Network," in *Wireless Pervasive Computing, 2008. ISWPC 2008. 3rd International Symposium on.*, 2008, pp. 218–222.
- [45] E. Sifuentes, O. Casas, and R. Pallas-Areny, "Wireless magnetic sensor node for vehicle detection with optical wake-up," *IEEE Sens. J.*, vol. 11, no. 8, pp. 1669–1676, 2011.

- [46] J. Gu, Z. Zhang, F. Yu, and Q. Liu, "Design and Implementation of a Street Parking System Using Wireless Sensor Networks," in *Industrial Informatics (INDIN), 2012 10th IEEE International Conference on.*, 2012, pp. 1212–1217.
- [47] N. Larisis, L. Perlepes, P. Kikiras, and G. Stamoulis, "U-Park : Parking Management System Based on Wireless Sensor Network Technology," in *International Conference on Sensor Technologies and Applications (SENSORCOMM)*, 2012, pp. 170–177.
- [48] Siemens AG, "Sitraffic Wimag," 2013. [Online]. Available: <https://www.mobility.siemens.com/mobility/global/SiteCollectionDocuments/en/road-solutions/urban/infrastructure/Sitraffic-Wimag-en.pdf>. [Accessed: 06-Oct-2016].
- [49] A. Haghani, S. Farzinfard, M. Hamedi, F. Ahdi, and M. K. Khandani, "Automated Low-Cost and Real-Time Truck Parking Information System," COLLEGE PARK, 2013.
- [50] D. Bong, K. Ting, and K. Lai, "Integrated approach in the design of car park occupancy information system (COINS)," *IAENG Int. J. Comput. Sci.*, no. February, 2008.
- [51] R. Yusnita, N. Fariza, and B. Norazwinawati, "Intelligent Parking Space Detection System Based on Image Processing," *Int. J. Innov. Manag. Technol.*, vol. 3, no. 3, pp. 232–235, 2012.
- [52] J. K. Suhr and H. G. Jung, "Sensor fusion-based vacant parking slot detection and tracking," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 21–36, 2014.
- [53] S. Valipour, M. Siam, E. Stroulia, and M. Jagersand, "Parking-stall vacancy indicator system, based on deep convolutional neural networks," *2016 IEEE 3rd World Forum Internet Things, WF-IoT 2016*, pp. 655–660, 2016.
- [54] X. Li, M. C. Chuah, and S. Bhattacharya, "UAV Assisted Smart Parking Solution," *Unmanned Aircr. Syst. 2017 Int. Conf.*, pp. 1006–1013, 2017.
- [55] C. Liao, "Using Truck GPS Data for Freight Performance Analysis in the Twin Cities Metro Area," 2014.
- [56] K. Haque, S. Mishra, R. Paleti, M. M. Golias, A. A. Sarker, and K. Pujat, "Truck Parking Utilization Analysis Using Truck GPS data," in *TRB 95th Annual Meeting Compendium of Papers*, 2016.
- [57] I. W. Ford Torrey and D. Murray, "Managing Critical Truck Parking Tech Memo \# 2 : Minnesota Case Study – Utilizing Truck GPS Data to Assess Parking Supply and Demand," Atlanta, GA, 2017.
- [58] B. Xu, O. Wolfson, J. Yang, L. Stenneth, P. S. Yu, and P. C. Nelson, "Real-time street parking availability estimation," *Proc. - IEEE Int. Conf. Mob. Data Manag.*, vol. 1, pp. 16–25, 2013.

- [59] J. Kopecky and J. Domingue, "ParkJam Crowdsourcing parking availability information with linked data (demo)," in *ESWC 2012 Satellite Events, LNCS 7540*, E. Simperl, B. Norton, D. Mladenic, E. Della Valle, I. Fundulaki, A. Passant, and R. Troncy, Eds. Springer, 2015, pp. 381–386.
- [60] W. Viriyasitavat, P. Sangaroonsilp, J. Sumritkij, and N. Tarananopas, "Mobile crowdsourcing platform for intelligent car park systems," *ICSEC 2015 - 19th Int. Comput. Sci. Eng. Conf. Hybrid Cloud Comput. A New Approach Big Data Era*, 2016.
- [61] S. Mathur, T. Jin, N. Kasturirangan, J. Chandrasekaran, W. Xue, M. Gruteser, and W. Trappe, "ParkNet: drive-by sensing of road-side parking statistics," *Proc. 8th Int. Conf. Mob. Syst. Appl. Serv. - MobiSys '10*, p. 123, 2010.
- [62] B. Hoh, T. Yan, D. Ganesan, K. Tracton, T. Iwuchukwu, and J. S. Lee, "TruCentive: A game-theoretic incentive platform for trustworthy mobile crowdsourcing parking services," *IEEE Conf. Intell. Transp. Syst. Proceedings, ITSC*, pp. 160–166, 2012.
- [63] Trucker Path, "<https://truckerpath.com/>." .
- [64] NATSO, "<http://www.natso.com/parkmytruck>." .
- [65] M. Caliskan, D. Graupner, and M. Mauve, "Decentralized discovery of free parking places," *Proc. 3rd Int. Work. Veh. ad hoc networks*, no. Vanet, pp. 30–39, 2006.
- [66] R. Lu, X. Lin, H. Zhu, and X. Shen, "SPARK: A new VANET-based smart parking scheme for large parking lots," *Proc. - IEEE INFOCOM*, pp. 1413–1421, 2009.
- [67] G. Y. Chang, J. P. Sheu, and C. Y. Chung, "Zooming: A zoom-based approach for parking space availability in VANET," *IEEE Veh. Technol. Conf.*, pp. 1–5, 2010.
- [68] Pilot/FlyingJ, "Pilot Flying J." [Online]. Available: <https://pilotflyingj.checkfront.com/reserve/>. [Accessed: 17-Apr-2017].
- [69] T. Rajabioun, B. Foster, and P. Ioannou, "Intelligent parking assist," *2013 21st Mediterr. Conf. Control Autom. MED 2013 - Conf. Proc.*, pp. 1156–1161, 2013.
- [70] Z. Schwartz and S. Hiemstra, "Improving the Accuracy of Hotel Reservations Forecasting: Curves Similarity Approach," *J. Travel Res.*, vol. 36, no. 1, pp. 3–14, 1997.