Eindhoven, August 2019

Forecasting the required tank container and trucking capacity for an intermodal logistic service provider

By Rijk van der Meulen

BSc. International Business Administration
Student identity number: 1231485

In partial fulfilment of the requirements for the degree of

Master of Science
in Operations Management and Logistics

Supervisors:

Dr. S.E.C. (Sarah) Gelper – Assistant Professor Eindhoven University of Technology
Prof. dr. T. (Tom) Van Woensel – Full Professor Eindhoven University of Technology
L. (Lex) Dogterom – Business Engineer H&S Foodtrans B.V

TU/E School of Industrial Engineering
Series Master Thesis Operations Management and Logistics

Subject headings: forecasting, advance demand information, Bayesian updating, hierarchical time-series, multiple linear regression, trucking capacity, tank container capacity, dynamic harmonic regression, artificial neural networks

Abstract

This research was motivated by two major operational challenges faced by intermodal logistics service providers: the repositioning of empty tank containers and the proactive planning of their drayage operations. To address these challenges, practitioners are in need of models that provide them with an accurate prediction of the number of loadings and deliveries and the corresponding capacity requirements in terms of trucks and tank containers. This study aimed to make a contribution to this problem area by investigating how forecasting can be used to predict the short-term (i.e. one to three weeks ahead) required tank container and trucking capacity. Since the required capacity is a derivative of the loadings and deliveries that need to be performed, this study began with forecasting the number of loadings and deliveries based on historical data. Subsequently, a Bayesian technique was developed to adjust the initial forecast based on advance demand information. The next step involved using the adjusted forecast to predict the required tank container and trucking capacity. With regards to the tank container capacity, a hierarchical top-down forecasting approach was employed that disaggregates the adjusted forecast for the number of loadings by type of tank container. Regarding trucking capacity, a multiple linear regression model was designed that utilizes the forecasted loadings and deliveries and other explanatory variables to predict the required trucking capacity. Results on a one-month test case for two planning regions showed that the Bayesian technique increased the accuracy of the initial forecasted loadings by 65% and 28%. Furthermore, the mean absolute error expressed in tank containers depended on the level of aggregation (i.e. the type of tank container) and varied between 0.57 and 2.72. Finally, regarding the required trucking capacity, the predictive accuracy in terms of mean absolute error during working days was shown to be 20.39 hours (i.e. approximately 1.62 trucks). The proposed forecasting methodology for predicting trucking capacity significantly outperformed the forecasting method that is currently employed at H&S: the difference in mean absolute error during working days turned out to be 41%. Taken together, the results of this study suggest that the proposed forecasting methodology accurately predicts the number of loadings and deliveries and the corresponding tank container and trucking capacity. As such, it is believed that the proposed forecast methodology presents a powerful tool that can assist planners at H&S Foodtrans in (1) making more efficient tank container repositioning decisions and (2) planning the drayage operations in a more proactive fashion.

Key words: forecasting, advance demand information, Bayesian updating, hierarchical time-series, multiple linear regression, trucking capacity, tank container capacity, dynamic harmonic regression, artificial neural network

Executive summary

Two key operational challenges faced by intermodal logistics service providers include the repositioning of empty tank containers and the proactive planning of drayage operations. These challenges can, partly, be addressed by accurately forecasting the short-term required tank container and trucking capacity. As such, the formal research question was defined as:

How can forecasting be used to accurately predict the short-term required tank container and trucking capacity for an intermodal logistics service provider?

Since the required capacity is a derivative of the loadings and deliveries that need to be performed, this study began by investigating how these loadings and deliveries can be forecasted based on historical data (Chapter 2). The data of 2017 and 2018 was used to estimate the parameters of the forecasting models and the data of the first three months of 2019 was used to evaluate the accuracy of the models. Subsequently, it was studied how this initial forecast (based on historical data) could be improved given the available advance demand information (Chapter 3). In other words, how can the information regarding the loadings and deliveries for a future time period, that are already known at present, be used to enhance the initial forecast? Finally, the last two steps of this thesis involved examining how the adjusted forecast for the number of loadings and deliveries could be employed to predict the required tank container capacity (Chapter 4) and trucking capacity (Chapter 5). Based on the analysis as described above, two forecasting methodologies were proposed to predict the short-term required tank container and trucking capacity. These methodologies are discussed below and visualized in Figure 1.

Proposed forecasting methodology for predicting tank container capacity

One of the fundamental building blocks of this proposed methodology is the idea that the total loadings in a certain region can be disaggregated by type of tank container. In other words, each loading at a certain region at a given day exactly represents the need of one tank container. Exploiting this observation, the following hierarchical forecasting methodology was proposed:

Step 1: Generate the initial 21 day ahead forecast for the number of loadings

In the first step, the initial (three weeks ahead) forecast is generated for the daily number of loadings in a certain region. In theory, all models developed in Chapter 2 can be used to perform this task. However, the simple mean method, dynamic harmonic regression, and artificial neural networks were the models that showed the highest predictive accuracy in this step and it is therefore recommended to use one of these models. To go one step further, for ease of implementation it is recommended to use the simple mean method in this step. This model is easy to understand and can be implemented for all planning regions with minimal effort. In case the ease of implementation is not an issue, one can also consider implementing the artificial neural network for each region. Although the difference compared to the simple mean method was often minimal, the neural networks did, on average, present the highest predictive accuracy.

Step 2: Adjust the initial forecast by applying the Bayesian algorithm

In the second step, the initial forecast as obtained in step 1 is adjusted to incorporate the advance demand information. The practitioner is left with two options to utilize the advance demand information. First, one can use the Bayesian algorithm which explores and analyses the possibility of using the expected number of orders for a future period as the variable to be estimated. This technique was shown to be very accurate, but it has also been subjected to criticism that it might be too complex for practitioners to implement. Hence, as an alternative, a combined forecast model was developed that combines the initial forecast with an inflator

algorithm. Although this model proved to be fairly accurate, the results were not as impressive as the Bayesian technique. Therefore, the practitioner needs to be mindful of the trade-off between ease of implementation and predictive accuracy. Moreover, another reason for preferring the Bayesian technique over the combined forecast might be that the Bayesian algorithm produces an entire forecast distribution, whereas the combined forecast only generated point forecasts. As such, the Bayesian technique allows for the computation of prediction intervals, whereas this is not the case for the combined forecast.

Step 3: Generate the preliminary forecast of the required tank container capacity
In the third step, a preliminary forecast is generated for the required tank container capacity based on historical data. Put differently, a three weeks ahead forecast is generated for the daily number of tank containers needed of each type at a certain region. This forecast can be obtained either by the simple mean method or an artificial neural network (as developed in Chapter 4). When choosing between the two, once again the trade-off between ease of implementation and accuracy needs to be considered. The simple mean method is easier to understand and implement, but the neural networks again showed a higher accuracy. It should be taken into account that for the sake of this research, the neural networks were used to generate the preliminary forecast.

Step 4: Compute the final forecast for the required tank container capacity
These preliminary forecasts derived in step 3 are not used directly and they are not coherent (i.e. they do not add up correctly). Instead, the preliminary forecasts are only used to calculate the forecasted proportions of each tank container type. Finally, the forecasted loadings (derived from step 2) can be disaggregated down the hierarchy based on these forecasted proportions.

Proposed forecasting methodology for predicting trucking capacity

Step 1: Generate the initial 7 day ahead forecast for the number of loadings
The actions to be performed, as well as the consideration as to which model to use, are similar to step 1 of the proposed model for predicting tank container capacity. The only exception is that here, a forecast is generated for the sub-daily number of loadings and deliveries for one week ahead. In other words, for each day, the number of loadings and deliveries are forecasted from midnight to noon (AM) and noon to midnight (PM) for one week ahead in total.

Step 2: Adjust the initial forecast by applying the Bayesian algorithm
In this step the exact same procedure is followed as in step 2 of the proposed model for predicting tank container capacity.

Step 3: Run the multiple linear regression model to generate the forecast the required trucking capacity
In the final step, the forecasted number of loadings and deliveries (derived in step 2) are used as primary explanatory variables in a multiple linear regression model (developed in Chapter 5). Together with a number of other explanatory variables, this model can be used to predict the sub-daily required trucking capacity in hours for one week ahead.

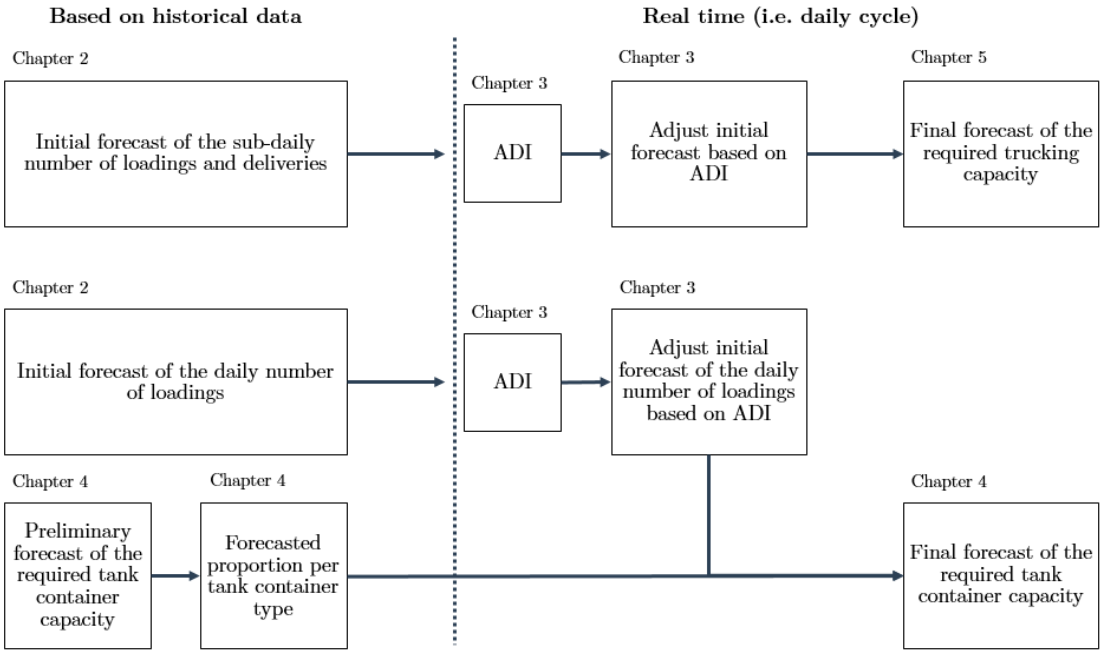


Figure 1 Proposed forecasting methodology

Accuracy during a test case in Rotterdam and Belgium / North France

Once the reader has grasped the steps involved in the proposed forecasting methodology, it is time to present the key results as to how the proposed methodology performed during a one-month test case. Particularly, the proposed methodology for the tank container capacity and trucking capacity were tested during a one-month period for Rotterdam and the Belgium and North France region (BFN), respectively. The main results are outlined below:

First, the Bayesian technique significantly improved the accuracy of the initial forecast of the number of loadings and deliveries: the accuracy in terms of mean absolute error improved by 65% and 28% for the BFN and Rotterdam region, respectively.

Second, the Bayesian technique particularly improved the initial forecasts for periods that lay nearer in the future. This is wat could be expected, since more advance demand information is known for these periods.

Third, although the performance of the combined forecast is not as accurate as the Bayesian technique, it still performed reasonably well and can therefore be seen as a serious alternative to the Bayesian technique if the latter is considered to be too complex by the practitioner. The adequate performance of the combined forecast also adds to the growing corpus of studies showing that combining multiple forecasts leads to increased forecast accuracy.

Fourth, the required tank container capacity can be forecasted with a fairly high accuracy by the proposed forecasting methodology. The accuracy in terms of mean absolute error expressed in number of tank containers depends on the level of aggregation (i.e. type of tank container) and varied between 0.57 and 2.72.

Fifth, although the proposed methodology for predicting the tank container capacity outperforms the preliminary forecast (based exclusively on historical data), the difference in predictive accuracy was fairly small.

Sixth, the required trucking capacity can also be forecasted with a reasonably high accuracy by the proposed forecasting methodology. The accuracy in terms of mean absolute error per period (i.e. half a day) in the BFN region during working days was shown to be 20.39 hours. Expressed in trucks (based on their average productivity), this translates to approximately 1.62 trucks.

Seventh, the proposed forecasting methodology for predicting trucking capacity performs significantly better than the current forecasting method: the difference in mean absolute error during working days for the Belgium and North France region turned out to be 41%.

Overall conclusion

All things considered, the findings of this research indicate that the proposed forecasting methodology accurately predicts the number of loadings and deliveries and the corresponding tank container and trucking capacity. As such, it is believed that the devised forecast methodology presents a powerful tool that can assist planners at H&S Foodtrans in (1) making more efficient tank container repositioning decisions and (2) planning the drayage operations in a more proactive fashion.

Limitations and practical suggestions

To conclude this executive summary, a number of limitations regarding this study are considered and it is discussed how these limitations might relate to fruitful directions for future research and fine-tuning of the proposed forecasting methodology. Since the findings of this study are currently being implemented in consultation with a data science consultant (CQM), these practical suggestions might be of particular importance.

The first limitation is concerned with the observation that the developed models are trained (and evaluated) based on the realized demand instead of direct customer demand. To address this limitation, an attempt could be made to gather the direct customer demand data and re-estimate the models based on this actual demand. More details on this topic are provided in Chapter 7.

Another limitation lies in the fact that the (historical) trucking capacity was estimated based on both actual data as well as theoretical assumptions. Nevertheless, despite the theoretical assumptions it is expected that the estimation of the (historical) trucking capacity is fairly robust. Still, if H&S aims for an even more robust estimation of their (historical) trucking capacity, they might want to consider recording the actual data of all trucking actions.

A final limitation is concerned with the relatively small size of the test case. Given this small sample size, caution must be exercised in dealing with absolute errors in isolation. Ideally, the promising findings of this study, regarding the predictive accuracy of the proposed methodology, should be replicated in a study with a larger test set on multiple planning regions. Fortunately, results regarding this topic are soon to be expected since the proposed forecasting methodology is currently being implemented in consultation with a data science consultant called CQM. Part of this project involves testing the accuracy of the proposed forecasting methodology on a larger test set including multiple planning regions.

In addition to the recommendations discussed above, a final interesting direction for further research might lie in the topic of collaborative forecasting. Instead of using only local information, there might be a great deal of untapped potential in using the forecasts of supply chain partners to enhance one's own predictive model. What makes this area of research particularly interesting for H&S Foodtrans, is that a small number of clients is responsible for a large part of the orders. Specifically, over the last six months, the largest 10% of the clients accounted for 70% of the orders. The prospect of harnessing this information and collaborating with the largest clients seems to be worthwhile exploring.

Preface

Here it is. Before you lies my thesis that has been written to fulfill the graduation requirements for the degree of Master of Science in Operations Management & Logistics at the Eindhoven University of Technology. Moreover, it also marks the end of a vibrant and enlightening five-year period of my life. This thesis, and my life as a student in general, would not have been the same without the inspiration and support of a number of wonderful individuals: my thanks and appreciation to all of you for contributing to this journey.

First, I want to express my deepest gratitude to Sarah Gelper, my mentor and first supervisor from the TU/e. Throughout the last two years, your continuous encouragement and sincere support enabled me to tackle (initial) challenges and excel in my studies. The trust you showed in my knowledge and capabilities made me feel confident to deliver a valuable and successful research project. I honestly could not have wished for a better mentor. I also want to thank Professor Tom van Woensel for his involvement in this research. I believe that your expertise and feedback near the end of the research were valuable in enhancing the quality of this work.

In addition, I gratefully acknowledge the contribution of several individuals within H&S Group. To Lex Dogterom, I really enjoyed having you as my company supervisor. In my opinion, you have struck the perfect balance between giving me a lot of freedom and responsibility, while simultaneously being actively engaged in the project. I am also glad that we discovered our mutual interest in reading (philosophical) books. This mutual interest resulted in interesting discussions during occasional breaks which I enjoyed a lot. To Siebe Dijkstra, thank you for guiding me through the data warehouse and helping me find my way in that massive architecture of data. I would also like to thank Marcel van Hulst for his supervision, especially during the early stages of the project.

Finally and foremost, I want to express my utmost gratitude to my family and friends. Throughout the years I gained numerous friendships that I will treasure the rest of my life. Thank you for making these years such a fun and inspiring time. To my parents, Linne and Julia, thank you for your continuous and unparalleled love, help, and support. You have always encouraged me to seek my own destiny and realize my ambitions, be it in sport, my personal life, or my studies. This journey would not have been possible if not for you.

Rijk van der Meulen

Eindhoven, August 2019

Table of Contents

List of Figures

List of Tables

List of Abbreviations

LSP	Logistics Service Provider
MMP	Multi Material Planning (i.e. tank container planning)
TCP	Truck Capacity Planning (i.e. planning of trucking units)
BFN	Belgium and North France planning region
GBN	Great Britain North planning region
AICc	Akaike Information Criterion
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
MAPE	Mean Absolute Percentage Error
ADI	Advance Demand Information
ANN	Artificial Neural Network

1 Introduction

Tell us what the future holds, so we may know that you are gods (Isaiah 41:23)

The practice of forecasting has enchanted people for thousands of years, sometimes seen as a felonious activity and sometimes being considered as an indication of divine inspiration. In ancient Babylon, about 600 BCE, future events were forecasted based on the distribution of maggots in a rotten sheep's liver (Hyndman & Athanasopoulos, 2018). 300 years later, people travelled all the way to Greece to consult the Oracle who predicted the future while intoxicated by ethylene vapours. Forecasters living in 357 CE had a tougher time with emperor Constantine issuing an injunction in which he forbade people to consult soothsayers, mathematicians and forecasters. More recently, in 1736 a comparable prohibition occurred in Great Britain. At that time, it became an offence to defraud by charging money for predictions. The punishment was three months of imprisonment with hard labour (Hyndman & Athanasopoulos, 2018). This thesis is the result of six (rather than three) months of hard labour, investigating how forecasting can be used to predict the short-term capacity requirements for an intermodal logistics service provider. This introductory chapter starts by introducing the main parties involved in this research. Thereupon, a problem description is given that is followed by an overall problem statement. Finally, the key deliverables are discussed, and the structure of this thesis is outlined.

1.1 Parties involved in the research

1.1.1 H&S Group B.V.

As the result of the interplay between producers and customers and the geographical distances that often separate them, producers of goods require transportation services to move raw materials or (intermediate) products in order to meet customer demands (Crainic & Kim, 2007). These transportation services are often carried out by logistics service providers (LSPs). Moreover, as a consequence of trade globalization, the conventional road mode is no longer an all-time feasible solution, necessitating other means of transportation such as intermodal containerized transportation. Intermodal transportation can be defined as the combination of at least two modes of transport in a single transport chain, without a change of container for the goods, with most of the route travelled by rail or ferry and with the shortest possible initial and final journeys by road transport (Macharis & Bontekoning, 2004). H&S is such an LSP in the intermodal bulk transport of liquid foodstuff. In particular, H&S Group B.V. is an encapsulating business entity that governs all aspects of providing logistics services in the (intermodal) liquid food industry and consists of five business units: Coldstores, Logistic Services, Cleaning, Transport and Foodtrans. Although these business units are organized under the same industrial group, they operate independently of one another. The transportation business unit can be subdivided into a part that is concerned only with the modal transport (H&S transport B.V.) and the part that is responsible for the intermodal transportation of liquid bulk (H&S Foodtrans B.V.). This business unit is mostly asset-based and fulfils the role of transporting liquid foodstuff throughout a pan-European network. This research was conducted specifically for H&S Foodtrans B.V. (in the remainder of this thesis referred to as H&S Foodtrans). In the section that follows, more detailed information is provided regarding this business unit.

1.1.1.1 H&S Foodtrans

As explained in the previous section, H&S Foodtrans is in charge of planning the intermodal transportation of liquid bulk. To arrange the intermodal transport for their clients, H&S Foodtrans needs both trucking units and tank containers. Trucking units are required for pre- and end-haulage activities that arrange the transport of containers by road between a container terminal, customer locations, and cleaning locations.

Furthermore, tank containers are needed to store the commodity during transportation. H&S Foodtrans consists of seven departments: Sales, Pricing, Account Management, Planning, Business Engineering, Quality, and IT. The role and responsibility of each department is briefly summarized below. The sales department is responsible for attracting new clients and re-establishing relationships with existing clients. Once the sales department has identified a new potential customer, the pricing department is responsible for determining the right tariff of the order. Thereupon, when a new client has signed a contract with H&S Foodtrans, they are assigned to an account manager who maintains the relationship with the clients. In particular, the account manager is the person whom the client should contact to place an order and keep informed about the progress of this order. Once the account manager has confirmed the order and it is also accepted by the technical system, it is transferred to the planning department. Next, the planning department plans and schedules the orders in such way that goods are delivered at the right time and place. More specifically, the planning department assigns a tank container and trucking unit to an order. This planning process is performed in a sequential manner. First a tank container is assigned to an order, and only when this has been completed, the trucking unit is scheduled. The part of the planning department that assigns the tank container is called the Multi Material Planning (MMP) and the part that plans the trucking unit is referred to as the Truck Capacity Planning (TCP). Finally, Business Engineering, Quality, and IT are all departments that support the core businesses of H&S Foodtrans.

1.1.2 Data2Move

Data2Move is the leading research community for logistics, supply chain management and big data. The initiative is part of the European Supply Chain Forum and was established to help companies in the industry exploit data to optimize and enhance their business processes. In particular, academics and students from the TU/e collaborate with prominent industry partners to deliver results that are relevant from both a practical and scientific point of view. Carrying out this research in the wider Data2Move context has proven to be very valuable. In particular, receiving feedback from both leading professors in the academia as well as the industry partners participating in Data2Move, turned out to be helpful in conducting this research.

1.1.3 CQM and Den Hartogh

At the time this study was completed, the management team of H&S decided they wanted to reap the benefits of this research by implementing the proposed forecasting methodology and integrate it in their planning software. To this end, a collaborative project between H&S, Den Hartogh and CQM was started to develop and integrate a forecasting methodology that works for both H&S and Den Hartogh. Den Hartogh is an LSP of bulk liquids and gasses for the chemical industry. H&S and Den Hartogh collaborate on many topics and even use the same tailor-made planning software (i.e. Transfusion) that has been designed specifically for them. CQM is a data science consultancy firm specialized in quantitative modelling. They have assisted both H&S as well as Den Hartogh in optimizing their operational processes throughout the years. More details regarding this implementation project is provided in Chapter 6.

1.2 Problem description

Given the extremely low profit margins in the intermodal transport industry, LSPs such as H&S Foodtrans are urged to work at low cost, while still maintaining high quality. Clearly, in this context, efficient and effective transportation planning is needed. Two main challenges in this transportation planning process include the proactive planning of drayage operations and efficient empty tank container repositioning.

1.2.1 Proactive planning of drayage operations

The first challenge is concerned with the efficient and pro-active planning of drayage operation. In other words, the planning of the pre- and end-haulage activities that arrange the transport of containers by road

between a container terminal, customer locations and cleaning locations. To arrange the drayage operation, H&S uses both own trucks as well as charters. Efficient planning of the drayage operation is of crucial importance. The reason for this is that, although the pre- and end-haulage activities only account for a small amount of total distance, drayage operations have been estimated to account for 25% to 40% of total intermodal container transportation costs (Funke & Kopfer, 2016). The current planning of the drayage operation within H&S Foodtrans occurs in a relatively ad hoc fashion. Due to their busy schedules, planners are often too carried away by the everyday humdrum to be engaged with estimating the number of trucking hours needed for tomorrow, let alone longer in the future. As a result, charters are often booked at the very last moment. This leads to a number of problems. First, sometimes there are no charters available anymore at the last moment. As a result, certain orders might have to be cancelled or delayed, reducing the on-time performance of H&S towards their clients. Second, according to the planners at H&S, the quality in terms of the drivers' qualifications and capabilities is lower for charters that are still available on such a short notice (i.e. the high performing charters are fully booked at an earlier stage). Third, it might sometimes also occur that a premium is asked by charters when they are booked on a short notice. Another result of the current myopic planning process is that almost no effort is made to proactively smooth the workload throughout the day and week. A smoothed workload is important as it leads to higher utilization of own trucks and less need for charters. Moreover, if the workload is divided more equally throughout the day, less trucks are needed since planners can make more efficient combinations. In sum, operational costs can be reduced by planning the drayage operations more proactively. An important prerequisite for moving to a more proactive planning, is insight in the expected number of trucking hours that are required to fulfil customers demand. In other words, research is needed as to how the number of loadings and deliveries and corresponding required trucking capacity can be accurately forecasted.

1.2.2 Repositioning of empty tank containers

The second challenge is concerned with repositioning decisions that need to be made for empty tank containers. This challenge arises due to the imbalance between product supply and demand and hence an imbalance in the container flow across different regions. This imbalance has become a major component in the strategic planning activities of LSPs because of its enormous potential effect on profitability (Fite, Don Taylor, Usher, English, Roberts, 2002). Suppose, for instance, an LSP has 80 loads per week inbound to the UK, but only 50 outbound loads. That leaves 30 containers per week in the UK with no loads to move. In other words, certain areas develop a surplus of containers, while others have a deficit (Braekers, Janssens & Caris, 2011). The imbalance in the container flow across regions, leads to empty container movements that do not generate revenue but do incur considerable operating costs. For land transportation of containers, it is estimated that 40% up to 50% of all container movements are empty container movements (Konings & Thijs, 2001; Branch, 2006). Efficient empty container repositioning, in order to be able to fulfil future demand while minimizing costs, is therefore one of the most pressing challenges for LSPs in containerized transport. Taking the imbalance in the container flow across regions as a given, the questions arises what should be done with empty containers in a certain region in order to minimize empty container movements, while still ensuring that demand can be met. To make informed repositioning decisions, one needs information about the number of loadings that are expected to occur in a certain region at a given time. Currently, repositioning decisions are made by two very experienced and capable planners at the MMP department at H&S Foodtrans. Despite their competence, solely relying on the experience of these planners has two major drawbacks. First, they have to store a lot of information in their heads and regardless of their competence, they might sometimes still be biased. Second, relying exclusively on the experience of two planners, poses a high risk for a large firm such as H&S. If one (let alone two) of these planners drop out for one reason or another, there is no alternative system on which repositioning decisions might be based. For these reasons, research is needed as to how the required tank container capacity can be accurately forecasted.

1.3 Problem statement

The previous section explained that two main challenges faced by H&S Foodtrans are (1) making efficient repositioning decisions and (2) planning the drayage operations in a more proactive fashion. Furthermore, it was argued that these challenges could be addressed by having an accurate forecast of the required tank container and trucking capacity in a certain region. For this reason, the main research question that this thesis aims to answer is:

How can forecasting be used to accurately predict the short-term required tank container and trucking capacity for an intermodal logistics service provider?

Based on this problem statement, a number of research questions were formulated to further narrow the field of research and guide this thesis:

RQ1: How can the number of loadings and deliveries in a given region be accurately forecasted from historical data?

The required tank container and trucking capacity are a derivative of the number of loadings and deliveries that need to be carried out in a certain region (i.e. customer demand). Hence, the first research question studies how the number of loadings and deliveries in a certain region can be forecasted based on historical data.

RQ2: How can advance demand information be utilized to enhance the initial forecast of the loadings and deliveries?

If one would predict the number of loadings and deliveries based exclusively on historical data, one would miss out on important information. Particularly, in the intermodal bulk transport industry, advance order information might be available. Advance demand information constitutes future orders (i.e. loadings and deliveries) that are already known at present. Demand for intermodal freight transportation is usually known at least some time in advance, since equipment needs to be scheduled and arranged. Therefore, the second research question investigates how the initial forecast, from historical data, can be enhanced by exploiting the advance demand information.

RQ3: How can the forecasted loadings be converted to the required tank container capacity?

The output of RQ2 is a forecasting model that uses the advance demand information and predicts the expected loadings and deliveries in a certain region. Subsequently, RQ3 studies how the forecasted number of loadings can be translated to the required tank container capacity per region.

RQ4: How can the forecasted loadings and deliveries be used to predict the required trucking capacity?

In roughly the same fashion as RQ3, the fourth research question investigates how the forecasted loadings and deliveries can be employed to predict the required trucking capacity per region.

1.4 Project scope and deliverables

Since H&S Foodtrans is engaged in intermodal bulk transport throughout a pan-European network, various planning regions can be distinguished. Consequently, a separate forecast of the required tank container and trucking capacity needs to be generated for each planning region. In order to demarcate the scope of this thesis, however, it was decided to focus on a few critical regions. In particular, in consultation with the TCP department, it was decided to focus the forecast of the required trucking capacity on the North of Great

Britain and Belgium / North France. These focus regions are from now on referred to as the GBN and the BFN regions, respectively. A specific deliverable for these regions is a forecasting model that predicts the sub-daily required trucking capacity for one week ahead. In other words: for each day, the trucking capacity is forecasted from midnight to noon (AM) and noon to midnight (PM) for one week ahead in total. Similarly, after consultation with the MMP department, it was decided to focus the forecast of the required tank container capacity on the Rotterdam planning region. The deliverable for this region is a daily forecast that predicts the tank container capacity three weeks ahead. The time-horizon is set at three weeks, as this time frame provides sufficient time to make repositioning decisions. By focusing the research on these critical regions, this thesis endeavours to produce a proof of concept that: (1) shows that capacity requirements can be forecasted over the short term and (2) propose a forecasting methodology that can be used to derive these forecasts. If proven successful, the proposed methodology can be extended to all other planning regions throughout the transportation network of H&S. After this research was completed, H&S indeed started with an implementation trajectory to use the findings of this research, applying the proposed forecasting methodology to all planning regions. For more details regarding this implementation project, the reader is referred to Chapter 6.

1.5 Structure of the research

The overall structure of this study takes the form of seven chapters. Chapter 1 introduced the research topic and problem statement, acquainting the reader with the main concepts and imperatives behind this thesis. Based on the problem statement, four research questions were presented that guide the structure of this thesis. In particular, in Chapter 2 to 5, the four research questions are addressed in chronological order.

Chapter 2 begins by laying out the characteristics of the data. Particularly, the various seasonal components in the data are identified and analysed. Second, the literature is consulted to identify forecasting models that can account for data with these characteristics. Third, in order to evaluate and compare forecasts, proper accuracy measures are identified. Fourth, the relevant models, identified in the literature, are developed in the context of this research and their predictive accuracy is evaluated.

Subsequently, in Chapter 3 an answer is provided as to how advance demand information can be utilized to enhance the initial forecast derived in Chapter 2. After acquainting the reader with the topic of advance demand information, Chapter 3 proceeds with analysing the specific characteristics of the advance demand information in the context of H&S Foodtrans. Consecutively, a Bayesian algorithm is developed that adjusts the initial forecast by incorporating information about future orders, that are already known at present. Finally, since Bayesian models are criticized on the grounds that their complexity requires a “specialist” to understand and implement them, another (simpler) model is developed that may serve as an alternative to the Bayesian model.

Thereupon, Chapter 4 is concerned with studying how the adjusted forecast (derived in Chapter 3) of the number of loadings and deliveries can be converted to the required tank container capacity. Specifically, Chapter 4 first describes the heterogeneous tank container fleet at H&S Foodtrans. Particular attention is devoted to examining the different characteristics of the tanks and a classification for the various type of tank containers is suggested. Second, the concept of hierarchical time series is introduced, and it is argued that the time series of the forecasted loadings can be disaggregated by type of tank container. Third, a number of forecasting methods for hierarchical time series are discussed and it is concluded that the top-down approach based on forecasted proportions seems most applicable to the context of this research. Fourth, this top-down method based on forecasted proportions is employed to generate a final forecast of the tank container capacity. Finally, the accuracy of the proposed forecasting model is tested and compared with a benchmark model.

In a roughly similar fashion, Chapter 5 studies how the adjusted forecast of the number of loadings and deliveries can be used to predict the required trucking capacity. First, some deeper insight is provided as to which specific actions in H&S' operation require trucking capacity. Second, the historical trucking capacity per region is estimated. This estimation is based on actual data as well as theoretical assumptions. Third, a multiple linear regression model is developed that generates forecasts for the required trucking capacity. Lastly, the accuracy of the proposed regression model is assessed and compared with a benchmark model.

Chapter 6 circles back to the two operational challenges faced by H&S Foodtrans (and intermodal LSPs in general). Specifically, this chapter particularizes how the devised forecasting methodology might help in addressing the operational challenges and quantifies some of the potential benefits of using the forecasting model. Moreover, attention is devoted to the implementation of this research.

Finally, Chapter 7 draws upon the entire thesis, tying up the conclusions of the previous chapters. Particularly, this chapter includes the synthetization of the key insights and a discussion of the implication of the findings to future research into this area.

2 Forecasting the loadings and deliveries

The goal of this chapter is to study how the number of loadings and deliveries in a given region can be forecasted based on historical data. In particular, six forecasting models that predict the number of loadings and deliveries are obtained. The first two models predict the daily number of loadings (model 1) and deliveries (model 2) in Rotterdam for three weeks ahead. At a later stage in the research, these models are used to determine the required tank container capacity. The second two models predict the sub-daily number of loadings (model 3) and deliveries (model 4) in the North of Great Britain (the GBN region) one week ahead. In other words, for each day the number of loadings and deliveries are forecasted from midnight to noon (AM), and noon to midnight (PM) for one week ahead in total. Finally, the last two models predict the sub-daily number of loadings (model 5) and deliveries (model 6) in Belgium and North France (the BFN region) one week ahead. At a later stage in the research the last four models were used as input to predict the required trucking capacity. The remainder of this chapter is structured as follows. First, the characteristics of the data are discussed. Specifically, the various seasonal periods are identified and analysed. Second, the literature is consulted to identify forecasting models that can account for data with these characteristics. Third, the relevant models identified in the literature are applied and implemented in the context of this research. Fourth, the various forecasting models are evaluated with respect to their predictive accuracy. Finally, the findings are synthesized, and an overall conclusion is drawn.

2.1 Characteristics of the data

In this section, the characteristics of the data are discussed. Since three different planning regions are examined in this research, this section specifically focusses on the characteristics of the data in Rotterdam, GBN and BFN.

2.1.1 Imbalance in loadings and deliveries

In the introductory chapter of this research, it was explained that there is often an imbalance between product supply and demand and hence an imbalance in the container flow across different regions. Rotterdam and GBN can be taken as perfect examples for this imbalance as in these regions there is an enormous surplus of loadings and deliveries, respectively. The number of loadings and deliveries in the BNF region is slightly more balanced, but this region too experiences a significant surplus of loadings relative to deliveries. The balance of the loadings and deliveries per region is visualized in Figure 2.1, 2.2, and 2.3. Another characteristic that might have immediately drawn the readers' attention in these figures, is the fact that the loadings and deliveries during weekdays differ substantially from the those during the weekends. This is just one of the various different seasonal patterns that might be present in the data. Therefore, these multiple seasonal patterns are examined in greater detail in the next section.

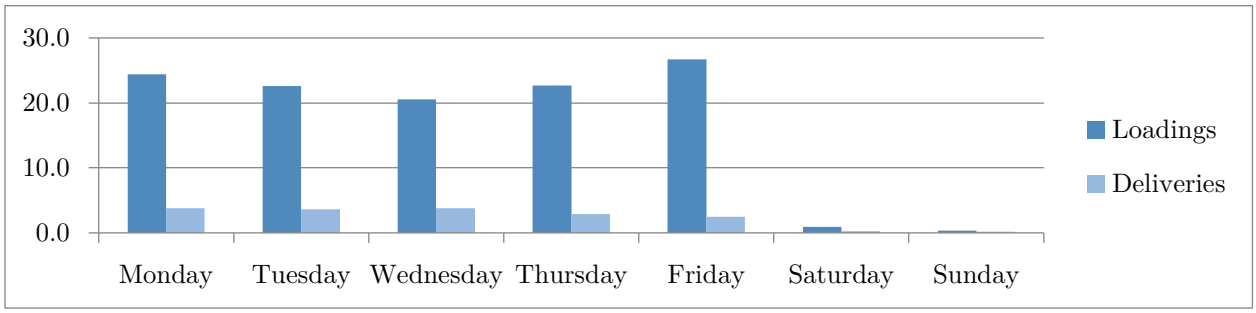


Figure 2.1 Average daily loadings and deliveries Rotterdam

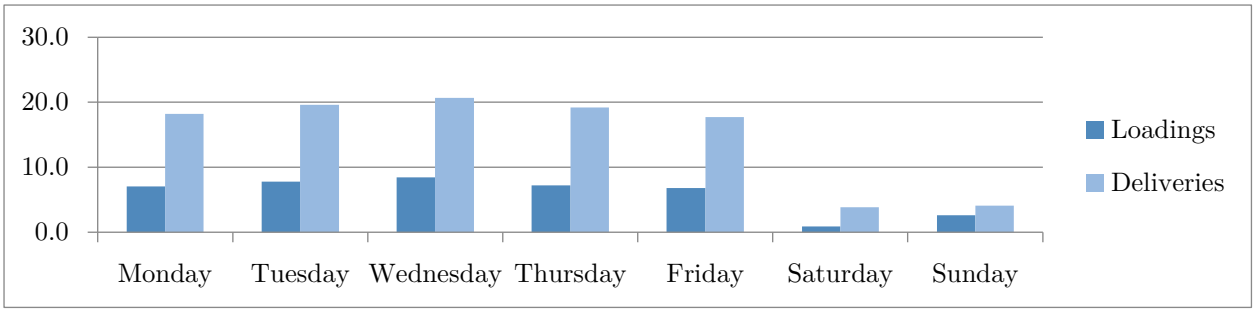


Figure 2.2 Average daily loadings and deliveries GBN

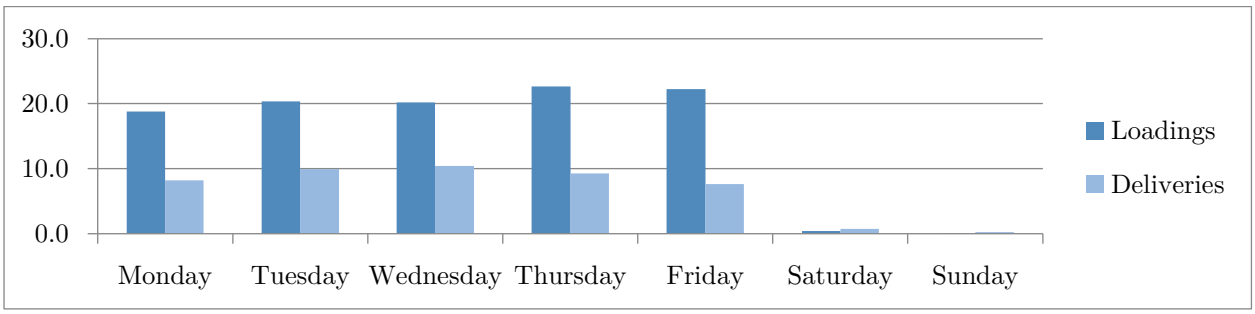


Figure 2.3 Average daily loadings and deliveries BFN

2.1.2 Seasonality

Since this research is concerned with deriving daily, and even sub-daily forecasts, the data is likely to exhibit multiple different seasonal patterns. It has been explained that for the GBN and BFN regions this research aims to derive a sub-daily forecast for one week ahead. More specifically, for each day, the number of loadings and deliveries is forecasted from midnight to noon (AM) and noon to midnight (PM), for one week ahead in total. Hence, for these two regions, the following seasonal periods can be distinguished: part of the day, day of the week, month of the year, and annual seasonality. Regarding the Rotterdam region, this thesis aims to derive a daily forecast that predicts container capacity three weeks ahead. Hence, for this region the potential seasonal periods are: day of the week, month of the year, and annual seasonality. Since the data exhibits a variety of seasonal patterns, it might be helpful to split the time series into various components, each representing an underlying pattern category. In order to extract the various components, a well-known technique is used: time series decomposition. Generally, a time series can be seen as comprising three components: a trend component, a seasonal component, and a remainder component (Hyndman & Athanasopoulos, 2018). In the context of this research, however, the data is likely to exhibit multiple seasonal components, rather than just one. A robust and successful method for general time series decomposition is the Seasonal and Trend decomposition using Loess (STL) developed by Cleveland, Cleveland, McRae & Terpenning (1990). This method can be extended to deal with data that comprises multiple seasonal components. Specifically, the STL with multiple seasonal components returns the various seasonal components, as well as a trend and remainder component. Figure 2.4 to 2.9 visualize the various components of all the time series considered in this research. To meticulously interpret these graphs, one should be mindful of the vertical scales. A general conclusion that seems to hold for all six time series, is that the trend component is fairly weak, since their ranges are rather narrow compared to the other components. Although still rather weak, the trend component seems to be slightly stronger for the Rotterdam and GBN deliveries series. Furthermore, considering the Rotterdam region, it can be observed that the components for the loadings time series and the deliveries time series are relatively similar. Moreover, it might be noticed that the weekly seasonality is the strongest, followed by the annual seasonality and the monthly seasonality. For the GBN region too, it may be noted that the weekly seasonality is strongest, for both the loadings as well as the deliveries time series. For a large part, this day of the week seasonality is due to the fact that the number

of loadings and deliveries are very low during weekend days. Regarding the variability within working days, it seems that every region has its own pattern, mainly determined by the ordering process of big clients within that particular region. Interestingly, the daily seasonality seems to be much stronger for the deliveries as compared to the loadings time series. This indicates that the deliveries in the GBN region depend to a large extend on the part of the day, whereas this pattern is much less apparent for loadings. Finally, the annual seasonality seems to be moderately important for both time series. Note that the monthly seasonality is not included in the decomposition since preliminary analysis showed that this component is very weak for both series. Lastly, for the BFN region, a relatively similar pattern seems to emerge as in the GBN region. Particularly, the weekly seasonality is the strongest component for both the loadings and the deliveries series. Moreover, the daily seasonality seems to be much more apparent in the deliveries than in the loadings time series. Finally, the annual component seems to be moderately important for both series.

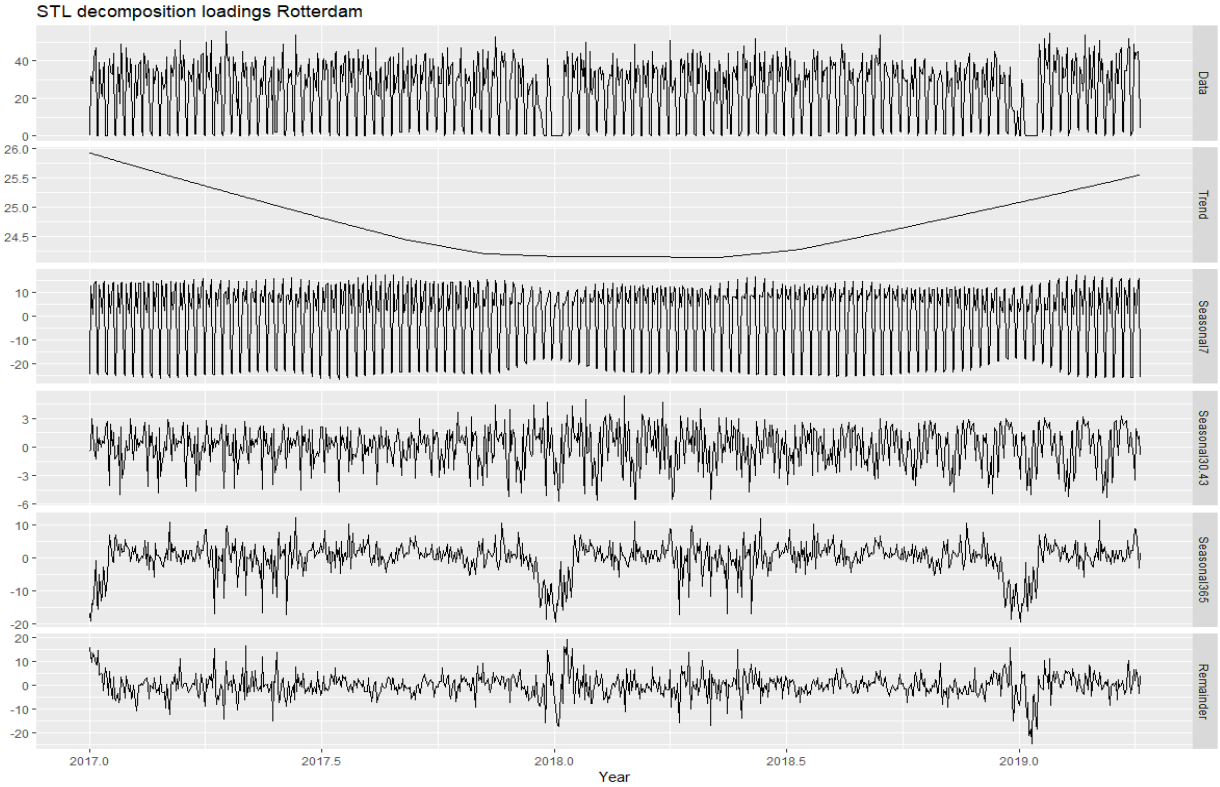


Figure 2.4 Various (seasonal) components loadings Rotterdam series

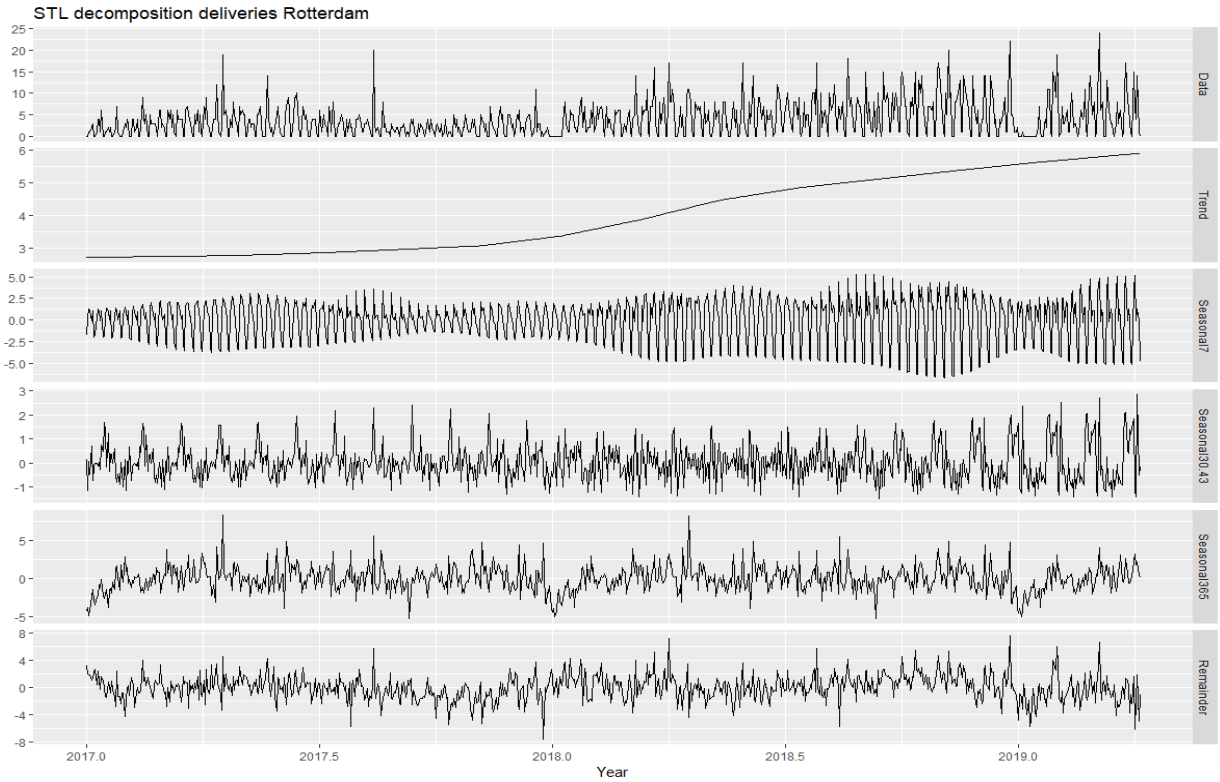


Figure 2.5 Various (seasonal) components deliveries Rotterdam series

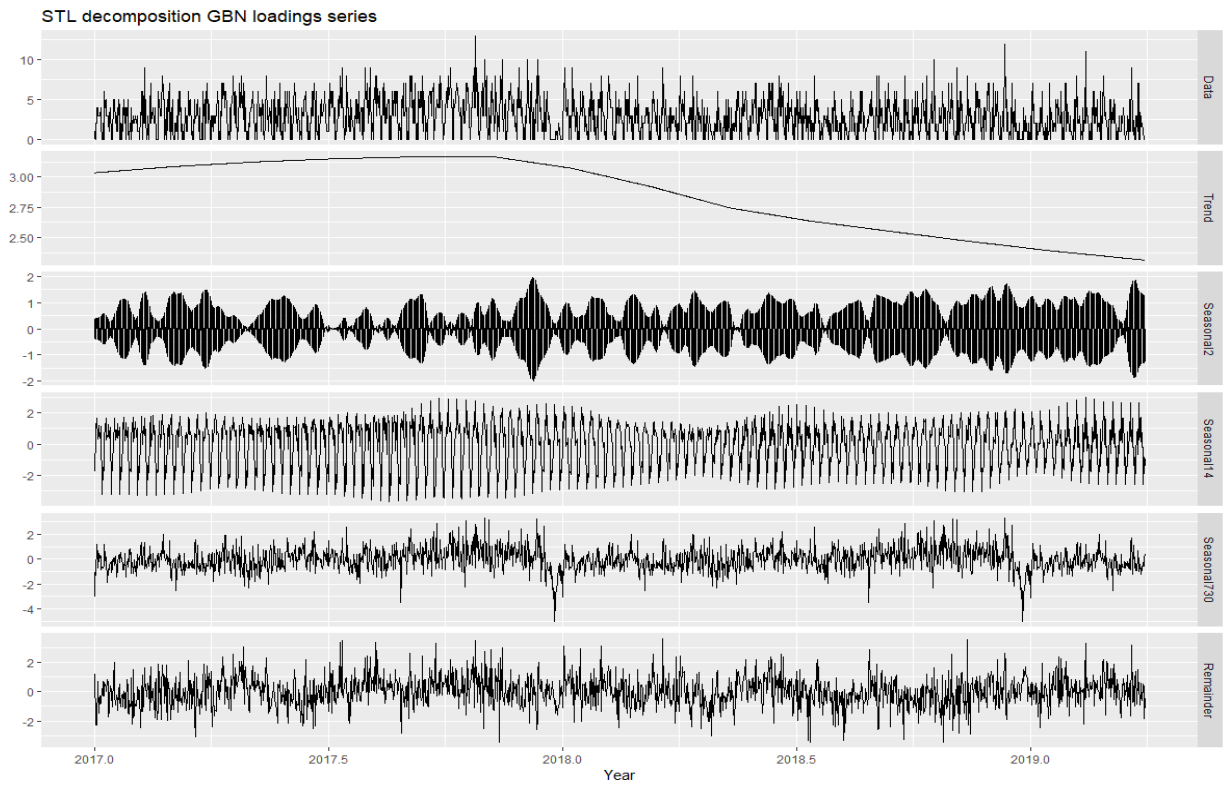


Figure 2.6 Various (seasonal) components GBN loadings series

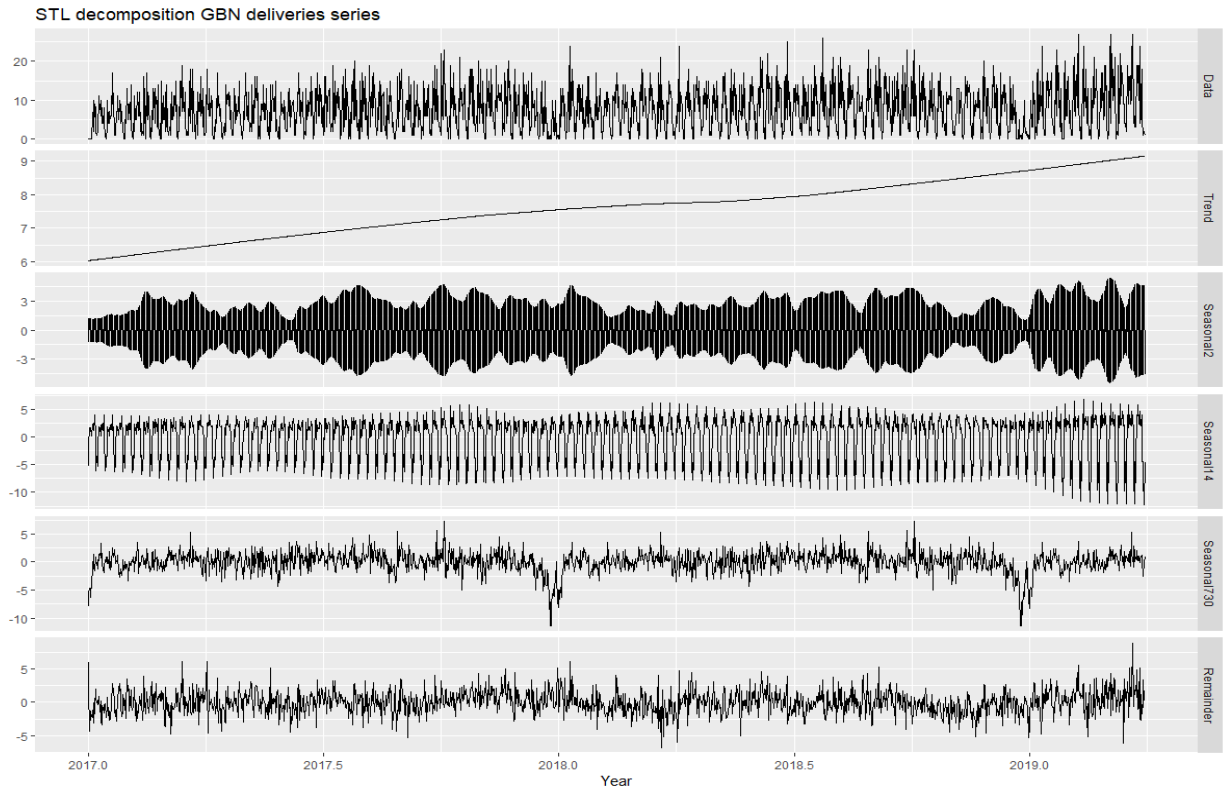


Figure 2.7 Various (seasonal) components GBN deliveries series

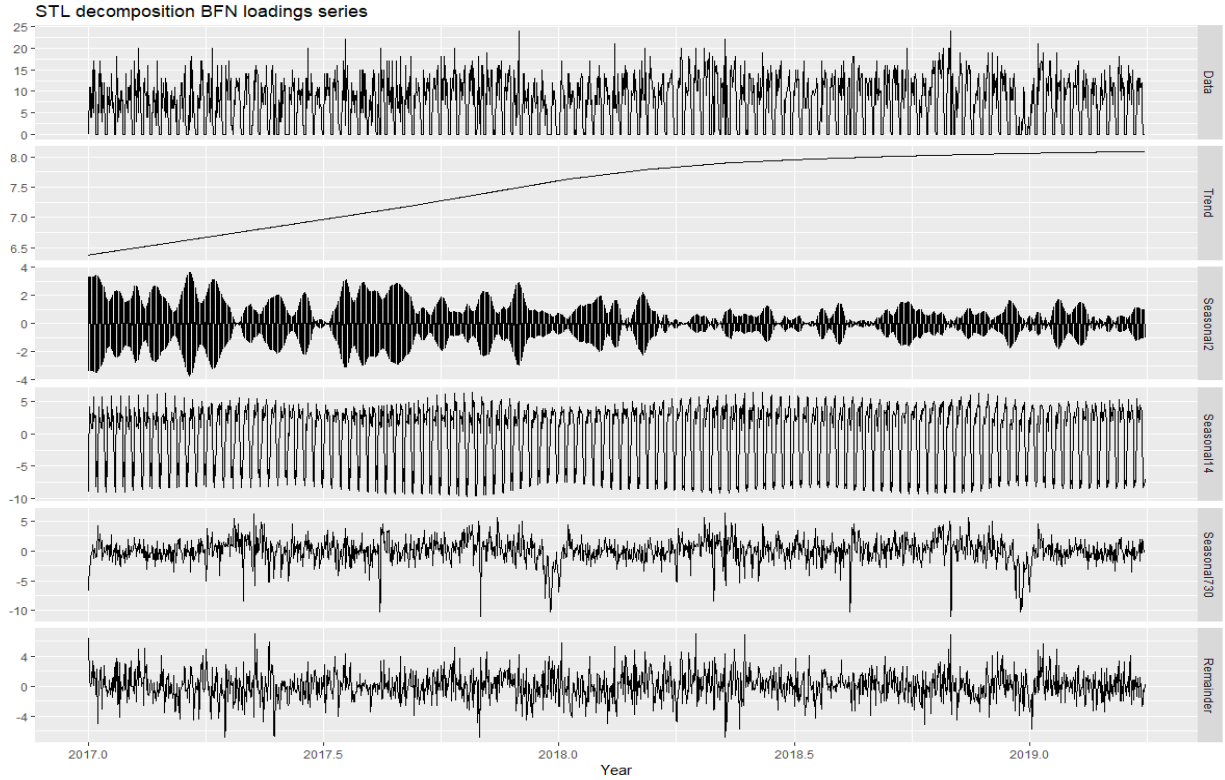


Figure 2.8 Various (seasonal) components BFN loadings series

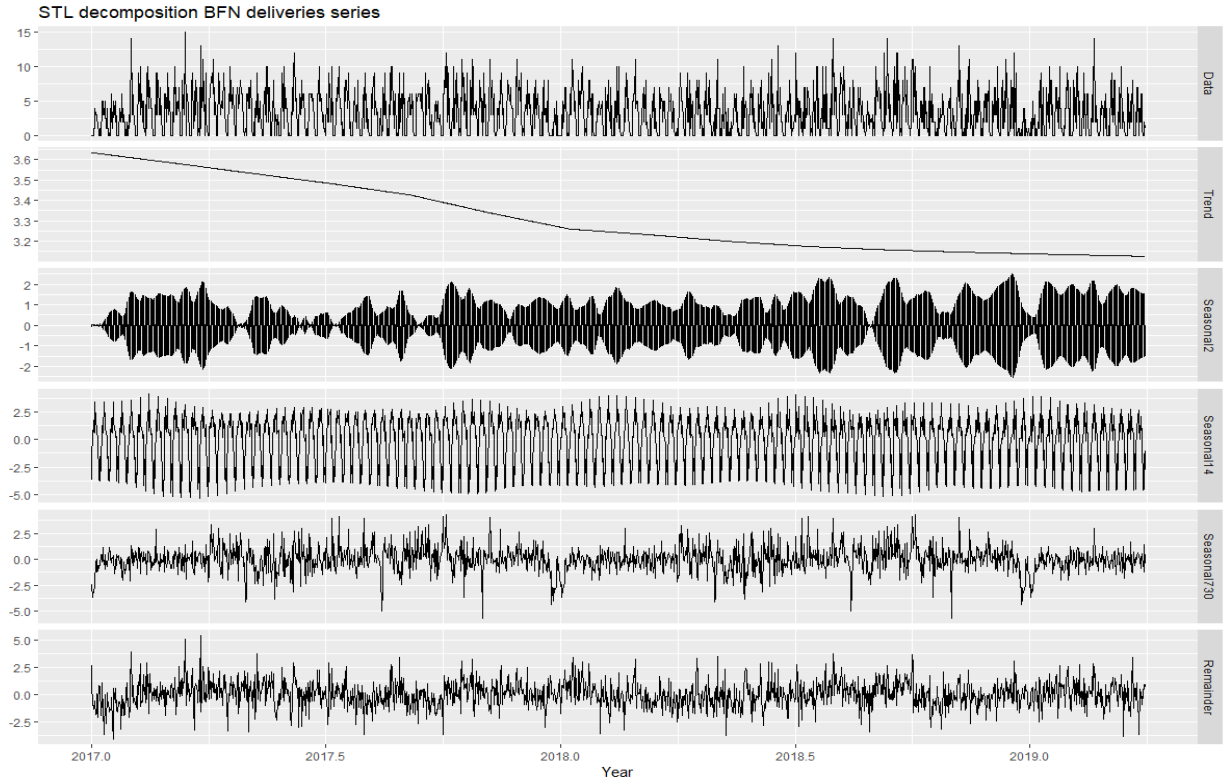


Figure 2.9 Various (seasonal) components BFN deliveries series

Visually inspecting the time series decomposition graphs is one way to determine the relative strength of the various (seasonal) components, but the strength of the components can also be measured in a more exact manner. Recall that the decomposition of a time series with multiple seasonal components can be written as

$y_t = T_t + \sum_{m=1}^{M} S_{t,m} + R_t$ where T_t is the trend component, $\sum_{m=1}^{M} S_{t,m}$ are the M different seasonal components, and R_t is the remainder component. Furthermore, also note that when the decomposition graphs were visually inspected, it was concluded that the trend component was relatively weak. It is therefore expected that the seasonally adjusted data should not exhibit much more variation than the remainder component. Put differently, the variances of the seasonally adjusted and the remainder component should be approximately the same. Exploiting this observation, the following measure of the strength of the trend component was defined by Hyndman & Athanasopoulos (2018):

$$F_t = max\left(0, 1 - \frac{Var(R_t)}{Var(R_t + T_t)}\right) \qquad \textit{Equation 2.1}$$

In this equation, F_t gives a measure of the strength of the trend component between 0 and 1. In other words, a time series with F_t close to zero exhibits almost no trend, whereas a series with F_t close to 1 indicates that the strength component is very strong. The minimal value of F_t is set to zero because in extreme circumstances, the variance of the remainder component might even be larger than the variance of the seasonally adjusted data (Hyndman & Athanasopoulos, 2018). In a similar vein, the strength of the M seasonal components can be determined by:

$$F_{S_m} = max\left(0, 1 - \frac{Var(R_t)}{Var(R_t + S_{t,m})}\right) \qquad \textit{Equation 2.2}$$

Equations 2.1 and 2.2 are thus employed to measure the strength of the various components of the time series considered in this research. These measures are then used to confirm whether or not the conclusions drawn on the basis of visually inspecting the decomposition graphs are still valid. Table 2.1 presents the measures for each time series. Not surprisingly, the strength measures of the various components confirm the earlier observations of the decomposition graphs.

Table 2.1 Strength of the various (seasonal) components

(Seasonal) Component	BFN loadings	BFN deliveries	GBN loadings	GBN deliveries	Rotterdam loadings	Rotterdam deliveries
Trend	0.082	0.019	0.096	0.190	0.008	0.229
All seasonal components combined	0.899	0.855	0.788	0.901	0.928	0.730
Annual seasonality	0.554	0.529	0.469	0.550	0.600	0.452
Monthly seasonality	-	-	-	-	0.227	0.226
Weekly seasonality	0.874	0.782	0.661	0.834	0.923	0.600
Daily seasonality	0.367	0.534	0.439	0.758	-	-

2.2 Literature on forecasting models that can deal with multiple seasonal patterns

In the previous section it was pointed out that the time series of the loadings and deliveries considered in this thesis exhibit multiple seasonal patterns. Many forecasting models are not equipped to deal with problems involving such high frequency time series data with complex multiple levels of seasonality. Therefore, the literature was consulted to examine various forecasting methods that can account for data with multiple seasonal patterns. Therefore, in the remainder of this section these models are discussed. Particular attention is devoted to the advantages and disadvantages of each method depending on a certain context.

It was noted earlier that time series decomposition can be useful to study the characteristics of the data and exploring the historical changes that have occurred over time. Time series decomposition can, however, also

be used in forecasting. One of the simplest approaches to deal with multiple seasonal patterns is, therefore, to use STL decomposition with multiple seasonal periods, along with a non-seasonal method applied to the seasonally adjusted data. In other words, a forecast can be made with each of the seasonal components being forecasted with a seasonal naïve method, and the seasonally adjusted data with for example an exponential smoothing model (Hyndman & Athanasopoulos, 2018).

Next, time series with multiple seasonal periods can also be forecasted by using an extension of standard exponential smoothing models: double exponential smoothing. Taylor (2003) first developed a double exponential smoothing method that can predict time series with two seasonal cycles: a short cycle that repeats itself many times within a longer one. The advantage of these models is that they are intuitive and often provide accurate forecasts in problem areas where the seasonal pattern is not too complex. In contexts where the seasonal pattern is more complicated (e.g. non-integer seasonality or time series with non-nested seasonal patterns), double exponential smoothing models are less appropriate to use. A further problem with double exponential smoothing models is that they assume the error process (ε_t) is serially uncorrelated. This, however, might not always be the case. Hence, it can be concluded that double exponential smoothing models provide an intuitive way to accurately forecast time series with multiple seasonal patterns as long as these patterns are not too complex. In problem areas where the seasonal pattern is more complex, methods such as dynamic regression, TBATS or artificial neural networks might be preferred.

Another way to account for long, multiple seasonal periods, might be the use of dynamic harmonic regression. In contrast to double exponential smoothing models, dynamic regression models do not assume that the error process is serially uncorrelated. Specifically, these types of models assume the error term in the regression equation follows an ARIMA process. Moreover, the dynamic harmonic regression model can also explicitly account for long, multiple seasonal periods by including Fourier terms. In the most basic sense, Fourier (or harmonic) analysis describes any data analysis procedure that describes or measures the fluctuations in a time series by comparing them with sinusoids (Bloomfield, 2004). Put differently, the Fourier decomposition of a series is a matter of explaining the series entirely as a composition of sinusoidal functions (Chianca, Ticona & Penna, 2005). By including Fourier terms, the model allows for any length of seasonality. A drawback, on the other hand, is that dynamic harmonic regression does not allow the seasonal patterns to change over time. In other words, the seasonal patterns are assumed to be fixed. Nevertheless, in practice the seasonal pattern is often remarkably constant, so it usually does not impose big problems (Hyndman & Athanasopoulos, 2018).

In situations where the fixed seasonal pattern is not a realistic assumption, an alternative could be the use of a combination of Fourier terms with an exponential smoothing model and a Box-Cox transformation. This approach was developed by De Livera, Hyndman, and Snyder (2011) and is called the TBATS model. In contrast to dynamic harmonic regression, a TBATS model does allow the seasonal period to change over time. A disadvantage of the TBATS methodology, on the other hand, is that they are usually quite slow to estimate (De Livera, Hyndman, & Snyder, 2011). In addition, TBATS models do not allow for the inclusion of explanatory dummy variables in the model. This can be problematic in contexts where moving seasonality effects exist (e.g. Easter). It can thus be concluded that dynamic harmonic regression is preferable if there are useful predictors that need to be added as additional regressors to the model. In contrast, TBATS is preferable in case the seasonal pattern changes over time.

Finally, a relatively recent technique that has proven to very accurate in forecasting short-term demand are Artificial Neural Networks (ANNs). Particularly in the problem area of predicting daily and even hourly electricity demand, ANNs are widely used nowadays (Hippert, Pedreira, & Souza, 2001). An ANN can be defined as a model of reasoning based on the human brain. The latter consists of densely interconnected information processing units, called neurons. Although an individual neuron has a very simple structure, all

the neurons combined possess a tremendous amount of processing power (Negnevitsky, 2005). The brain can be seen as a bundle of highly complex, nonlinear, and parallel information processing units. Furthermore, our brain also exhibits plasticity. Plasticity enables neurons, in response to a certain stimulation pattern, to demonstrate long term changes in the strength of their connections (Negnevitsky, 2005). In other words, connections between neurons leading to the "right" answers are strengthened while connections leading to the "wrong" outcomes are weakened. ANNs are thus machine learning tools that try to replicate the way humans learn. One attribute that characterizes ANNs is that they allow for complex non-linear relationships between the variable to be forecasted and its predictors. Neural networks also have certain serious limitations, however. Firstly, one of the problems that often occur during the design of ANNs is overfitting. When overfitting occurs, the error on the training set is driven to a very small value, but when new data is presented to the network, the error is large as the network is not able to deal with new data (Negnevitsky, 2005). Overfitting particularly occurs when the data is of limited size. The implication of this is that, in order to be able to design a suitable network, the data requirements are quite extensive (Negnevitsky, 2005). A second problem with ANNs seems to be their lack of explanatory capabilities and lack of a proper building methodology to define the network architecture (Portugal, 1995). Finally, as there are no statistical considerations involved in the modelling process of ANNs, it can only produce point forecasts (Negnevitsky, 2005).

2.3 Implementation and evaluation of the forecasting models

2.3.1 Introduction

In section 2.2 the literature was consulted to learn which forecasting models are appropriate to deal with the demand characteristics of the time series considered in this thesis. In particular, various forecasting methods were identified that can account for data with multiple seasonal patterns. The obvious next step is to implement the forecasting models, apply them to the context of this research and evaluate their predictive accuracy. The remainder of this section is therefore structured as follows. Firstly, the data that is used in forecasting the number of loadings and deliveries is briefly described. Secondly, the models that are implemented are concisely outlined. Thirdly, every model is evaluated, and some general conclusions are drawn.

2.3.2 Data description

It is probably clear by now that this chapter strives to derive six forecasts. That is, the number of loadings and deliveries are separately forecasted for the three selected regions: Rotterdam, GBN and BFN. Straightforwardly, this means that six data sets are used to derive these forecasts. More precisely, for every time series, 27 months of data is gathered. These 27 months correspond to the complete years of 2017 and 2018 and the first three months of 2019. For the Rotterdam region, the time series consists of the daily number of loadings and deliveries and for the GBN and BFN regions, the time series consists of the sub-daily number of loadings and deliveries. It is critical to assess the forecasting accuracy using genuine forecasts instead of relying on the residuals of the training set. The residuals of the data that are used to estimate the model, are not a reliable indication of how large true forecast errors are likely to be. The real accuracy of the forecast should instead be determined by considering how well a model performs on new data, that was not used to fit the model. For this reason, each time series is subdivided into two parts: a training set (the 2017 and 2018 data), and a test set (the first three months of 2019). The training set is used to estimate any parameters of a forecasting method and the test data is used to evaluate its accuracy. Since the test set is not used in fitting the model, it should provide a reliable indication of how well the model is likely to forecast on new data.

2.3.3 Implemented forecasting models

2.3.3.1 Simple Mean Method

The simple mean method provides an intuitive benchmark against which other (technically more advanced) models can be compared. The method is based on the earlier observation that the day of the week seasonality is, in all time series, the strongest component. The method works as follows. For the Rotterdam region, where the aim is to predict the daily number of loadings and deliveries, the mean number of loadings and deliveries of every day is calculated for the most recent year (i.e. 2018). The mean number of loadings or deliveries on a given day of the week is then automatically the forecast for that day of the week in the future. In a similar vein, for the GBN and BFN region, where the aim is to predict the sub-daily number of loadings and deliveries, the mean of every period is calculated and is used as the prediction for future periods. For example, the mean number of loadings and deliveries for Monday AM is calculated and constitutes the forecast for all future occurrences of Monday AM. The reason why this model is also developed in addition to the more advanced methods is that in forecasting it frequently occurs that simple, intuitive methods still outperform complex models. The equation for the GBN and BFN region is formally denoted by Equation 2.3. Note that the equation for the Rotterdam region is almost the same. The only difference is that for the Rotterdam region the AM/PM distinction is not applicable. Consequently, m therefore disappears out of the equation and the number of dummy variables is therefore reduced to seven, only representing the days of the week.

$$\hat{X}_t = \sum_{k=1}^{7} \sum_{m=1}^{2} \mu_{k,m} * x_{k,m,t} \qquad \textit{Equation 2.3}$$

Where,

$\hat{X}_t$	the predicted number of loadings or deliveries on period t
$\mu_{k,m}$	the historical mean number of loadings or deliveries on day k and part of the day m
$x_{1,1}$	1 if period t is a Sunday AM, 0 otherwise
$x_{1,2}$	1 if period t is a Sunday PM, 0 otherwise
$x_{2,1}$	1 if period t is a Monday AM, 0 otherwise
$x_{2,2}$	1 if period t is a Monday PM, 0 otherwise
$x_{3,1}$	1 if period t is a Tuesday AM, 0 otherwise
$x_{3,2}$	1 if period t is a Tuesday PM, 0 otherwise
$x_{4,1}$	1 if period t is a Wednesday AM, 0 otherwise
$x_{4,2}$	1 if period t is a Wednesday PM, 0 otherwise
$x_{5,1}$	1 if period t is a Thursday AM , 0 otherwise
$x_{5,2}$	1 if period t is a Thursday PM , 0 otherwise
$x_{6,1}$	1 if period t is a Friday AM, 0 otherwise
$x_{6,2}$	1 if period t is a Friday PM, 0 otherwise
$x_{7,1}$	1 if period t is a Saturday AM, 0 otherwise
$x_{7,2}$	1 if period t is a Saturday PM, 0 otherwise

2.3.3.2 STL with multiple seasonal periods

Another relatively simple approach to deal with multiple seasonal patterns is to use STL decomposition with multiple seasonal periods along with a non-seasonal method applied to the seasonally adjusted data. Specifically, for all six time series in this thesis, each of the seasonal components is forecasted using a seasonal naïve method (in which the forecast simply equals the last value of the same season), and the seasonally adjusted data is forecasted using a simple exponential smoothing model. The Akaike information criterion (AICc) statistic is used to determine the model parameters of the exponential smoothing model. The AICc is

useful for distinguishing between various models in the same class since it is an estimator of the relative quality of statistical models for a given set of data (Neusser, 2016). Given a collection of models for the data, AICc estimates the quality of each model, relative to each of the other models.

2.3.3.3 Double Exponential Smoothing

From the literature it was concluded that time series with multiple seasonal periods can also be forecasted by using an extension of standard exponential smoothing models: double exponential smoothing. A constraint with using this method, however, is that it can only be applied to time series with two seasonal cycles: a short cycle that repeats itself within a longer one. If one wants to apply this model to the problem area of this research this thus leads to the following two complications: (1) there are more than two seasonal periods and (2) not all the longer seasonal periods are a multiple of the shorter periods. For the GBN and BFN regions, all three seasonal components were relatively strong and for this reason it is decided to not implement the double exponential smoothing model for those regions. For the Rotterdam region, on the other hand, the monthly seasonality is not very strong. Hence, for the estimation of the double exponential smoothing, it is decided to drop this seasonal component. The problem that remains, however, is that the annual seasonality (365) is not a multiple of the weekly seasonality (7). To overcome this problem, the frequency of the annual seasonality is approximated to 364 for estimating this model. The smoothing parameters and seed values are automatically estimated using least squares in R. The double exponential smoothing model of Taylor (2003) for the Rotterdam region can be written as:

$$\hat{X}_t = l_{t-1} + b_{t-1} + s_{t-7}^{(1)} + s_{t-364}^{(2)} + \varepsilon_t \qquad \textit{Equation 2.4}$$

Where,

$l_t = l_{t-1} + b_{t-1} + \alpha\varepsilon_t$

$b_t = b_{t-1} + \beta\varepsilon_t$

$s_t^{(1)} = s_{t-7}^{(1)} + \gamma_{d_1}\varepsilon_t$

$s_t^{(2)} = s_{t-364}^{(2)} + \gamma_{d_2}\varepsilon_t$

2.3.3.4 Dynamic Harmonic Regression

In the literature review conducted earlier, it was concluded that by including Fourier terms, dynamic harmonic regression is a very robust method to explicitly account for long, multiple seasonal periods. The standard equation for dynamic harmonic regression with one seasonal component (with frequency *m)* is given by:

$$\hat{X}_t = \beta_0 + \sum_{k=1}^{K} [\alpha_k s_k(t) + \gamma_k c_k(t)] + \eta_t \qquad \textit{Equation 2.5}$$

Where,

$s_k(t) = sin\left(\frac{2\pi kt}{m}\right)$ *and* $c_k(t) = \cos\left(\frac{2\pi kt}{m}\right)$

Moreover, in Equation 2.5 the error term, η_t, is assumed to follow an ARIMA process. In this research, however, there are three seasonal periods that need to be considered, rather than one. In particular, for the Rotterdam region, a weekly (frequency is 7), monthly (frequency is 30.42) and annual (frequency is 365) seasonality are considered. Similarly, for the GBN and BFN regions, a daily (frequency is 2), weekly (frequency is 14) and annual (frequency is 730) seasonality are considered. Hence, the dynamic harmonic

regression equation for Rotterdam region and the GBN and BFN regions are denoted by Equations 2.6 and 2.7, respectively. Both equations thus have three types of Fourier terms with different frequencies. The number of Fourier terms *H, I* and *J* are chosen to minimize the AICc. The values of *H, I* and *J* that minimize the AICc, are found by iterating over all possible values. Note that *H, I* and *J* are not allowed to be greater than the corresponding seasonal period divided by two (e.g. in Equation 2.7 *I* cannot be larger than 7). This naturally limits the number of values considered in the iteration.

$$\hat{X}_t = \beta_0 + \sum_{h=1}^{H} \sum_{i=1}^{I} \sum_{j=1}^{J} \left[\alpha_h sin\left(\frac{2\pi ht}{7}\right) + \gamma_h cos\left(\frac{2\pi ht}{7}\right) + \beta_i sin\left(\frac{2\pi it}{30.42}\right) + \theta_i cos\left(\frac{2\pi it}{30.42}\right) + \delta_j sin\left(\frac{2\pi jt}{365}\right) + \rho_j cos\left(\frac{2\pi jt}{365}\right) \right] + \eta_t \quad \text{Equation 2.6}$$

$$\hat{X}_t = \beta_0 + \sum_{h=1}^{H} \sum_{i=1}^{I} \sum_{j=1}^{J} \left[\alpha_h sin\left(\frac{2\pi ht}{2}\right) + \gamma_h cos\left(\frac{2\pi ht}{2}\right) + \beta_i sin\left(\frac{2\pi it}{14}\right) + \theta_i cos\left(\frac{2\pi it}{14}\right) + \delta_j sin\left(\frac{2\pi jt}{730}\right) + \rho_j cos\left(\frac{2\pi jt}{730}\right) \right] + \eta_t \quad \text{Equation 2.7}$$

2.3.3.5 TBATS

Another forecasting method that came forward during the literature review, is the TBATS methodology developed by De Livera, Hyndman, and Snyder (2011). This uses a state space model that is essentially a generalization of those underpinning exponential smoothing (De Livera, Hyndman, & Snyder, 2011). Furthermore, it also allows for a Box-Cox transformation and ARMA errors. To estimate the TBATS models, the *tbats()* function in R is used that automatically fits a TBATS model applied to a time series as described in De Livera, Hyndman & Snyder (2011). The output of this function is a $TBATS(omega, \{p, q\}, phi, \{\langle m1, k1\rangle \dots, \langle mJ, kJ\rangle\})$ model. In this model summary, $omega$ refers to the Box-Cox parameter, $\{p, q\}$ refers to the $ARMA(p, q)$ process that is used to model the error term and phi is the damping parameter. Finally, $m1, \dots mJ$ and $k1, \dots kJ$ denote the seasonal periods used in the model and the corresponding number of Fourier terms used for each seasonality (De Livera, Hyndman & Snyder, 2011).

2.3.3.6 Artificial Neural Networks

Finally, in the literature ANNs have been praised for their ability to model complex non-linear relationships between the variable to be forecasted and its predictors. As a first step in developing an ANN, the input data and target data need to be determined. Evidently, the target data in the context of this research are the number of loadings and deliveries in one of the three focus regions. For the Rotterdam region, the input data consist of seven dummy variables, representing the days of the week. Additionally, one dummy variable is included that accounts for holiday effects. The input data for the GBN and BFN regions includes three extra variables on top of the input data used for the Rotterdam series. That is, two extra dummy variables are included to account for the AM / PM distinction. Finally, a last variable is included that represents the serial number of the date. This variable turned out to enhance the accuracy for these series. This variable, however, did not enhance the accuracy for the Rotterdam region. Hence, for this region the variable was excluded in the final ANN for the Rotterdam region. The next step in estimating the ANN is to find the right number of neurons in the models and train the network. The neurons in the network are connected via links and each link has a certain numerical weight associated with it. These weights can be seen as the long-term memory of the network and the network 'learns' through repeated adjustments to these weights. In other words, training the network includes initializing the weights of ANN and update the weights from a set of training examples.

As a training function Bayesian regularization backpropagation is used. Although this algorithm requires slightly more time compared to other training functions, it often results in good generalization for difficult and small datasets. Training stops according to adaptive weight minimization (MacKay, 1992). In this back-propagation ANN, the learning algorithm consists out of two phases. Firstly, a certain input training set is given to the ANN input layer. The ANN then propagates the input pattern from layer to layer until the output pattern is generated by the output layer (Negnevitsky, 2005). It is then checked if the pattern corresponds to the desired output. In case the pattern does not corresponds to the desired output, an error is calculated and then propagated backwards through the ANN from the output layer to the input layer. Consecutively, the weights are adjusted as the error is propagated (Negnevitsky, 2005).

2.3.4 Evaluation of the implemented models

2.3.4.1 Appropriate accuracy measures

"All models are wrong, but some are useful" is a famous quote by the British statistician George Box. The intuition of the quote is that every single model will be wrong in the sense that it will never represent the exact real behaviour of a certain process. Still, even if a model cannot describe exactly the reality, it can be very helpful if it is close enough. When a forecasting model is estimated, its errors can be measured and used to determine how accurately the model predicts the actual demand (Granger & Pesaran, 2000). A forecast error is defined as the difference between the observed (actual) demand and its forecast. In other words, is refers to the unpredictable part of an observation (Hyndman & Athanasopoulos, 2018). Let Y_t denote the observation at time *t* and F_t the corresponding forecast. The error, e_t, can then be defined as $e_t = Y_t - F_t$. Accuracy measures based on e_t are formally referred to as scale-dependent measures because the forecast errors are on the same scale as the data. Generally, these measures are useful when comparing various forecasting methods on the same set of data as is the case in this chapter. Hence, the two scale-dependent accuracy measures that were used to evaluate the accuracy of the implemented models are:

$$Mean\ Absolute\ Error\ (MAE) = mean(|e_t|)$$

$$Root\ Mean\ Square\ Error\ (RMSE) = \sqrt{mean(e_t^2)}$$

The RMSE has been extensively used, largely due to its theoretical relevance in statistical modelling (Hyndman & Koehler, 2006). The RMSE is often preferred over the mean squared error since it is on the same scale as the actual data. The disadvantage of the RMSE compared to the MAE is that it is more sensitive to outliers. Ideally, a measure based on percentage or relative errors should also be included. This is not possible, however, since the time series considered in this chapter contain zeros.

2.3.4.2 Accuracy of the implemented forecasting models

After the abovementioned models were implemented, the next step was to assess and compare their forecasting accuracy. Table 2.2 to 2.4 show how the various models perform on the six time series. A number of conclusions can be drawn on the basis of these tables.

Considering the accuracy of the various models on six time series, it can be concluded that the simple mean method, dynamic harmonic regression, and ANNs outperform the other three methods. The only exception to this, is that the STL method seems to perform remarkably well for the Rotterdam deliveries series. The finding that the simple mean method can be found amongst the three best performing methods, outperforming more complex methods such as TBATS, is quite surprising. Recall that this method was implemented to provide an intuitive benchmark against which other (more advanced) models could be compared. As it turns out, however, it proves to be one of the most accurate methods. Especially in the BFN

loadings series, the performance of the simple mean method is impressive, showing the highest accuracy in terms of MAE and RMSE of all models. Although in most other series, the ANN proves to be slightly more accurate, the difference is relatively small. On the basis of the remarkable performance of the simple mean method, the conclusion seems warranted that a lot of the variation in the data can be explained by the weekly and daily seasonal components, confirming the earlier observation in section 2.1.2. Moreover, recall that the simple mean method uses the averages of the combination between a certain day of the week and part of the day (i.e. AM or PM). In other words, it utilizes the interaction effects between these seasonal components. The findings therefore suggest that these seasonal components are not only important in isolation, but that their interaction effect is also of significant importance.

Furthermore, the relatively accurate performance of the dynamic harmonic regression models indicates that the seasonal variation in the data can be modelled reasonably well using Fourier terms. However, in most series, the performance of the dynamic harmonic regression is not significantly better (and sometimes even worse) compared to the simple mean method. Finally, in most series, the ANN outperforms the simple mean method. Nonetheless, also in this case the difference in accuracy is fairly small. To be more specific, the difference in accuracy in terms of MAE of the simple mean method and the ANN is at most 5.2%.

2.3.4.2.1 Accuracy Rotterdam loadings series

The Rotterdam loadings series might demand some further examination. The reader might already have noticed that the MAE and RMSE of the Rotterdam loadings series are fairly high compared to other series. This can to a large extent be explained by the fact that the volume of loadings is very high in Rotterdam. However, for each model, the accuracy measures on the training set were significantly lower than the ones in the test set. Recall that earlier analysis showed that the trend component in all series was virtually non-existent in 2017 and 2018. Hence, when training the data on the 2017 and 2018 data, the implemented models do not anticipate that the first quarter of 2019 will look significantly different than the first quarter of 2017 or 2018. Closer examination of the data of the first quarter of 2019 in Rotterdam showed, however, that the number of loadings was, in fact, substantially higher during this quarter compared to the same time of the year in 2017 and 2018. To be precise, the number of loadings during the first quarter of 2019 (i.e. the test set) were 22% and 27% higher as compared to the same time in 2017 and 2018, respectively. This can be seen as a peculiar increase, since a comparison between the second quarter of 2019 and the same time during 2017 and 2018 shows no increase in the number of loadings. More careful examination of the data showed that this sudden increase during the first quarter of 2019 can be explained by the fact that a large number of loadings were carried out for another internal business; WHS Logistics BV. This internal business operates independently but outsourced a lot of orders to H&S Foodtrans. As a result of all this, it might be expected that the accuracy of the forecasting models for the Rotterdam series will be higher for predictions on more recent data. Results from another test set (May 2019) indeed appear to confirm this hypothesis, but more on this in section 3.4.

Table 2.2 Accuracy for the GBN loadings and deliveries time series

	Loadings series		Deliveries Series	
Model	MAE	RMSE	MAE	RMSE
Simple Mean Method	1.16	1.57	3.03	4.26
STL with multiple seasonal periods	1.17	1.66	6.02	7.86
Dynamic Harmonic Regression	1.12	1.56	3.62	4.93
TBATS	1.26	1.74	4.90	6.65
Artificial Neural Network	1.15	1.50	2.87	4.05

Table 2.3 Accuracy for the BFN loadings and deliveries time series

	Loadings series		Deliveries Series	
Model	MAE	RMSE	MAE	RMSE
Simple Mean Method	1.86	2.90	1.24	1.70
STL with multiple seasonal periods	3.00	4.29	1.45	2.19
Dynamic Harmonic Regression	1.94	3.06	1.22	1.70
TBATS	2.69	3.92	1.70	2.24
Artificial Neural Network	1.89	2.72	1.20	1.58

Table 2.4 Accuracy for the Rotterdam loadings and deliveries time series

	Loadings series		Deliveries Series	
Model	MAE	RMSE	MAE	RMSE
Simple Mean Method	5.18	7.36	2.02	3.26
STL with multiple seasonal periods	6.77	8.76	1.91	3.18
Double exponential smoothing	5.62	7.78	2.77	4.41
Dynamic Harmonic Regression	5.95	7.96	2.07	3.26
TBATS	6.72	8.90	2.21	3.34
Artificial Neural Network	4.91	6.92	1.97	3.37

2.4 Synthesis of findings regarding forecasting loadings and deliveries

The goal of this chapter was to provide an answer to the question how the loadings and deliveries could be forecasted from historical data. To achieve this, first, an exploratory analysis of the data was conducted. Analysis of various seasonal components showed that the day of the week seasonality was the strongest seasonal component for each series. For the deliveries series, the part of the day seasonal component also turned out to be very strong, whereas for the loadings series it was shown to be only moderately important. Furthermore, the annual seasonality was shown to be moderately important for all series. Finally, the trend component was almost non-existent for all series, except for the Rotterdam and GBN deliveries series. This indicates that the overall movement of the series (ignoring the seasonality and random fluctuations) is largely constant. After this exploratory analysis, various forecasting models were implemented that can account for data that exhibits multiple seasonal patterns. After evaluating the accuracy of the implemented models, a number of conclusions were drawn.

To begin with, it was concluded that the simple mean method, dynamic harmonic regression, and ANNs outperform the other three methods in almost all series. The only exception to this, is that the STL method seems to perform remarkably well for the Rotterdam deliveries series. Secondly, it was also shown that, although the ANNs outperform the simple mean method (and the dynamic harmonic regression models) in most series, the difference in accuracy was relatively small. On the basis of the remarkable performance of the simple mean method, the conclusion seems warranted that a lot of the variation in the data can be explained by the weekly and daily seasonal components and their interaction effects. Moreover, the striking performance of the simple mean method also adds to the growing corpus of research showing that complex

models do not always generate more accurate forecasts (Green & Armstrong, 2015).

Furthermore, these findings might have important practical implications. It was explained earlier that eventually, a separate forecast of the required tank container and trucking capacity needs to be generated for each planning region. Going through the effort of having to implement a complex forecasting model for each planning region might not be worthwhile if the difference in accuracy in comparison with the simple mean method is relatively small. In addition to the ease of implementation, the simple mean method might also be preferred for the reason that it is very easy to explain to planners, who actually need to use the forecast.

3 Adjusting the initial forecast by utilizing advance demand information

In Chapter 2, various forecasting models were developed that predict the number of loadings and deliveries based on historical data. It was concluded that the simple mean method, dynamic harmonic regression, and ANNs were the most accurate methods to generate this initial forecast. This chapter investigates how the initial forecast can be adjusted by exploiting advance demand information. The remainder of this chapter is therefore structured as follows. First, to acquaint the reader with what exactly advance demand information entails, some more information is provided on the topic. Second, the order flow at H&S Foodtrans is analysed to learn more about the characteristics of the advance demand information at H&S Foodtrans. Third, a Bayesian technique, which can be used to adjust an initial forecast when some future demands are known, is explained and implemented to derive a final forecast for each time series. Fourth, the performance of the final forecasts is evaluated. Finally, since Bayesian models are criticized on the grounds that their mathematical complexity makes them difficult to understand and implement for practitioners, a simpler model is developed that can serve as an alternative to the Bayesian model.

3.1 Advance Demand Information

In the context of this research, Advance Demand Information (ADI) denotes the orders that have been placed in advance. In other words, it constitutes the future number of loadings and deliveries that are already known at present. Demand for intermodal freight transportation is usually known at least some time in advance, since equipment needs to be scheduled and arranged. The fact that loadings and deliveries are known some time in advance can thus be exploited to enhance the initial forecast. It is useful to make a distinction between *perfect* ADI and *imperfect* ADI. Perfect ADI refers to orders placed in advance that are materialized without any change, whereas imperfect ADI refers to orders that are subject to changes over time (e.g. the initial order is cancelled or postponed) (Tan, 2008). The term ADI was coined by Hariharan and Zipkin (1995) and is mostly investigated in the context of production and inventory management. Hariharan and Zipkin (1995), for instance, show that (perfect) ADI enhances the performance of a continuous-time inventory system in the same way as a reduction in lead-times. Furthermore, Gallego and Ozer (2001) studied the optimality of a state-dependent order-up-to policy in a discrete time setting, where the states are formed by perfect ADI and inventory. Moreover, Karaesmen, Buzacott and Dallery (2002) researched a capacitated problem under perfect ADI and stochastic lead times. In the context of forecasting, however, the use of advance demand information has not been thoroughly researched (Tan, 2008). Abuizam and Thomopoulos (2005), however, propose a Bayesian technique to utilize ADI in forecasting. In section 3.3. this technique is described and adjusted to fit the purposes of this research.

3.2 Analysis of the Advance Demand Information at H&S Foodtrans

3.2.1 Introduction

The goal of this chapter is to utilize ADI to enhance the accuracy of the initial forecasts developed in Chapter 2. Before delving into the method that is employed to adjust the initial forecasts, the current order flow at H&S Foodtrans is analysed. In particular, it is analysed to what extent the order flow constitutes perfect or imperfect ADI and how long in advance orders are known on average.

3.2.2 Quality of the Advance Demand Information at H&S Foodtrans

Earlier, the distinction between perfect and imperfect ADI was discussed. It was pointed out that perfect ADI refers to orders placed in advance that are materialized without any change, whereas imperfect ADI refers to orders that are subject to changes over time. The Bayesian technique that is discussed in the next section assumes perfect ADI. Hence, it is important to investigate to what extent the orders at H&S Foodtrans also constitute perfect ADI. Unfortunately, there is no data available to quantitatively investigate how often previously placed orders are altered, postponed, or cancelled at a later stage. For this reason, we are forced to rely on the judgement of experienced planners at H&S Foodtrans. Interviews with a number of planners revealed that occasionally loading and delivery dates are altered due to planners pro-actively rescheduling them to balance the workload and due to unforeseen delays and disruptions during transportation. These alterations do not really matter in the context of this research, however. The reason for this is that these alterations do not really represent customers' demand for loading and deliveries. It is more problematic when loading and delivery dates are altered by customers themselves, since in this case, the alternations do represent customers' demand. According to the planners, this also happens occasionally. It can thus be concluded that the transportation demand at H&S Foodtrans constitutes imperfect ADI. This is a limitation of using the Bayesian technique, since this technique assumes perfect ADI. Nevertheless, as will become clear in later sections, the technique seems to perform remarkably well.

3.2.3 How long are orders known in advance?

In order to effectively utilize the ADI, it is important to get a better picture of how long in advance, on average, orders are known. This section, therefore analyses what percentage of orders is known at a certain point in time. Figure 3.1 to 3.6 visualize the average percentage of loadings and deliveries for future period τ that are already known at present, for each day of the week. From these figures it can be observed that, on average, orders are known fairly long in advance. This is also in accordance with the policy at H&S Foodtrans, which specifies that orders should be entered into the system at least 72 hours in advance. However, note that these figures say nothing regarding how often orders are changed after they have initially been placed. Nonetheless, the implication is that there is a lot of advance demand information that can be utilized by the method discussed in the next section. Moreover, it can also be concluded that deliveries are known longer in advance than loadings. This is of course not very surprising since at the moment an order is placed, the loading date lies closer in the future than the delivery date. Until now, the average percentage of loadings and deliveries for future period τ, that are already known at present have been discussed, without differentiating between the specific day that is taken as starting point. This does not capture the entire picture, as Figure 3.1 to 3.6 also show the percentage of future orders already known at present vary for different days. To understand the rationale behind this, consider the following example. The average percentage of orders for the coming three days, that are already known at present is likely to be different for Fridays than for Mondays. The reason for this is that during the weekend fewer orders come in than during working days. Finally, the strange "peaks" in Figure 3.4 demand some further attention as they might look odd at first glance. These peaks can be explained, however, by the fact that there are (almost) no loadings at the BFN region on Sundays. The result of this is that if τ days in the future is a Sunday, all loadings are already in the system (since there are often no loadings). In the next section it is discussed how these percentages are used to adjust the initial forecast generated in Chapter 2.

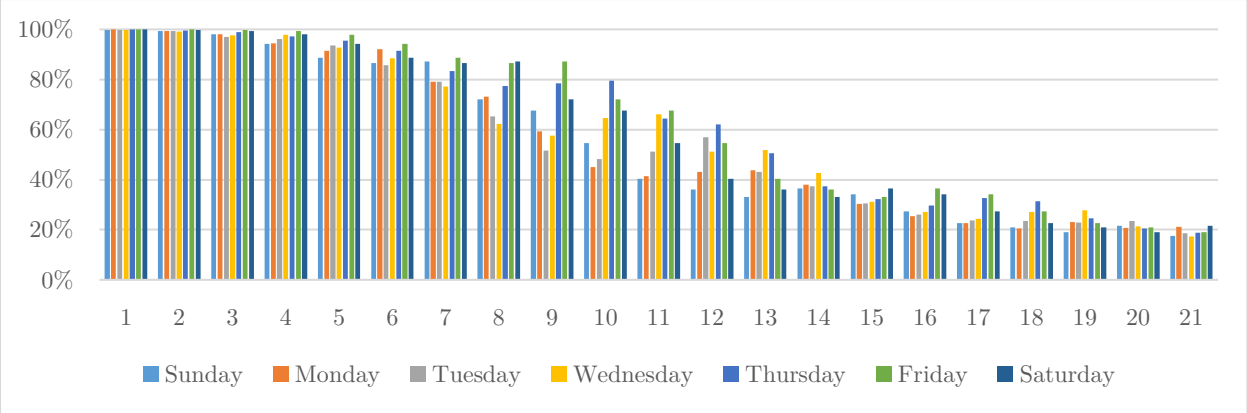


Figure 3.1 Percentage of deliveries known τ days in advance in the GBN region

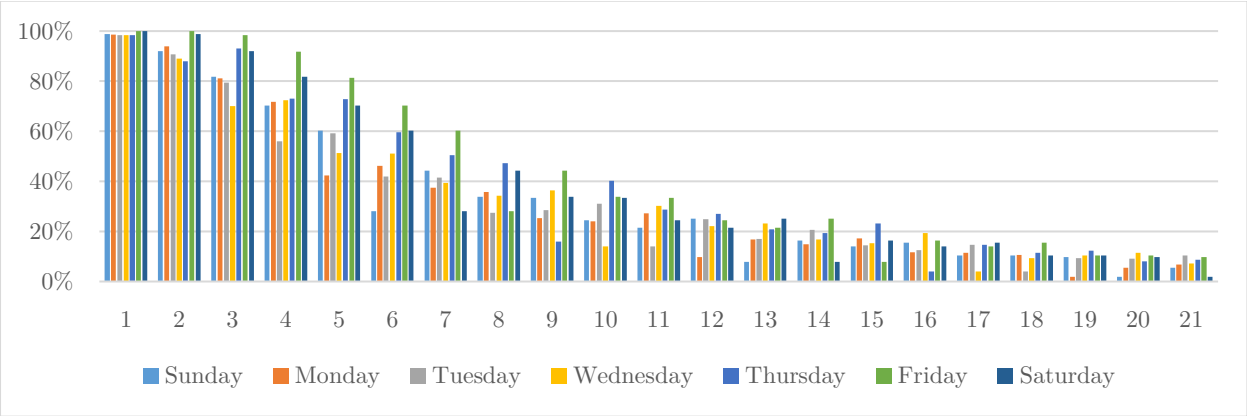


Figure 3.2 Percentage of loadings known days in advance in the GBN region

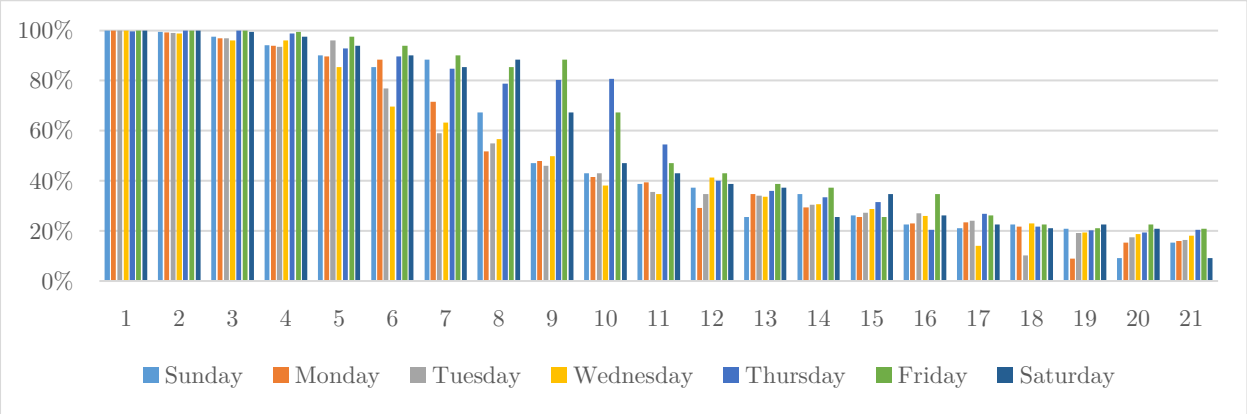


Figure 3.3 Percentage of deliveries known τ days in advance in the BFN region

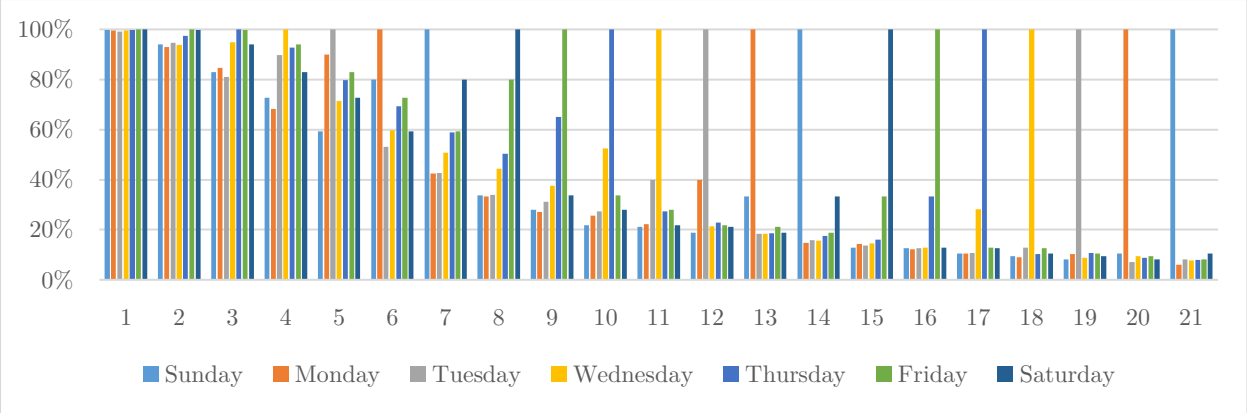


Figure 3.4 Percentage of loadings known τ days in advance in the BFN region

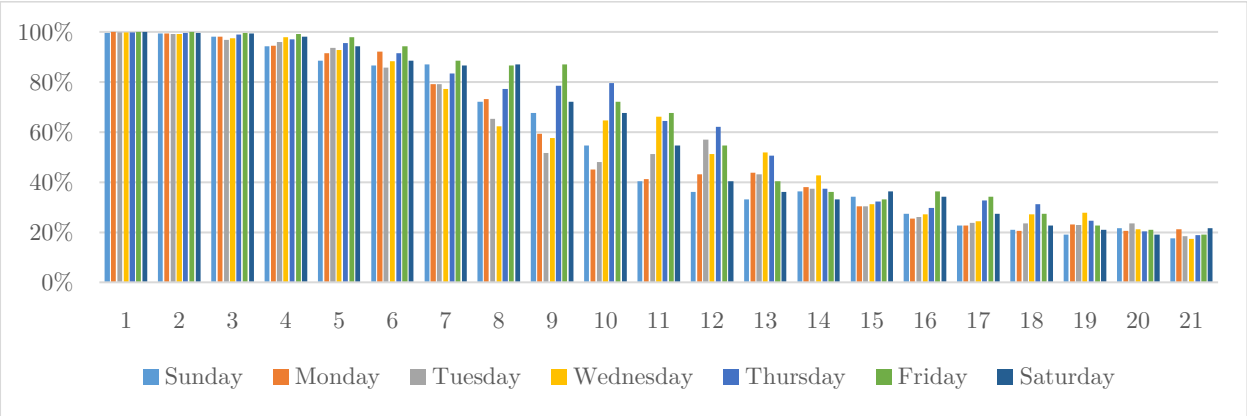


Figure 3.5 Percentage of deliveries known τ days in advance in the Rotterdam region

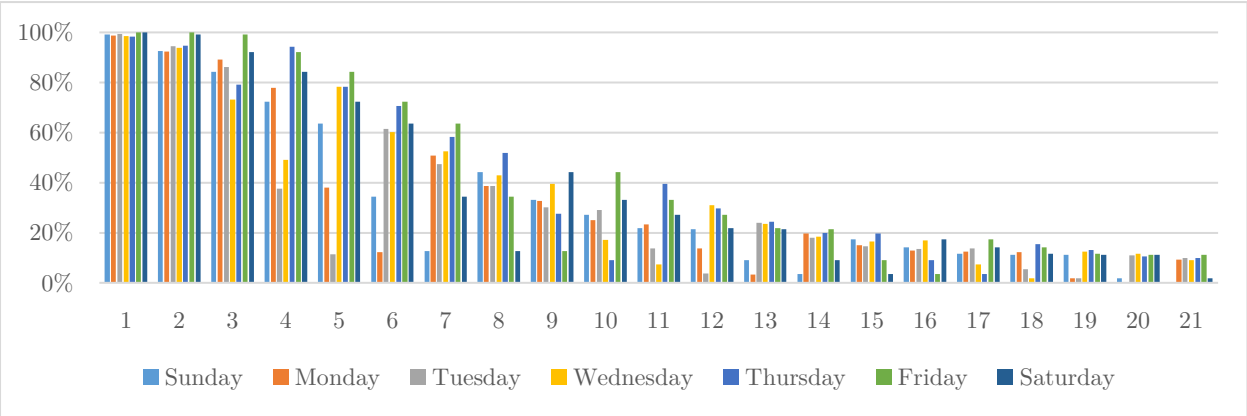


Figure 3.6 Percentage of loadings known τ days in advance in the Rotterdam region

3.3 Adjusting the initial forecast: a Bayesian technique

3.3.1 Predicting the future with Bayes' theorem

Bayes' Theorem is the workmanship of an 18th-century statistician named Thomas Bayes. It was first released in a paper entitled "An Essay Towards Solving a Problem in the Doctrine of Chances". An important implication of Bayes' line of thinking is that one must continuously update one's probability estimates when new information comes in. In the previous section it was shown that, on average, orders are known relatively long in advance. The question that now arises, is how to incorporate this information and alter the initial forecast to reflect the new information. As such, the purpose of this section is to provide a model which can

be used to adjust the original forecasts derived in Chapter 2. The model that is developed in this section is a (slightly) altered version of a model developed by Abuizam and Thomopoulos (2005). The model explores and analyses the possibility of using the expected number of orders for a future period as the variable to be estimated. Thereupon, the Bayesian estimate of the expected number of orders is used to derive the adjusted forecast (Abuizam & Thomopoulos, 2005).

3.3.2 Model development

The first step is to derive the initial forecast which comes from one of the forecasting models developed in Chapter 2. This initial forecast, at the most current time period T for τ time units in the future, is denoted as $\hat{X}_T(\tau)$. Now suppose that at time period T, information becomes available about some advanced number of orders for future period τ, denoted by $n_{a,\tau}$. Next, the adjusted forecast for period τ, can be defined as:

$$n_\tau = n_{a,\tau} + n_{R,\tau} \qquad \textit{Equation 3.1}$$

Where,

n_τ denotes the expected number of orders for future period τ (i.e. the adjusted forecast for future period τ as of time T)

$n_{a,\tau}$ denotes the number of advanced orders in future period τ as of time T

$n_{R,\tau}$ denotes the estimate for the extra orders that are expected to arrive on top of $n_{a,\tau}$

The steps of the algorithm that should be followed to ultimately result in the expected number of orders for future period τ, denoted by n_τ, are outlined below.

Step 1: Estimate of the probability distribution for the number of orders (n) and their corresponding probabilities.

Abuizam and Thomopoulos (2005) recommend using the Poisson probability distribution if the estimate of the mean number of orders is less than ten. Since the mean number of orders is based on the initial forecast, $\hat{X}_T(\tau)$, without any knowledge of the advanced demand, the mathematical formula of the Poisson probability distribution is given by:

$$P(\mathrm{n} = x) = \frac{(\hat{X}_T(\tau))^x e^{-\hat{X}_T(\tau)}}{x!} \qquad \textit{Equation 3.2}$$

In contrast, when the estimate of the mean number of orders is greater or equal to ten, it is recommended to use the Normal probability distribution with parameters $(\hat{X}_T(\tau), \hat{\sigma}_n)$, where $\hat{\sigma}_n$ is the historical estimate of the standard deviation of the number of orders (Abuizam & Thomopoulos, 2005). Recall that in the context of this research, the expected number of orders varies depending on the day of the week and part of the day. Due to this fact, a separate standard deviation is calculated for every day (for the Rotterdam series) and part of the day (for the GBN and BFN series), rather than a single one. These standard deviations are shown in Table 3.1 and 3.2.

Table 3.1 Standard deviation of the GBN and BFN series

Standard deviation	Region			
	GBN		BFN	
	Loadings	*Deliveries*	*Loadings*	*Deliveries*
Monday AM	2.02	3.97	3.65	2.60
Monday PM	1.42	2.66	4.59	2.35
Tuesday AM	2.12	3.62	4.30	2.72
Tuesday PM	1.70	2.58	3.81	1.68
Wednesday AM	2.05	4.54	3.19	2.99
Wednesday PM	1.51	2.41	4.02	1.97
Thursday AM	1.83	4.02	3.95	2.92
Thursday PM	1.67	2.73	4.00	1.63
Friday AM	1.85	3.19	3.57	2.11
Friday PM	1.40	2.22	3.76	1.36
Saturday AM	0.96	2.02	0.61	0.61
Saturday PM	0.33	0.79	0.27	0.64
Sunday AM	1.23	1.58	0.14	0.27
Sunday PM	0.68	0.88	0.00	0.60

Table 3.2 Standard deviation of the Rotterdam series

Standard deviation	Rotterdam	
	Loadings	*Deliveries*
Monday	7.84	3.17
Tuesday	5.81	3.32
Wednesday	6.41	3.32
Thursday	5.56	2.58
Friday	6.69	2.93
Saturday	0.98	0.53
Sunday	0.78	0.50

Step 2: Estimate the probability that an order for future period τ is already known as of time T

The probability that an order for future period τ is already known at time T, is denoted by θ_τ. In their study, Abuizam and Thomopoulos (2005) use three estimates of θ_τ, each occurring with a certain probability. In this research, however, a single estimate for θ_τ is used, as derived from the analysis of the order flow at H&S Foodtrans in section 3.2.3. In that section, it was calculated what percentage of total orders for period τ was already known at present. Note that θ_τ differs per forecasting region and the day of the week on which the forecast is generated. To illustrate, consider the situation in which one wants to adjust an initial seven day ahead forecast for the number of loadings in Rotterdam (i.e. $\tau = 7$). From Figure 3.6 it can easily be observed that the corresponding estimate for the probability that an loading for τ is known at present, θ_7, is 48.27%.

Step 3: Estimate the conditional probability of the advanced number of orders, for each possible value that the number of orders for period τ can take

Recall that n_j denotes the possible value j that the actual number of orders for period τ might take. The conditional probability of the advanced number of orders for each value of n_j, can then be calculated using the binomial probability distribution:

$$P\langle n_{a\tau} \mid n_j \rangle = \binom{n_j}{n_{a\tau}} {\theta_\tau}^{n_{a\tau}} (1-\theta_\tau)^{n_j - n_{a\tau}} \qquad \textit{Equation 3.3}$$

Step 4: Estimate the probability of the number of orders, given the advanced number of orders

With all the information from the previous steps, it is now possible to estimate the probability of the number of orders, given the advanced number of orders. To calculate this probability, Bayes' Theorem is used as follows:

$$P\langle n_j \mid n_{a\tau}\rangle = \frac{P\langle n_{a\tau} \mid n_j\rangle * P\langle n_j\rangle}{P\langle n_{a\tau}\rangle} \qquad \text{Equation 3.4}$$

Where,

$$P\langle n_{a\tau}\rangle = \sum_j P\langle n_{a\tau} \mid n_j\rangle * P\langle n_j\rangle$$

Step 5: Estimate the expected number of orders for future period τ

Finally, the expected number of orders for period τ is determined. As was pointed out earlier, this number represents the adjusted forecast. This forecast is derived by taking the expected value of n_j given $n_{a\tau}$, knowing that this is a Bayesian estimate of n. The final, adjusted forecast for period τ as of time T, is denoted by $\tilde{Y}_T(\tau)$. As such:

$$\tilde{Y}_T(\tau) = E\langle n \mid n_{a\tau}\rangle = \sum_j n_j P\langle n_j \mid n_{a\tau}\rangle \qquad \text{Equation 3.5}$$

Since this research is dealing separately with forecasting loadings and deliveries, the final, adjusted forecast for the expected number of *loadings* for period τ as of time T, is denoted by $\tilde{Y}_T^{Lo}(\tau)$. Similarly, the final, adjusted forecast for the expected number of *deliveries* for period τ as of time T, is denoted by $\tilde{Y}_T^{De}(\tau)$.

3.3.3 Application of the algorithm: an example

Consider the situation in which one wants to forecast the number of loadings in Rotterdam ten days from now (i.e. $\tau = 10$), and that today is a Monday. Furthermore, assume that the initial forecast at the most current time period T for ten days in the future, $\hat{X}_T(10)$, equals 21 loadings. Finally, it is known that at time T, 11 loadings are already in the system. Since the estimate of the mean number of orders is greater than ten, the normal distribution is used to estimate the probability distribution for the number of orders, n. The mean of the distribution is based on the initial forecast, $\hat{X}_T(10)$, and thus equal to 21. In order to decide the standard deviation of the distribution, the day of the week needs to be considered. Since it was assumed that today is a Monday, ten days in the future will be a Thursday. It can then be checked in Table 3.2 that the corresponding standard deviation equals 5.56. Column two in Table 3.3 presents the estimate of the distribution of the number of loadings and their corresponding probabilities. Note, as the Bayesian technique assumes perfect ADI, the value of n should be at least equal to the number of advanced loadings (i.e. $n \geq n_{a\tau}$). Next, the value of θ_{10}, the probability that a loading for ten days from now is already known at time T, needs to be estimated. The reader can check in Figure 3.6 that when a forecast is generated on a Monday, θ_{10} equals 31.42%. Thereupon, the conditional probability of the advanced number of orders (column three in Table 3.3) is calculated using Equation 3.3. Next, the probability of the number of orders, given the advanced number of orders (column five in Table 11) is determined by applying Equation 3.4. Finally, Equation 3.5 is used to compute that the expected number of orders equals 26.36. Ergo, the adjusted forecast for the number of loadings in Rotterdam ten days from now, $\tilde{Y}_T^{Lo}(10)$, is equal to 26.36.

Table 3.3 Example of the application of the Bayesian technique

n_j	$P\langle n_j \rangle$	$P\langle n_{a\tau} \mid n_j \rangle$	$P\langle n_{a\tau} \mid n_j \rangle * P\langle n_j \rangle$	$P\langle n_j \mid n_{a\tau} \rangle$	$n_j P\langle n_j \mid n_{a\tau} \rangle$
0					
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11	1,43E-02	2,94E-06	4,20E-08	1,08E-06	1,19E-05
12	1,94E-02	2,42E-05	4,70E-07	1,21E-05	1,45E-04
13	2,55E-02	1,08E-04	2,76E-06	7,08E-05	9,21E-04
14	3,25E-02	3,46E-04	1,12E-05	2,89E-04	4,04E-03
15	4,01E-02	8,89E-04	3,56E-05	9,17E-04	1,37E-02
16	4,79E-02	1,95E-03	9,34E-05	2,40E-03	3,84E-02
17	5,54E-02	3,79E-03	2,10E-04	5,40E-03	9,18E-02
18	6,20E-02	6,69E-03	4,15E-04	1,07E-02	1,92E-01
19	6,72E-02	1,09E-02	7,32E-04	1,88E-02	3,58E-01
20	7,06E-02	1,66E-02	1,17E-03	3,01E-02	6,02E-01
21	7,17E-02	2,39E-02	1,71E-03	4,41E-02	9,26E-01
22	7,06E-02	3,28E-02	2,31E-03	5,95E-02	1,31E+00
23	6,72E-02	4,31E-02	2,90E-03	7,45E-02	1,71E+00
24	6,20E-02	5,46E-02	3,38E-03	8,70E-02	2,09E+00
25	5,54E-02	6,68E-02	3,70E-03	9,52E-02	2,38E+00
26	4,79E-02	7,95E-02	3,80E-03	9,78E-02	2,54E+00
27	4,01E-02	9,20E-02	3,69E-03	9,48E-02	2,56E+00
28	3,25E-02	1,04E-01	3,38E-03	8,68E-02	2,43E+00
29	2,55E-02	1,15E-01	2,93E-03	7,53E-02	2,18E+00
30	1,94E-02	1,24E-01	2,41E-03	6,19E-02	1,86E+00
31	1,43E-02	1,32E-01	1,88E-03	4,84E-02	1,50E+00
32	1,02E-02	1,38E-01	1,40E-03	3,60E-02	1,15E+00
33	7,00E-03	1,42E-01	9,95E-04	2,56E-02	8,44E-01
34	4,68E-03	1,44E-01	6,73E-04	1,73E-02	5,89E-01
35	3,02E-03	1,44E-01	4,35E-04	1,12E-02	3,92E-01
36	1,89E-03	1,42E-01	2,69E-04	6,92E-03	2,49E-01
37	1,15E-03	1,39E-01	1,59E-04	4,09E-03	1,51E-01
38	6,73E-04	1,34E-01	9,02E-05	2,32E-03	8,81E-02
39	3,82E-04	1,28E-01	4,89E-05	1,26E-03	4,91E-02
40	2,10E-04	1,21E-01	2,55E-05	6,55E-04	2,62E-02
41	1,12E-04	1,13E-01	1,27E-05	3,27E-04	1,34E-02
42	5,78E-05	1,05E-01	6,09E-06	1,57E-04	6,58E-03
43	2,88E-05	9,72E-02	2,80E-06	7,20E-05	3,10E-03

44	1,39E-05	8,88E-02	1,24E-06	3,18E-05	1,40E-03
45	6,52E-06	8,06E-02	5,26E-07	1,35E-05	6,09E-04
46	2,96E-06	7,27E-02	2,15E-07	5,53E-06	2,54E-04
47	1,30E-06	6,51E-02	8,44E-08	2,17E-06	1,02E-04
48	5,51E-07	5,79E-02	3,19E-08	8,20E-07	3,94E-05
49	2,27E-07	5,12E-02	1,16E-08	2,98E-07	1,46E-05
50	9,02E-08	4,50E-02	4,06E-09	1,04E-07	5,22E-06
51	3,48E-08	3,94E-02	1,37E-09	3,52E-08	1,80E-06
52	1,30E-08	3,42E-02	4,45E-10	1,14E-08	5,95E-07
53	4,69E-09	2,96E-02	1,39E-10	3,58E-09	1,90E-07
54	1,64E-09	2,55E-02	4,19E-11	1,08E-09	5,82E-08
55	5,56E-10	2,19E-02	1,22E-11	3,13E-10	1,72E-08
56	1,83E-10	1,87E-02	3,41E-12	8,76E-11	4,91E-09
57	5,80E-11	1,59E-02	9,20E-13	2,37E-11	1,35E-09
58	1,78E-11	1,34E-02	2,39E-13	6,16E-12	3,57E-10
59	5,31E-12	1,13E-02	6,01E-14	1,54E-12	9,11E-11
60	1,53E-12	9,51E-03	1,45E-14	3,74E-13	2,24E-11
			$\sum_j P\langle n_{a\tau} \mid n_j\rangle * P\langle n_j\rangle = 0{,}0389$		$\sum_j n_j P\langle n_j \mid n_{a\tau}\rangle = 26.36$

3.4 Performance of the Bayesian Technique in the context of H&S

3.4.1 Data collection "Bayesian test set"

In this chapter, a Bayesian technique was introduced. The objective of this technique was to utilize advance demand information to improve the initial forecasts derived in Chapter 2. To assess whether the Bayesian technique in fact enhances the initial forecast, the accuracy of the adjusted forecast needs to be tested. However, a complication in testing the algorithm is that one needs the *most recent* number of advanced orders in future period τ as of time T for all future periods τ that one wants to include in the forecast. The problem with this, is that this data cannot be retrieved from the data warehouse, as this information is continuously overwritten. To overcome this problem, the required data for the algorithm is manually retrieved from the data warehouse for a total of one month. To be more precise, the number of advanced orders in future period τ were manually retrieved from the data warehouse each day at 16:00 for the period 27/04/2019 to 25/05/2019. As can be imagined, it is rather labour intensive to gather all the data in this manner. Hence, it was decided to only gather the data for two series: the Rotterdam and BFN loadings series. The reason that the data is collected for both these regions is that these regions have a different forecasting horizon. Recall that for the Rotterdam series, a daily forecast is generated for three weeks ahead in total, whereas for the BFN region a sub-daily forecast is generated for one week in total. This distinction results in different values of τ. I.e. the value of τ for the Rotterdam series is one day, whereas the value of τ for the BFN series is 0.5 days. Note that the data gathered in this section differs from the original test set on which the forecasting models in Chapter 2 were evaluated. To prevent confusion, the data gathered in this section is referred to as the "Bayesian test set" in the remainder of this research. Note that this test set is also used at a later stage to evaluate the accuracy of the final forecast for the required trucking and tank container capacity.

3.4.2 Assessing the accuracy of the Bayesian Technique

In order to assess whether the Bayesian technique improves the initial forecast, the accuracy of the Bayesian technique is evaluated in this section. As a first step, the required data was manually retrieved from the data

warehouse according to the steps described in the previous section. Thereupon, the Bayesian technique was applied to adjust the initial forecast for these periods. Finally, the adjusted forecast is compared with the initial forecast. Additionally, the technique is compared with the forecast accuracy if one were to use solely the advance demand information (this is the information that planners at H&S are currently using). In this case, the forecast for period τ as of time T simply equals the number of advanced orders in future period τ as of time T. Equation 3.6 is from now on referred to as the *ADI method*.

$$\tilde{Y}_T(\tau) = n_{a\tau} \qquad \textit{Equation 3.6}$$

3.4.2.1 BFN Region

Recall that for the BFN region, the simple mean method was shown to be the most accurate forecasting model for predicting both the number of loadings and deliveries. Hence, this model is used to generate the initial forecast for this region. To assess the accuracy of the Bayesian adjustment, the obvious first step might be to compare the MAE of the initial forecast, the ADI method, and the Bayesian adjustment. Table 3.4 summarizes these results.

Table 3.4 Average accuracy over τ Bayesian adjustment BFN loadings

Method	MAE
Initial forecast	2.87
ADI method	1.92
Bayesian adjustment	1.01

A number of conclusions can be drawn on the basis of Table 3.4. First and foremost, it shows that the Bayesian adjustment significantly improves the initial forecast, reducing the MAE from 2.87 to 1.01 Moreover, it also significantly outperforms the ADI method. Secondly, Table 3.4 shows that the ADI method (which only uses the information that is already at hand when the forecast is generated) outperforms the initial forecast. This is not entirely surprising, as the vast majority of the orders is already known relatively long in advance. What is surprising, on the other hand, is the high MAE of the initial forecast. Recall that the MAE of the simple mean method in the test set (first three months of 2019) was 1.86. In contrast, the MAE during this application was 2.87. The most likely explanation for this might be that the test set gathered for this application (one month of data) is relatively small, giving rise to more variable results. The absence of a larger test set might thus be an important limitation that makes it difficult to arrive at very robust conclusions with regard to absolute errors of the Bayesian algorithm. However, the goal of this real-world application of the Bayesian technique lies not so much in determining the absolute error. Instead, interest lies in assessing the degree to which the Bayesian adjustment enhances the initial forecast. As shown by Table 3.4, the results demonstrate that the Bayesian adjustment significantly improves the initial forecast on the test set as a whole.

If one would only consider the MAE on the test set as a whole (as presented in Table 3.4), one would miss out on important information, however. It is very likely, for example, that the accuracy of the ADI method and the Bayesian adjustment will vary depending on τ. More specifically, it expected that the forecast accuracy of these models will reduce in τ (i.e. if one generates a forecast for longer in the future). In contrast, since the initial forecast does not use ADI at all, it is expected that this method does not depend on τ. It might therefore be useful to check the MAE for each level of τ (recall that for the BFN region, τ is measured in 0.5 days to account for the AM/PM distinction). Figure 3.7 presents the accuracy of the three methods for each level of τ. As expected, Figure 3.7 shows that for the smaller values of τ, the ADI method and the Bayesian adjustment show approximately the same level of accuracy. Intuitively, however, the accuracy of the two models diverge as τ increases. Furthermore, it can also be observed that the initial forecast does not depend on τ, as it does not use ADI. It can therefore be expected that, if τ is increased further, the Bayesian

adjustment will eventually align with the initial forecast. In the next section, the accuracy for the Rotterdam series is assessed. For this series, a forecast is generated for 21 days in the future. Hence, for this series it can, indeed, be observed that the Bayesian adjustment aligns with the initial forecast for larger values of τ.

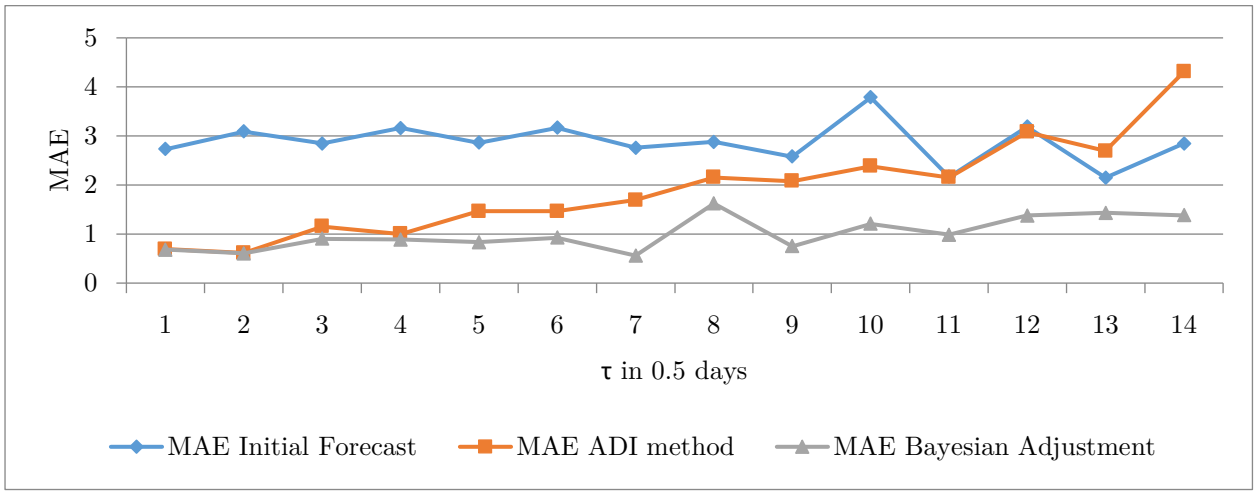


Figure 3.7 Accuracy as a function of τ BFN loadings

3.4.2.2 Rotterdam Region

In Chapter 2, the ANN was shown to be the most accurate for the Rotterdam loading series. Hence, this model is used to generate the initial forecast. Similarly as for the BFN region, it is observed that the Bayesian adjustment significantly improves the initial forecast, reducing the MAE from 2.87 to 2.08 (Table 3.5). Furthermore, unsurprisingly, the accuracy of ADI method diminishes tremendously as τ increases due to the fact that less order information is known for days that lay further in the future. For the same reason, the Bayesian adjustment aligns with the initial forecast for larger values of τ.

Table 3.5 Average accuracy over τ Bayesian adjustment Rotterdam loadings

Method	MAE
Initial forecast	2.87
ADI method	9.97
Bayesian adjustment	2.08

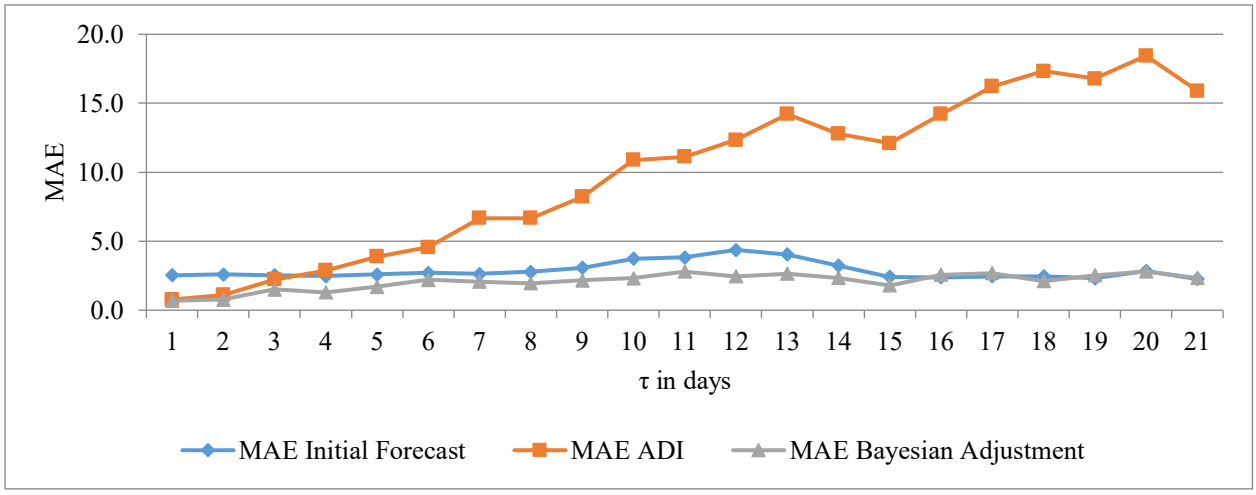


Figure 3.8 Accuracy as a function of τ Rotterdam loadings

3.5 Alternative method to utilize ADI

3.5.1 Limitation of the Bayesian technique

If a manager gets the opportunity to adopt a (complicated) forecasting model that could improve forecast accuracy, will he or she take it? Chances are this will, to a large extend, depend on the degree to which it is feasible to actually implement the model. In the previous section it was shown that the Bayesian technique proved to be rather accurate, outperforming both the ADI method and the initial forecast. However, despite its accuracy and mathematical elegance, Bayesian models are widely criticized on the grounds that their complexity requires a "specialist" to understand and implement them (Guerrero & Elizondo,1997; Tan, 2008; Utley & May, 2010). For this reason, a relatively simple model is developed in this section that can serve as alternative for the Bayesian technique.

3.5.2 Model development: The Combined Forecast

A very intuitive and quick forecasting model that utilizes ADI is the inflator algorithm. Suppose that at a certain day, 20 orders are already in the system for 5 days in the future. Moreover, from historical data, the proportion of orders that is already known 5 days in advance can be estimated (e.g. 80%). The five day ahead forecast can now be computed as $\frac{20}{0.8} = 25$. Recall that in section 3.2.3. the average percentage of loadings and deliveries for future period τ that are already known at present was calculated for each day of the week. These percentages were denoted by θ_τ and can be used in the inflator algorithm:

$$\tilde{Y}_T(\tau) = \frac{n_{a\tau}}{\theta_\tau} \qquad \textit{Equation 3.7}$$

Note that in Equation 3.7, θ_τ is different for each day of the week. That is, the percentage of loadings and deliveries for future period τ that are already known at present, is different for each day of the week. The greatest strength of the inflator algorithm is that it very intuitively captures the ADI. The main problem with this approach, on the other hand, is that its variability tends to increase dramatically for larger values of τ (i.e. if a forecast is generated for the more distant future). How this works is easily illustrated with an example. Suppose that at a certain day, 5 orders are already in the system for 21 days in the future. Furthermore, the proportion of orders that is already known 21 days in advance, is 5% on average. According to the inflator algorithm, this would result in a forecast of 100 orders. Merely increasing the ADI with one order (from 5 to 6) would, instead, result in a forecast of 120. Similarly, decreasing the ADI with one order (form 5 to 4) would result in a forecast of 80. This example illustrates that, although the inflator algorithm does capture important information regarding advance orders, its variability increases tremendously as θ_τ decreases. Note that the Bayesian technique discussed earlier does not suffer from this problem as it fits a distribution with mean equal to the initial forecast. As such, the Bayesian technique is much more stable, even when θ_τ decreases. To resolve the problem of increased variability for larger values of τ inherent in the inflator algorithm, this section suggests combining the inflator algorithm with the initial forecast derived in Chapter 2. Let $\tilde{Y}_T^a(\tau)$ and $\tilde{Y}_T^b(\tau)$ denote the initial forecast and the forecast generated by the inflator algorithm, respectively. The combined model can then be defined as:

$$\tilde{Y}_T(\tau) = \alpha_\tau * \tilde{Y}_T^a(\tau) + (1 - \alpha_\tau) * \tilde{Y}_T^b(\tau) \qquad \textit{Equation 3.8}$$

In Equation 3.8, α_τ represents a parameter that determines how the weight is distributed over the initial forecast and the forecast generated by the inflator algorithm. It was already mentioned that the forecast accuracy of the inflator algorithm decreases in τ. In contrast, the accuracy of the initial forecast does not depend on τ. Hence, for larger values of τ, more weight should be given to the initial forecast. Put differently,

α_τ should increase for larger values of τ. Although it is clear that α_τ should increase in τ, the question remains what values should be assigned to $\alpha_1, \dots, \alpha_{21}$ exactly. In other words, the aim is to find weights that are likely to yield low errors for the combined forecast. According to a procedure described by Bates and Granger (1969) the weights can be determined from past known errors of the two series. Note that to estimate the weights, different historical data is used compared to the data set on which the Bayesian algorithm was tested earlier. Since different data is used to estimate the weights, the performance of the combined forecast can later be contrasted to the Bayesian technique on the same data set, resulting in a fair comparison. In line with the procedure coined by Bates and Granger (1969), let E_τ^a and E_τ^b denote the summation of the squared forecast errors for the τ step ahead forecast of the initial forecast and the forecast generated by the inflator algorithm, respectively. Henceforth, the weights can simply be computed as follows:

$$\alpha_\tau = \frac{E_\tau^b}{E_\tau^b + E_\tau^a} \qquad \textit{Equation 3.9}$$

3.5.3 Assessing the accuracy of the Combined Forecast

The combined forecast as described in the previous section was developed as a quick and intuitive alternative to the Bayesian technique in case that the latter proves to be too difficult or time consuming to implement. In this section, the performance of the combined forecast is assessed and compared to the Bayesian technique. Similarly as in section 3.4, the forecast is tested for the Rotterdam and BFN loadings series. The weights for the Rotterdam and BFN loading series were computed, following the procedure Bates and Granger (1969) and employing Equation 3.9 and are depicted in Table 3.6. As expected, it can be observed that α_τ increases in τ, shifting the emphasis gradually from the inflator algorithm to the initial forecast. Note that for the BFN loadings series the weights have only been computed for the first seven days since for this region the goal of this thesis is to forecast one week ahead. Moreover, also recall that for the BFN region a sub-daily forecast is generated, distinguishing between AM and PM.

Table 3.6 estimated weights for the combined forecast

Rotterdam loadings series		BFN loadings series	
τ in days	**α_τ**	**τ in 0.5 days**	**α_τ**
1	0.002	1	0.0004
2	0.024	2	0.0004
3	0.051	3	0.0582
4	0.067	4	0.0582
5	0.167	5	0.0982
6	0.237	6	0.0982
7	0.311	7	0.1968
8	0.313	8	0.1968
9	0.440	9	0.3414
10	0.468	10	0.3414
11	0.519	11	0.3366
12	0.520	12	0.3366
13	0.592	13	0.3985
14	0.673	14	0.3985
15	0.651		
16	0.600		
17	0.596		
18	0.578		
19	0.594		
20	0.672		
21	0.728		

Now that the weights have been determined, the accuracy of the combined forecast can be assessed and compared to the Bayesian technique. Similarly as in section 3.4, the total accuracy is evaluated as well as how the accuracy changes as τ increases. In addition to the accuracy of the models discussed earlier, Table 3.7 presents the overall MAE of the combined forecast for the BFN and Rotterdam region. For the BFN region, the MAE of the combined forecast is only 0.04 higher than the Bayesian adjustment. Considering the Rotterdam region, it can be observed that the difference in accuracy between the combined forecast and the Bayesian adjustment is slightly larger. That is, the MAE of the combined forecast is 2.5, whereas it was earlier shown that the MAE of the Bayesian adjustment was only 2.08.

Table 3.7 Average accuracy over τ Combined Forecast BFN loadings series

Method	MAE Belgium and North France loadings	MAE Rotterdam loadings
Initial Forecast	2.87	2.87
Bayesian Algorithm	1.01	2.08
Combined Forecast	1.05	2.50

The question that arises based on the results in Table 3.7 is why the difference in accuracy between the Bayesian adjustment and the combined forecast in the BFN loadings series is almost negligible, whereas they differ significantly in the Rotterdam loadings series. A partial answer to this question might be found by inspecting how the accuracy changes in τ (Figure 3.9 and 3.10). Based on Figure 3.10 it seems reasonable to conclude that for small values of τ, the difference in accuracy between the combined forecast and the Bayesian adjustment seems to be (slightly) smaller. Another conclusion that can be drawn by visually inspecting Figure 3.9 and 3.10, is that the accuracy of both models seems to comove together. This is also what one would expect, since both models use the ADI as well as the initial forecast. All in all, the conclusions seems warranted that, although the combined forecast is not as accurate as the Bayesian adjustment, it can still be considered as a serious alternative to the Bayesian adjustment.

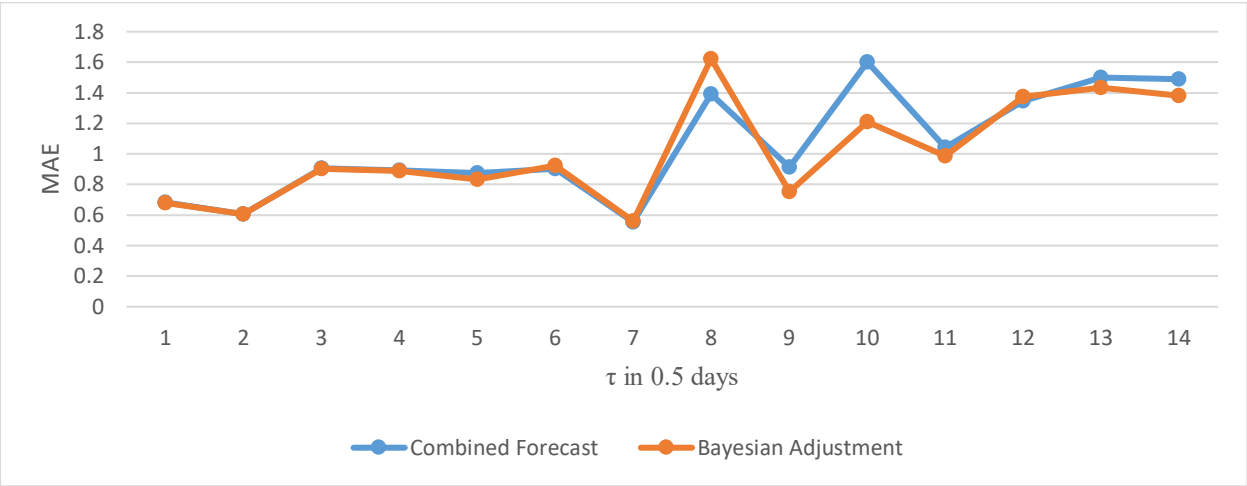


Figure 3.9 Accuracy as a function of τ BFN loadings

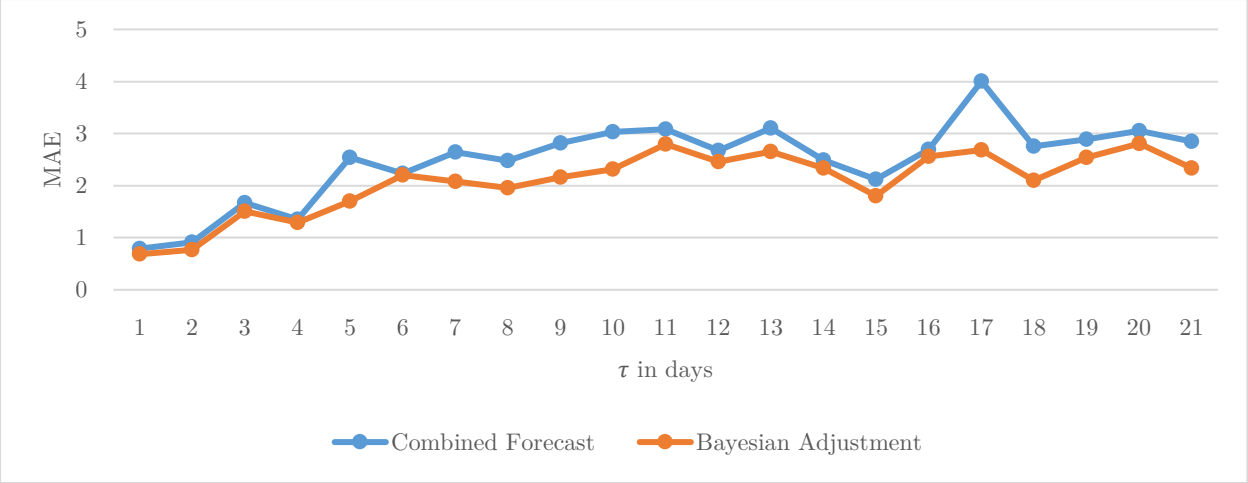


Figure 3.10 Accuracy as a function of τ Rotterdam loadings

3.6 Synthesis of findings regarding utilization of ADI

This chapter investigated how the initial forecast (derived in Chapter 2) can be adjusted by exploiting advance demand information. In particular, a Bayesian algorithm was developed that explores and analyses the possibility of using the expected number of orders for a future period as the variable to be estimated. Thereupon, the Bayesian estimate of the expected number of orders is used to derive the adjusted forecast. At a later stage, the accuracy of the Bayesian technique was evaluated, and it was shown that it significantly improves the initial forecast, especially for smaller values of τ. Finally, this chapter also acknowledged that Bayesian models are criticized for their mathematical complexity. For this reason, a relatively simple model that combines the initial forecast with an inflator algorithm was developed which can serve as alternative for the Bayesian technique. Although this combined forecast did not outperform the more complex Bayesian adjustment, it was shown to be quite accurate, especially for smaller values of τ. It was therefore concluded that the combined forecast can be considered as a serious alternative to the Bayesian adjustment if the latter proves to be too difficult or time consuming to implement. On the other hand, if practitioners care deeply about the accuracy of the forecast, it is still recommended to employ the Bayesian algorithm. Furthermore, another reason for preferring the Bayesian technique over the combined forecast might be that the Bayesian algorithm produces an entire probability distribution. In the context of forecasting this probability distribution is referred to as the forecast distribution and represents a set of values that the random variable (i.e. the loadings and deliveries) could take, along with their relative probabilities (Hyndman & Athanasopoulos, 2018). In contrast, the combined forecast only generates point forecasts. Having the entire forecast distribution might considered to be an advantage since it allows for the computation of an interval within which the forecasted variable is expected to lie with a specified probability. In other words, it allows for the computation of prediction intervals. The rationale behind calculating prediction intervals is that they express the uncertainty in the forecasts.

Finally, it is also worth mentioning that the adequate performance of the combined forecast lends support to the growing body of research suggesting that combining multiple forecasts leads to increased forecast accuracy (Armstrong, 2001; Hyndman & Athanasopoulos, 2018). Taken together, the findings of this chapter suggest that the number of loadings and deliveries can be forecasted with a fairly high level of accuracy. In the two chapters that follow, it is studied how this forecast can be translated to the required capacity in terms of trucking units and tank containers.

4 Tank Container Capacity

In Chapter 2, forecasts were produced for the number of loadings and deliveries based on historical data. To briefly recapitulate, the regions Rotterdam, GBN, and BFN were considered and for each region the number of loadings and deliveries were predicted, resulting in forecasts for six time series. Thereupon, in Chapter 3, these initial forecasts were adjusted to incorporate advance demand information, using a Bayesian technique. The aim of this chapter is to use the final forecast for the number of loadings in a certain region to predict the daily required tank container capacity for three weeks ahead. An accurate forecast for the required tank container capacity in a certain region at a given time should ultimately assist planners within H&S Foodtrans in making better decisions regarding the repositioning of empty tank containers. The remainder of this chapter is structured as follows. First, the heterogeneous tank container fleet at H&S Foodtrans is described. Particularly, attention is devoted to examining the different characteristics of the tanks, and a classification for the various type of tank containers is suggested. Secondly, the concept of hierarchical time series is introduced, and it is argued that the time series of the loadings in a certain region can be disaggregated by type of tank container. Thirdly, a number of forecasting methods for hierarchical time series are discussed, and it is concluded that the top-down approach based on forecasted proportions seems most applicable in the context of this research. Fourth, this top-down method based on forecasted proportions is employed to generate a final forecast of the tank container capacity. Finally, the accuracy of the proposed forecasting model is tested and compared with a benchmark model.

4.1 Description of the tank container fleet at H&S Foodtrans

A tank container can be defined as an intermodal container for the transport of liquids, gasses, and powder as bulk cargo. More specifically, they consist of a pressure vessel (i.e. the tank) supported and protected within an insulation and protective layer frame (Figure 4.1). This frame is made according to the International Standards Organization (ISO) and uses a type of corner castings to enable lifting and stacking. The tank design is thus governed by international regulations ensuring the safe transport of a wide range of bulk liquids (International Tank Container Organization, 2011). A standard tank container carries 25,000 litres and has a maximum gross weight of 36 metric tonnes. Moreover, usually a discharge valve is mounted at the rear end, and access for loading, cleaning, and maintenance is at the top (International Tank Container Organization, 2011). Until now, this research referred to tank containers as if there exists just one single type. This is not the case, however, as in practice, tank containers are manufactured in a range of capacities, with various configurations of valves and fitting. In the intermodal transport of liquid foodstuff, this distinction between various different types of tank containers is particularly important, since many products require very specific handling and thus a particular type of tank container. In the remainder of this section, the tank container fleet at H&S Foodtrans is outlined. In particular, the characteristics are discussed on the basis of which the various type of tank containers differ from one another. Thereupon, six mutually exclusive types are suggested, such that each container belongs to one of these groups. These six types are selected in consultation with the ones responsible for the MMP (i.e. the tank container planning) within H&S Foodtrans. This classification is also the aggregation level at which the tank container capacity is forecasted at a later stage.

Figure 4.1 Typical tank container of H&S Foodtrans

4.1.1 Characteristics of the tank container fleet at H&S Foodtrans

4.1.1.1 Food and cross tanks

Recall that H&S Foodtrans is an intermodal LSP that is mainly engaged in the transportation of liquid foodstuff. Although not their main focus, they occasionally also provide intermodal transportation of liquid non-food (e.g. silage) or light chemical products (e.g. glycerine). Within H&S Foodtrans, non-food and chemical products are referred to as "Cross" commodities. Regulation requires that each individual container may only be used for either food or cross products. Hence, in the most basic sense, the container fleet at H&S Foodtrans can be subdivided in containers for food and cross products.

4.1.1.2 Volume and size

It was earlier mentioned that a standard tank container carries 25,000 litres. However, the heterogeneous container fleet of H&S Foodtrans consists of tanks for which the content ranges from 23,000 to 36,000 litres. In addition to the volume that tanks can carry, they also differ in the number of departments. Although the vast majority of tank containers are equipped with just one department, the number of departments in a tank range from 1 to 3. Tanks with more than one department are referred to as split tanks. Finally, the containers also differ in size. Specifically, three categories can be identified: 20 foot, 23 foot and 26 foot.

4.1.1.3 Insulation, cooling, heating and other technical features

Since H&S Foodtrans is mainly involved in the intermodal transportation of liquid foodstuff, the temperature of the commodities that are transported needs to be closely regulated. Depending on, amongst other things, the type of commodity, the temperature outside and the duration of the trip, the commodity needs to be either cooled or heated. As such, an important distinction is the sub-division between insulated and not insulated tanks. Moreover, most tanks are equipped with a steam heating mechanism that produces high- and low-pressure steam to heat a commodity to the desired temperature. In addition to steam heating, tanks might also be equipped with a glycol heating, which enables continuous heating and storing of the product in the adjusted temperature. In other words, in tanks with a glycol mechanism, a commodity can be both heated and cooled during transportation. Tanks with a glycol heating mechanism are formally referred to as Reefers. Lastly, other technical features are can be considered. First, some tanks (albeit a very small minority) are equipped with a sterile filter that protects the commodity from harmful submicron particle contamination and bacterial transfer or growth. Secondly, a tank can either be allowed or not allowed for the transportation of dangerous commodities. These commodities are classified as "ADR", and only ADR certified tanks are allowed to ship these products. In the tank container fleet at H&S Foodtrans, more than 95% of the tanks are ADR certified tanks.

4.1.2 Classification of the tank container fleet at H&S Foodtrans

As has become clear from the previous section, the fleet of tank containers at H&S Foodtrans is very heterogeneous. It does not seem reasonable to derive a forecast for every single possible container type present within the container fleet at H&S Foodtrans, as the volume would become very low, tremendously increasing the variability in the forecast. In consultation with the head of the MMP department, it is therefore decided to sub-divide the container fleet at H&S Foodtrans into six mutually exclusive categories depicted in Table 4.1. According to the head of the MMP department, this classification represents the most relevant categories while still maintaining a high enough volume per category. Moreover, Figure 4.2 depicts the usage various tank container types in the Rotterdam region over the years 2017 and 2018.

Table 4.1 Classification of tank container fleet at H&S Foodtrans

Type of tank container	Number of tank containers
Cross tanks	182
Big (> 30,000 ltr) and insulated	298
Big (> 30,000 ltr) and not insulated	162
Small (< 30,000 ltr) and insulated	231
Small (< 30,000 ltr) and not insulated	88
Reefers	117

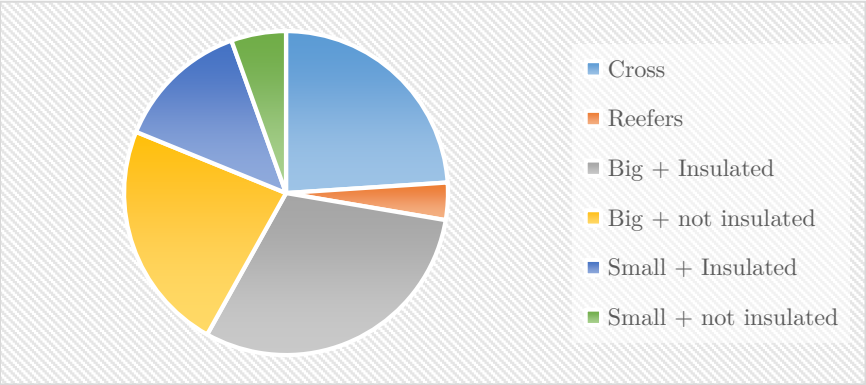


Figure 4.2 Usage of the various tank container types in Rotterdam

4.2 Considering loadings as a hierarchal time series

4.2.1 Introduction to hierarchical time series

Time series can sometimes be intuitively disaggregated by various features of interest. Consider, for instance, the time series of the total number of books sold by a book shop. This time series can be naturally disaggregated by genre such as non-fiction, fiction, young adults, et cetera. In turn, each of these can be disaggregated into even narrower categories. To illustrate, non-fiction books can be divided into science, philosophy and history and so on. The main point here is that these categories are nested within the larger group categories. As such, the collection of time series follows a hierarchical aggregation structure (Hyndman & Athanasopoulos, 2018). In a similar vein, it is argued that the time series of the loadings in Rotterdam (and any other planning region for that matter) can be disaggregated by type of tank container (Figure 4.3). For this to be the case, a number of assumptions need to be made. First, each tank can only perform a single loading at a time. This seems to be a valid assumption since H&S Foodtrans is a full truckload carrier, meaning that the shipment will take up an entire container by itself. Secondly, it is assumed that one tank can perform at most one loading per day at a single region. In other words, it is not possible to load a commodity at a certain region, deliver it at another region, clean the tank and load again in the region of origin. Although this assumption might occasionally, in extreme circumstances be violated, it is expected

that this will not affect the results too much. More specifically, over the last 840 days, there were only 12 occurrences in which a single tank performed two loadings in Rotterdam on a single day (thus accounting for less than 1.5%). In the sections that follow, some general notation regarding hierarchical time series is introduced and the literature is consulted for methods that can produce coherent forecasts across the aggregation structure. The term coherence refers to the requirement that the disaggregated forecasts need to add up in a manner that is consistent with the aggregation structure of the collection of time series as is presented in Figure 4.3.

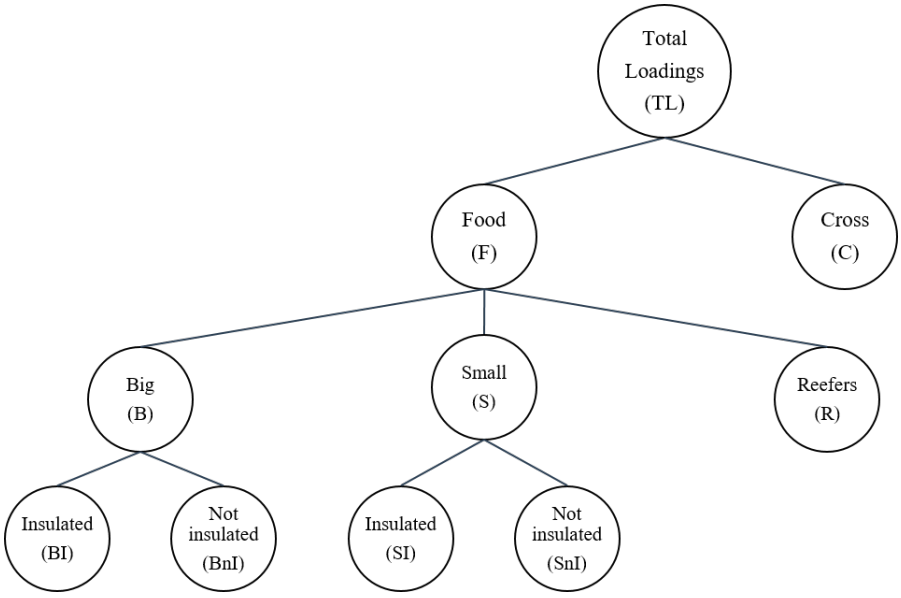


Figure 4.3 Hierarchical tree diagram of the number of loadings

4.2.2 Notation in hierarchical time series

In this section, some convenient notation is introduced that is used at a later stage when forecasts are derived for each level of aggregation. Consider the hierarchical tree diagram in Figure 4.3. The completely aggregated time series at the top is denoted as level 0. Similarly, the first level of disaggregation is referred to as level 1, and so on down to the bottom level K, which denotes the most disaggregated time series (Athanasopoulos, Ahmed & Hyndman, 2009). It is important to note that for the completely aggregated series at the top (i.e. level 0), the forecast is already known. Recall that this is the (adjusted) forecast for the total number of loadings derived in Chapter 3 (i.e. $\tilde{Y}_T^{Lo}(\tau)$). It can be observed that the hierarchical series as presented in Figure 4.3, is a $K = 3$ level hierarchy. Furthermore, let $Y_{j,t}$ denote the tth observation, of series Y_j which refers to node j on the hierarchical tree diagram (Athanasopoulos, Ahmed & Hyndman, 2009). So, to illustrate, $Y_{B,t}$ corresponds to the tth observation of the series corresponding to node B at level 2, and so on. Additionally, m_i denotes the total number of series for level i, and m thus corresponds to the total number of series in the hierarchy.

For any time t, the observations at a certain level of the hierarchy sum to the observations of the series above. For instance:

$$Y_{TL,t} = Y_{F,t} + Y_{C,t}$$

and

$$Y_{F,t} = Y_{B,t} + Y_{S,t} + Y_{R,t}$$

These equations can be considered summing equalities or aggregation constraints and can therefore be more

efficiently expressed using matrix notation (Hyndman & Athanasopoulos, 2018). Hence, a $m * m_K$ matrix is constructed, which is denoted by $\boldsymbol{S}$. Note that m_K denotes the total number of series for the bottom-level of the hierarchy. Looking at Figure 4.3, the mistake is easily made to only consider nodes $BI, BnI, SI,$ and SnI as the bottom level-series. It is important to understand, however, that the nodes C and R also belong to the bottom-level series as they are not disaggregated further. The interpretation of this matrix is that it shows how the bottom-level series are aggregated. Furthermore, let $\boldsymbol{Y}_{i,t}$ denote a vector that contains all the observations in level i at time t. Finally, let $\boldsymbol{Y}_t$ refer to a column vector containing all observations of all series at time t, such that $\boldsymbol{Y}_t = \left[Y_{TL,t}, \boldsymbol{Y}'_{1,t}, \dots, \boldsymbol{Y}'_{K,t}\right]'$ (Athanasopoulos, Ahmed & Hyndman, 2009). It is now possible to construct Equation 4.1:

$$\boldsymbol{Y}_t = \boldsymbol{S}\boldsymbol{Y}_{K,t} \qquad \textit{Equation 4.1}$$

Applying this notation to the hierarchy depicted in Figure 4.3 yields:

$$\begin{bmatrix} Y_{TL,t} \\ Y_{F,t} \\ Y_{C,t} \\ Y_{B,t} \\ Y_{S,t} \\ Y_{R,t} \\ Y_{BI,t} \\ Y_{BnI,t} \\ Y_{SI,t} \\ Y_{SnI,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Y_{C,t} \\ Y_{R,t} \\ Y_{BIt} \\ Y_{BnI,t} \\ Y_{SI,t} \\ Y_{SnI,t} \end{bmatrix}$$

Since interest lies in working with forecasts instead of the actual observations, some further notation needs to be introduced. Let $\hat{Y}_{j,T}(\tau)$ denote the τ step ahead forecast for series Y_j derived at current period T. In addition, all τ step ahead base forecasts for level i and the base forecast for the whole hierarchy are denoted by $\hat{\boldsymbol{Y}}_{i,T}(\tau)$ and $\hat{\boldsymbol{Y}}_T(\tau)$, respectively. Finally, in this research, let $\boldsymbol{P} = (p_j, \dots, p_{m_K})$ be a set of disaggregation proportions which dictate how the forecasts of the aggregated series are to be distributed to obtain forecasts for each series at the bottom-level of the structure. The revised forecasts, $\tilde{\boldsymbol{Y}}_T(\tau)$, can now be written as (Hyndman & Athanasopoulos, 2018):

$$\tilde{\boldsymbol{Y}}_T(\tau) = \boldsymbol{S}\boldsymbol{P}\hat{\boldsymbol{Y}}_T(\tau) \qquad \textit{Equation 4.2}$$

4.2.3 Forecasting hierarchical time series

One of the most simple, but widely used forecasting methods for hierarchical time series, is the bottom-up approach (Hyndman & Athanasopoulos, 2018). As the name suggests, this method starts with producing forecasts for each series at the bottom-level. Thereupon, these forecasts are summed to generate forecasts for all series in the hierarchical structure. This method has, amongst others, been applied by Dunn Williams and DeChaine (1976), Dangerfield and Morris (1992), and Zellner and Tobias (1998). An advantage of the bottom-up approach is that, since it starts at the bottom-level, no information is lost due to aggregation. A disadvantage, however, is that bottom-level data is often rather noisy, especially in case of low count data (Hyndman & Athanasopoulos, 2018). More importantly, this approach is not suitable for this research since the forecast on the most aggregate level (i.e. the total number of loadings) is already generated by the forecasting model from Chapter 2 and 3. Interest, therefore, lies in disaggregating this top-level forecast down the hierarchy. A more relevant approach for this research is therefore the top-down approach. This approach involves first generating a forecast for the most aggregated series (as shown in Chapter 2 and 3),

and then disaggregating this forecast down the hierarchy. This method has also been extensively applied to forecast hierarchical time series (Lütkepohl,1984; McLeavey & Narasimhan, 1985; Gross & Sohl, 1990; Fliedner & Mabert, 1992). The most common and simple top-down methods generate disaggregation proportions based on historical proportions of the data and were shown to be quite accurate by Gross and Sohl (1990). Particularly, Gross and Sohl (1990) specified two methods. The first method is based on average historical proportions where each proportion captures the average of the historical proportions of the bottom-level series, relative to the total aggregate (Equation 4.3) (Gross & Sohl, 1990). The second method is based on the proportions of the historical averages, where each proportion reflects the average historical value of the bottom-level series, relative to the average value of the total aggregate (Equation 4.4) (Gross & Sohl, 1990).

$$p_j = \sum_{t=1}^{T} \frac{Y_{j,t}}{Y_t} * \frac{1}{T}$$ *Equation 4.3*

$$p_j = \frac{\sum_{t=1}^{T} \frac{Y_{j,t}}{T}}{\sum_{t=1}^{T} \frac{Y_t}{T}}$$ *Equation 4.4*

In general, the simplicity of these top-down methods is their greatest strength. Moreover, these approaches seem to generate rather reliable predictions for the aggregate levels, and they are in particular applicable with low count data (Athanasopoulos, Ahmed & Hyndman, 2009). A general disadvantage, on the other hand, is the loss of information due to aggregation (Athanasopoulos, Ahmed & Hyndman, 2009). A more important disadvantage, related to the context of this research, is that these proportions are calculated on the whole series, thereby not taking into account seasonality effects of individual series (i.e. individual tank containers). The fractions of tanks used might, for instance, differ per day of the week or season of the year. This disadvantage can be overcome by another type of top-down approach, developed by Athanasopoulos, Ahmed and Hyndman (2009). Rather than disaggregating the aggregate forecast down the hierarchy based on historical proportions, this method derives forecasted proportions which are used to distribute the aggregated forecast. This approach thus seems especially appropriate for this research and will be explained in greater detail below.

4.2.3.1 Top-down approach based on forecasted proportions

This method, developed by Athanasopoulos, Ahmed and Hyndman (2009), first produces a τ step ahead forecast for the aggregate series. In the context of this research, the aggregate forecast is generated by adjusting the initial forecast derived in Chapter 2 by the Bayesian technique covered in Chapter 3. Secondly, a preliminary τ step ahead forecast is generated for all other series. These forecasts are not used directly, and they are not coherent (i.e. they do not add up correctly). The next step is to obtain the forecast proportions. These are obtained by computing the proportions of each τ step ahead preliminary forecast at the bottom level, relative to the aggregate of all τ step ahead preliminary forecasts at this level (Athanasopoulos, Ahmed & Hyndman, 2009). This course of action is then repeated for each node, going from the top, to the bottom level. Formulated more formally, let $\hat{Y}_{j,T}^{(i)}(\tau)$ denote the τ step ahead forecast of the series that corresponds to the node that is i levels above j. Moreover, let $\hat{S}_{j,T}^{i}(\tau)$ be the sum of the τ step ahead preliminary forecasts below the node that is i levels above node j, and are directly connected to that node. These forecast proportions disaggregate the initial forecast of the most aggregate series to get a set of coherent forecasts of the bottom-level series. Mathematically, the algorithm for obtaining the forecasting proportions is written as:

$$p_j = \prod_{i=0}^{K-1} \frac{\hat{Y}_{j,T}^{(i)}(\tau)}{\hat{S}_{j,T}^{(i+1)}(\tau)} \qquad \text{Equation 4.5}$$

The hierarchy in Figure 4.3 is now used as an illustration how Equation 4.5 is applied. In this example the "tilde" notation is again used to indicate the revised, coherent forecast. Note that since this is an top-down approach, the preliminary forecast of the total (i.e. most aggregated) series, equals the revised coherent forecast for the most aggregated series (i.e. $\hat{Y}_{TL,T}(\tau) = \tilde{Y}_{TL,T}(\tau) = \tilde{Y}_T^{Lo}(\tau)$). Continuing the example, the revised final (coherent) forecasts, moving down the farthest left branch of the hierarchy are:

$$\tilde{Y}_{F,T}(\tau) = \left(\frac{\hat{Y}_{F,T}(\tau)}{\hat{S}_{F,T}^{(1)}(\tau)}\right)\tilde{Y}_{TL,T}(\tau) = \left(\frac{\hat{Y}_{F,T}(\tau)}{\hat{Y}_{F,T}(\tau) + \hat{Y}_{C,T}(\tau)}\right)\tilde{Y}_{TL,T}(\tau) \qquad \text{Equation 4.6}$$

$$\tilde{Y}_{B,T}(\tau) = \left(\frac{\hat{Y}_{B,T}(\tau)}{\hat{S}_{B,T}^{(1)}(\tau)}\right)\tilde{Y}_{F,T}(\tau) \qquad \text{Equation 4.7}$$

$$= \left(\frac{\hat{Y}_{B,T}(\tau)}{\hat{Y}_{B,T}(\tau) + \hat{Y}_{S,T}(\tau) + \hat{Y}_{R,T}(\tau)}\right)\tilde{Y}_{F,T}(\tau)$$

$$\tilde{Y}_{BI,T}(\tau) = \left(\frac{\hat{Y}_{BI,T}(\tau)}{\hat{S}_{BI,T}^{(1)}(\tau)}\right)\tilde{Y}_{B,T}(\tau) \qquad \text{Equation 4.8}$$

$$= \left(\frac{\hat{Y}_{BI,T}(\tau)}{\hat{Y}_{BI,T}(\tau) + \hat{Y}_{BnI,T}(\tau)}\right)\tilde{Y}_{B,T}(\tau)$$

Note that p_{BI} can thus be written as:

$$p_{BI} = \left(\frac{\hat{Y}_{BI,T}(\tau)}{\hat{S}_{BI,T}^{(1)}(\tau)}\right)\left(\frac{\hat{Y}_{BI,T}^{(1)}(\tau)}{\hat{S}_{BI,T}^{(2)}(\tau)}\right)\left(\frac{\hat{Y}_{BI,T}^{(2)}(\tau)}{\hat{S}_{BI,T}^{(3)}(\tau)}\right) \qquad \text{Equation 4.9}$$

All other proportions and revised forecasts can be obtained in a similar way.

4.3 Forecasting the required tank container capacity

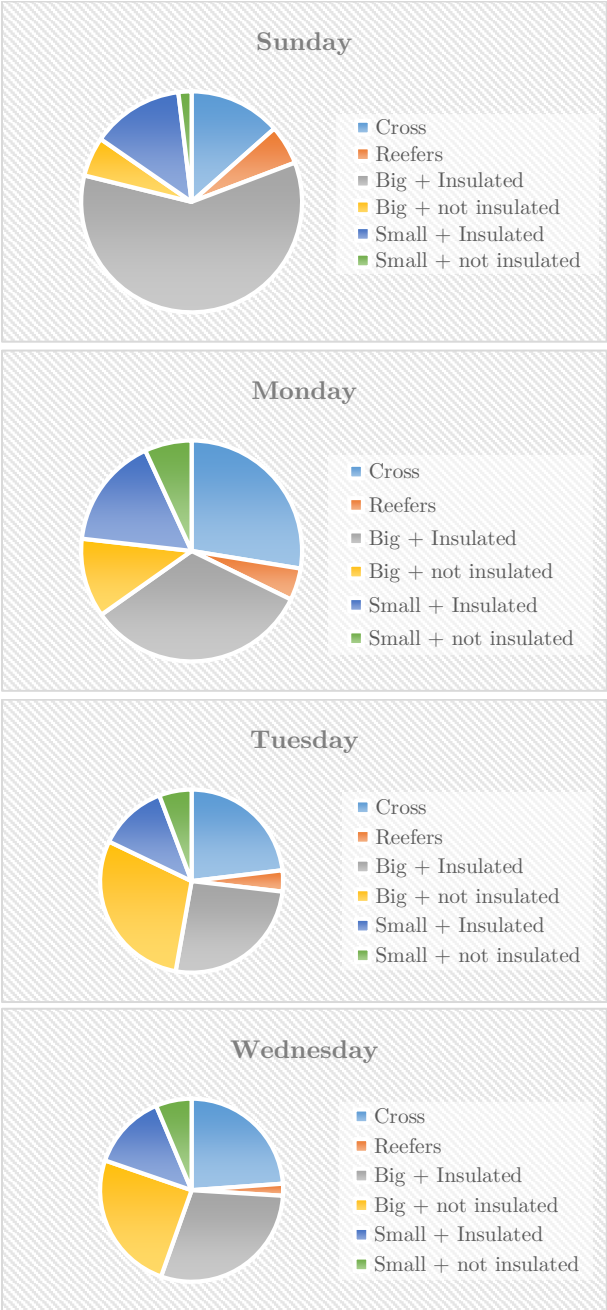
4.3.1 Seasonality in the use of tank containers

In Chapter 2, it was discussed that since this research is concerned with deriving daily, and even sub-daily forecasts, the data exhibits multiple different seasonal patterns. It was also concluded that the strongest seasonal component in the Rotterdam series was the day of the week component. As this chapter aims to disaggregate the initial forecast of the most aggregate series in order to obtain a set of coherent forecasts of the bottom-level series, a different type of seasonality needs to be examined. In particular, in Chapter 2 the seasonal effect on the total number of loadings and deliveries was analysed, whereas in this chapter the interest lies in the seasonal effect on the proportion of container types. Put differently, the question is whether the fraction of loadings carried out by a certain type of tank container affected by seasonal factors?

Figure 4.4 shows the usage of the various types of tank containers per day of the week over the last two years. The most important conclusion that can be drawn from Figure 4.4, is that working days (i.e. Monday to Friday) differ substantially from weekend days (i.e. Saturday and Sunday). It seems, for instance, that the number of big and insulated tanks is significantly larger during the weekends, whereas the number of big

and not insulated tanks is significantly smaller. In a similar vein, the difference between the small and insulated tanks and small and not insulated tanks is considerably larger in the weekends. Finally, another difference between working days and weekends is that the fraction of cross tanks seems to be somewhat smaller during the weekends, indicating that the percentage of non-food and chemical products shipped during weekends is somewhat smaller than during working days.

Figure 4.5 shows the employment of the various tank types per meteorological season. There seems to be surprisingly little variation between the various meteorological seasons. It therefore appears valid to conclude that this type of seasonality is negligible.



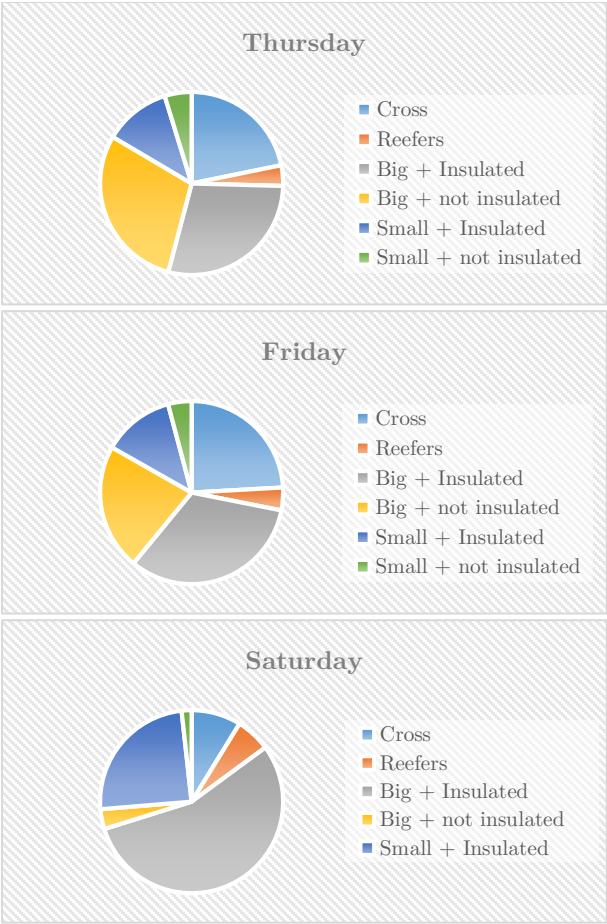
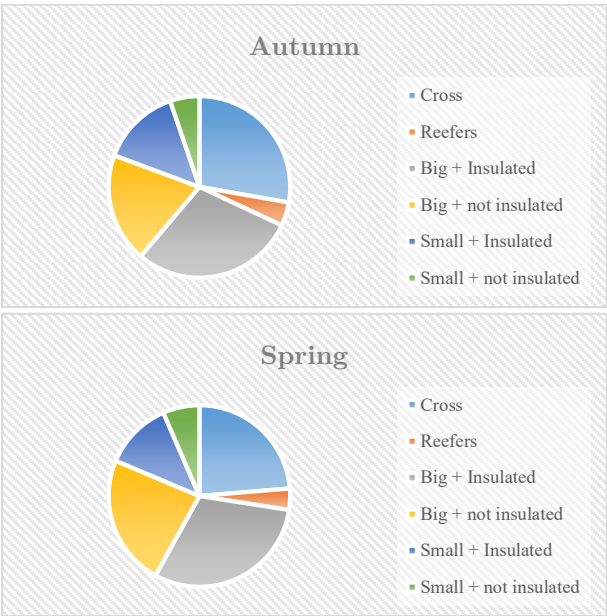


Figure 4.4 Usage of the various tank types per day of the week



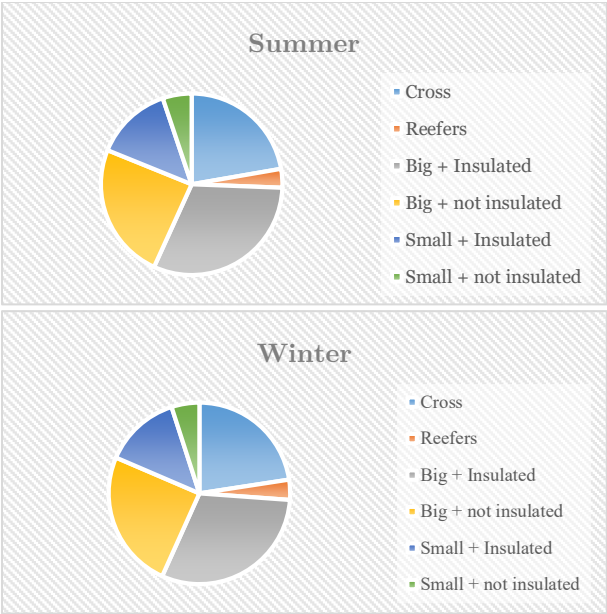


Figure 4.5 Seasonality tank container type per meteorological season

4.3.2 Methodology

Earlier this chapter, the tank container fleet at H&S Foodtrans was examined and in consultation with the MMP department, the container fleet was subdivided into six main container types. Moreover, in section 4.2 it was argued that the time series of the loadings in a certain region can be disaggregated by type of tank container. In other words, the loadings in a particular region can be considered a hierarchical time series. In the same section, also multiple forecasting approaches were discussed that can be applied to hierarchical time series. From this discussion of the various approaches, it was concluded that the top-down approach based on forecasted proportions seems most applicable in the context of this research. In this section, the top-down approach based on forecasted proportions is applied to the hierarchical time series of the loadings in Rotterdam (Figure 4.3). Recall that since this is an top-down approach, the preliminary forecast of the total (i.e. most aggregated) series, equals the revised coherent forecast for the most aggregated series (i.e. $\hat{Y}_{TL,T}(\tau) = \tilde{Y}_{TL,T}(\tau) = \tilde{Y}_{T}^{Lo}(\tau)$). Moreover, also recall that this most aggregated forecast is derived by adjusting the initial forecast from Chapter 2 by the Bayesian technique, covered in Chapter 3. Next, in accordance with the approach designed by Athanasopoulos, Ahmed and Hyndman (2009), a preliminary τ step ahead forecast is generated for all other levels of aggregation depicted in Figure 4.3 (so nine in total). These forecasts for all other series are based on the historical allocation of tanks to certain loadings.

To obtain these preliminary forecasts for all series, two forecasting methods are considered: the simple mean method and artificial neural networks. The simple mean method is selected since it has shown to deliver fairly accurate results in Chapter 2. Moreover, the method is also very intuitive and therefore easy to implement for practitioners at H&S. The second method, ANN, is selected since this method generated the most accurate forecasts for the Rotterdam loadings series in Chapter 2. To estimate the ANN, the same input variables have been used as in ANN developed in Chapter 2. The next step is to calculate the forecast proportions. These are obtained by computing the proportions of each τ step ahead preliminary forecast at the bottom level, relative to the aggregate of all τ step ahead preliminary forecasts at this level (Athanasopoulos, Ahmed & Hyndman, 2009). This course of action is then repeated for each node, going from the top to the bottom level according to the methodology as described in section 4.2.3.1.

4.3.3 Accuracy of the preliminary forecasts

In the section above it was discussed that the intuitive simple mean method and ANNs are used to generate the preliminary forecasts. Recall that these preliminary forecasts are, in the subsequent step, used to obtain the forecasted proportions. Table 4.2 summarizes the accuracy of the preliminary forecast for the various series within the hierarchy. Similarly as in Chapter 2, the simple mean method again generates fairly accurate forecasts. However, Table 4.2 also shows that, although in some series the difference is rather small, ANNs outperform the simple mean method. For this reason, it is decided to use ANNs in the remainder of this thesis to generate the preliminary forecasts.

Table 4.2 Accuracy of the preliminary forecasts

Series	MAE Simple Mean Method	RMSE Simple Mean Method	MAE ANN	RMSE ANN
Food ($\hat{Y}_F$)	4.294	6.258	3.890	4.980
Cross ($\hat{Y}_C$)	1.591	2.216	1.601	2.117
Big ($\hat{Y}_B$)	3.416	5.377	2.830	3.987
Small ($\hat{Y}_S$)	1.352	1.980	1.340	1.987
Reefers ($\hat{Y}_R$)	0.835	1.253	0.698	1.000
Big + Insulated ($\hat{Y}_{BI}$)	2.797	4.374	1.980	3.010
Big + not insulated ($\hat{Y}_{BnI}$)	1.289	1.946	1.123	1.980
Small + Insulated ($\hat{Y}_{SI}$)	1.039	1.403	1.010	1.449
Small + not insulated ($\hat{Y}_{SnI}$)	0.866	1.507	0.690	1.058

4.3.4 Summary of the complete forecasting model for predicting tank container capacity

Recall that with regards to tank container capacity, the goal of this research was to design a forecasting method that predicts the daily number of tanks of a certain type, three weeks ahead. At this point, all separate steps to generate such a forecast have been discussed. Particularly, Chapter 2 discussed in detail how an initial forecast of the expected number of loadings can be generated. Thereupon, Chapter 3 examined how this initial forecast can be adjusted, based on the advance demand information. Finally, this chapter has shown how the adjusted forecast for the expected number of loadings can be translated to the required capacity for each type of tank container. In the remainder of this section, the integrated forecasting method of how to obtain the final forecast for each type of tank container is concisely outlined.

Step 1: Generate the initial τ step ahead forecast for the expected number of loadings $\forall \boldsymbol{\tau} \in \{\mathbf{1}, \mathbf{2}, \dots, \mathbf{21}\}$

In the first step, the initial (three weeks ahead) forecast is generated for the daily number of loadings in Rotterdam. To generate this initial forecast, the ANN is used as developed in Chapter 2. Note that in principle all models developed in Chapter 2 can be used do generate this initial forecast, but that the ANN is selected as it proved to be the most accurate model. This initial forecast, at the most current time period T for τ time units in the future is denoted as $\hat{X}_T(\tau)$.

Step 2: Adjust the initial forecast by applying the Bayesian technique $\forall \boldsymbol{\tau} \in \{\mathbf{1}, \mathbf{2}, \dots, \mathbf{21}\}$

In the second step, the initial forecast as obtained in step 1 is adjusted to incorporate the advance demand information. The Bayesian estimate of the expected number of orders is used to derive this adjusted forecast. The steps of the algorithm that should be followed to ultimately result in the expected number of loadings for future period τ is outlined in Chapter 3. Note that this algorithm is followed separately for all $\tau \in \{1,2, \dots, 21\}$. The output of this step is the adjusted three weeks ahead forecast for the daily number of loadings in Rotterdam. This adjusted forecast, at the most current time period T for τ time units in the future

is denoted as $\tilde{Y}_T^{Lo}(\tau)$.

Step 3: Generate the preliminary forecasts, $\hat{\boldsymbol{Y}}_{j,T}(\boldsymbol{\tau})\ \forall\ \mathbf{j} \in \boldsymbol{J}, \forall\ \boldsymbol{\tau} \in \{\mathbf{1}, \mathbf{2}, \dots, \mathbf{21}\}$

In accordance with the approach designed by Athanasopoulos, Ahmed and Hyndman (2009), a preliminary τ step ahead forecast is generated for all series depicted in Figure 4.3. It is important to note that the preliminary forecast of the most aggregated series simply equals the adjusted forecast of the total loadings as obtained in step 2 (i.e. $\hat{Y}_{TL,T}(\tau) = \tilde{Y}_T^{Lo}(\tau)$). For all other series, the preliminary forecast is generated using the ANNs as derived earlier in this chapter. Note that a separate ANN is developed for each node in the hierarchy, resulting in nine different neural networks in total.

Step 4: Compute the forecast proportions $(\mathbf{p}_j)$ ***and the final, coherent forecasts,*** $\tilde{\boldsymbol{Y}}_{j,T}(\boldsymbol{\tau}), \forall\ \mathbf{j} \in \boldsymbol{J}, \forall\ \boldsymbol{\tau} \in \{\mathbf{1}, \mathbf{2}, \dots, \mathbf{21}\}$

The final step is to compute the forecast proportions and the revised, coherent forecasts. These are obtained by computing the proportions of each τ step ahead preliminary forecast at level i, relative to the aggregate of all τ step ahead preliminary forecasts at this level (Athanasopoulos, Ahmed & Hyndman, 2009). This course of action is then repeated for each node, going from the top to the bottom level according to the methodology as described in section 4.2.3.1.

4.3.5 Real world application and results

In order to assess the accuracy of the total forecasting model as summarized in the previous section, the model is tested on the same test set that was used before to evaluate the accuracy of the Bayesian technique (see Section 3.4.1). In short, the test set contains the data of the Rotterdam region ranging from 27/04/2019 to 25/05/2019. The accuracy of the forecasting model proposed by this research, is contrasted with the preliminary forecasts (see step 3 in section 4.3.4.), which thus serve as a benchmark for the proposed forecasting model. Recall that these preliminary forecasts are obtained using the ANNs as derived earlier this chapter.

Figure 4.6 presents the MAE per container type for both the proposed model and the benchmark model. Broadly translated, these findings indicate that for most tank container types, the proposed forecasting methodology slightly enhances the preliminary forecast. An exception to this is the Big and insulated aggregation level, for which the preliminary forecast provided more accurate results.

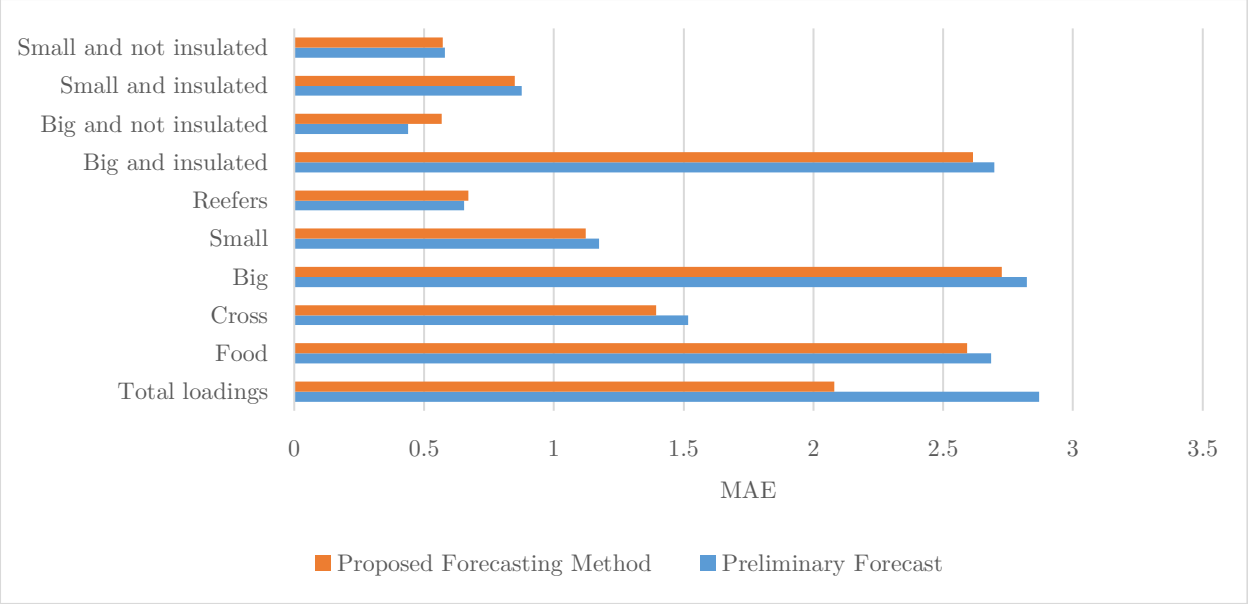


Figure 4.6 Accuracy of the proposed forecasting model vs. the preliminary forecasts

4.4 Synthesis of findings regarding forecasting tank container capacity

The goal of this chapter was to develop a forecasting methodology that can be used to predict the daily required tank container capacity for three weeks ahead. This forecast could then be employed by practitioners to make more informed and efficient repositioning decisions of empty tank containers. To achieve this goal, first the concept of hierarchical time series was introduced, and it was argued that the time series of the loadings in a certain region can be disaggregated by type of tank container. Furthermore, a top-down method based on forecasted proportions was employed to generate a final forecast of the tank container capacity. In this method, one first generates a preliminary forecast per tank container type from historical data. To obtain these forecasts, an ANN was developed for each tank container type. These forecasts are not used directly, and they are not coherent (i.e. they do not add up correctly). Instead, the preliminary forecasts are only used to calculate the forecast proportions of each tank container type. Finally, the forecasted loadings (derived from Bayesian algorithm) can be disaggregated down the hierarchy based on these forecasted proportions.

From the practitioners' point of view, all levels of aggregation are important. Higher levels of aggregation might be used to make more overarching decision such as the number of Food or Cross tanks needed in a certain region. On the other hand, the more disaggregated forecasts can be used for more detailed decisions regarding the type of tank container needed in a certain region at a given time. Applying the proposed model to a real-world test case led to the conclusion that the required tank container capacity can be forecasted with a fairly high accuracy by the proposed forecasting methodology. The accuracy in terms of mean absolute error per day expressed in number of tank containers depends on the level of aggregation and varies between 0.57 and 2.72.

Moreover, it was shown that the proposed methodology performs slightly better than the benchmark model. Be that as it may, it was also acknowledged that the difference in predictive accuracy is fairly small, especially for the more disaggregated series. Although it is difficult to arrive at any definite conclusions as to why the difference is relatively small, one possible explanation might be found in the relatively small size of the data on which the proposed model could be tested. To be more specific, the accuracy of the preliminary forecast on the Bayesian test set was rather high in comparison to the accuracy on the larger test set (see Chapter 2). As a result, the potential for improvement by the proposed forecasting method is also more limited. Although

speculative, it might therefore be expected that the difference in accuracy between the proposed forecasting methodology and the benchmark model will be greater in situations where more extreme demand patterns occur. The reason for this, is that the proposed methodology takes into account ADI, whereas the benchmark model does not utilize this information. For this reason, the proposed model is able to adapt to more unexpected demand patterns.

All in all, the conclusion seems warranted that the proposed forecasting methodology for predicting tank container capacity provides a valuable addition to the knowledge of the experienced planners at the MMP department in making repositioning decisions for empty tank containers. Moreover, it also serves as a back-up on which less experienced planners might rely in case the two experienced planners (temporarily) drop out.

5 Trucking Capacity

In Chapter 2, forecasts were produced for six time series in total. To briefly recapitulate, the regions Rotterdam, GBN, and BFN were considered and for each region the number of loadings and deliveries were predicted, resulting in forecasts for six time series. Thereupon, in Chapter 3, these initial forecasts were adjusted to incorporate advance demand information, using a Bayesian technique. Also recall that the planning regions for which this research aims to forecast the required trucking capacity are BFN and GBN. Hence, in this chapter attention is restricted to the forecasts for these regions. The aim of this chapter is to use the forecast for the number of loadings and deliveries to predict the required trucking capacity. To achieve this, the remainder of this chapter is structured as follows. First, some deeper insight is provided as to which specific actions in H&S' operation require trucking capacity. Second, the historical trucking capacity per region is estimated. This estimation is based on actual data as well as theoretical assumptions. Third, a multiple linear regression model is developed which generates forecasts for the required trucking capacity. Finally, the accuracy of the proposed regression model is assessed and compared with a benchmark model.

5.1 Actions that require trucking capacity

In order to be able to translate the (forecasted) number of loadings and deliveries into the required trucking capacity, it is essential to examine the various actions associated with a loading or delivery that demand the use of trucks. Particularly, five different trucking actions can be distinguished:

- *Pick-up action*: pick-up a tank at a certain location (e.g. terminal or depot)
- *Cleaning action*: clean the tank at a cleaning station
- *Loading action*: loading a commodity at the origin customer site
- *Drop action*: drop the tank at a certain location (e.g. terminal or depot)
- *Delivery action*: deliver the commodity at the destination customer site

Also note that in addition to the abovementioned actions, a truck needs to drive from one location (e.g. the loading location) to the next one (e.g. the drop location), which also requires trucking capacity.

Furthermore, recall that H&S Foodtrans is engaged in *intermodal bulk* transport, meaning that most of the route is travelled by rail or ferry, with the shortest possible initial and final journeys carried out by road transport (Figure 5.1). The simplified picture in Figure 5.1 shows that for one order, trucking capacity is required in both the region of origin and the region of destination. This trucking of containerized freight in the pre- and end-haulage is formally referred to as the drayage operation. More specifically, it was explained earlier that the drayage operation in intermodal transport consist of pre- and end-haulage activities that arrange the transport of containers by road between a container terminal, customer locations, and cleaning locations. That is, certain containers need to be loaded at a customer location and transported to the terminal and vice versa (Funke & Kopfer, 2016).

Figure 5.1 Intermodal shipment from loading to delivery

Although useful for illustrative purposes, Figure 5.1 does not capture the entire need for trucks in a certain region. The reason for this is that there are more trucking actions that need to be performed and thus demand trucking capacity. These actions can occur before the loading at the region of origin (e.g. pick-up the tank, clean the tank, and drive to the customer location) or after the delivery at the region of destination (e.g. cleaning, repositioning, and driving to next customer location). Hence, in order to derive an accurate estimate of the required trucking capacity, these actions also need to be considered. It is, therefore, useful to consider the full sequence from a loading at the region of origin to the *next* loading. This sequence is visualized in Figure 5.2 and 5.3. More specifically, two different sequences are possible. These two possibilities relate to the imbalance between loadings and deliveries that exists in many regions. The first sequence (Figure 5.2) depicts the situation in which, after delivery, the empty tank container is repositioned to another region by an intermodal route. In contrast, the second sequence (Figure 5.3) depicts a situation in which, after delivery, the tank container performs another loading in the same region. The dashed arrows in Figure 5.2 denote that somewhere in between the delivery and the next loading, a cleaning action needs to be performed. This can either occur in between the delivery and the drop or in between the pick-up and the second loading.

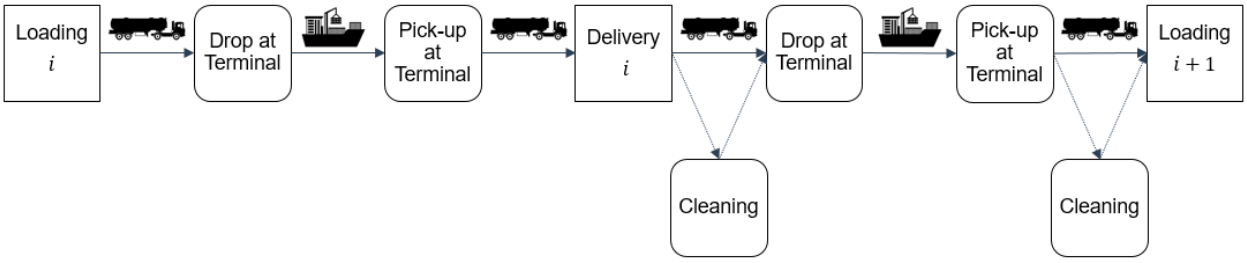


Figure 5.2 Order with empty container repositioning

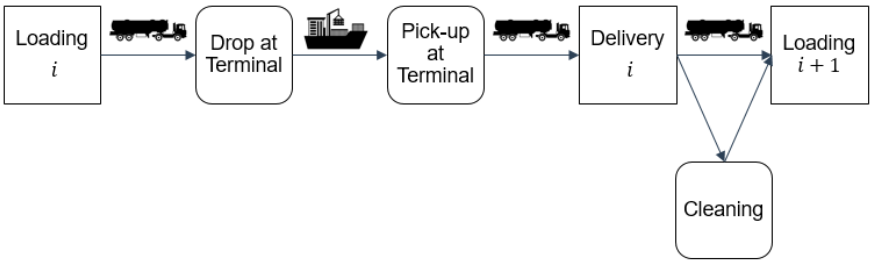


Figure 5.3 Order with return loading

Note that some intermediate stops might be added to these sequences. For instance, in between the pick-up at the terminal and the second loading, the truck might drive to a depot to wait for the second loading. Furthermore, it might also be possible that intermodal part of the trajectory is not travelled by rail or ferry, but by truck. This might for instance be the case if the client wishes an emergency shipment, as road transport is often still faster than rail or ferry. In the next section it is shown how the number of required trucking hours in a certain region can be estimated based on the various trucking actions that have just been considered.

5.2 Estimation of the historical trucking capacity

5.2.1 Available data and theoretical assumptions

In the previous section, various actions were distinguished, each requiring a certain amount of trucking capacity. The next step is to estimate the number of trucking hours required to execute these actions. This estimation is partly based on the actual data and partly on a set of theoretical assumptions. The reason for this, is that H&S Foodtrans does not have all the actual data related to how many trucking hours were needed

to perform certain set of actions.

5.2.1.1 Actual recorded data

H&S Foodtrans records the actual data of loading and delivery actions. In other words, for a loading, the time is recorded from the moment a truck starts loading the commodity at the origin customer location to the moment the truck leaves the customer location again. Similarly, for a delivery, the time is recorded from the moment the truck enters the destination customer location, to the moment at which the truck leaves the customer location again. Furthermore, H&S Foodtrans also possesses the data of the exact sequence of actions that was followed for every historical order. In other words, for every historical order, the specific trucking actions carried out by a certain truck are known. Consider, for example, the imaginary order Q depicted in Figure 5.4. For order Q the exact sequence of actions is known. In other words, it is recorded that first, a tank container was picked up at a known location. Thereupon, the truck drove to another known location where a cleaning action was performed, and so on. Moreover, the actual time it took to perform the loading and delivery action is also recorded and this data can be extracted for all historical orders.

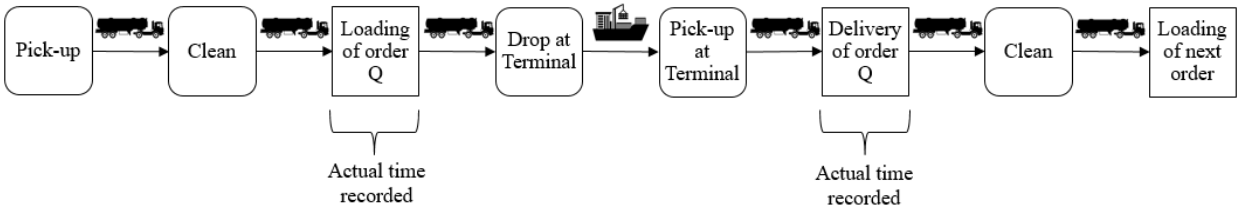


Figure 5.4 Example of imaginary order Q

5.2.1.2 Assumptions to obtain the missing data

To recap, for every historical order, the exact sequence of actions is known. Moreover, the actual historical trucking hours for the loadings and deliveries are documented and can thus be extracted for all historical orders. More information is needed, however, to approximate the total trucking hours that were needed to execute a given order. In particular, assumptions need to be made regarding the trucking hours required for a pick-up action, drop action and cleaning action. Finally, an assumption must be made regarding the trucking hours needed to drive to the next location. The assumptions in this research are made in consultation with the planners at the TCP (the planning department responsible for truck planning) and were later also checked with an industry partner; Den Hartogh. Particularly, the following four assumptions were derived:

Theoretical assumptions

(1)	*Trucking hours required for pick-up action*	45 minutes
(2)	*Trucking hours required for drop action*	45 minutes
(3)	*Trucking hours required for cleaning action*	60 minutes
(4)	*Trucking hours required for driving between locations*	$\frac{Total\ distance\ between\ consequtive\ locations}{60\ km/h}$

5.2.2 Estimation of the historical trucking capacity

Integrating the actual data and the theoretical assumptions, it is now possible to estimate the historical trucking capacity at a certain region at a given time. All trucking actions are known and the actual data for the loading and delivery actions can be retrieved from the data warehouse. Combining this with the theoretical assumptions as mentioned earlier, the historical trucking hours in a certain region can be estimated. Figure 5.5 presents an example of the estimated trucking hours of 2018, aggregated on a weekly level, for the GBN region. A distinction is made between H&S' own trucks and external charters.

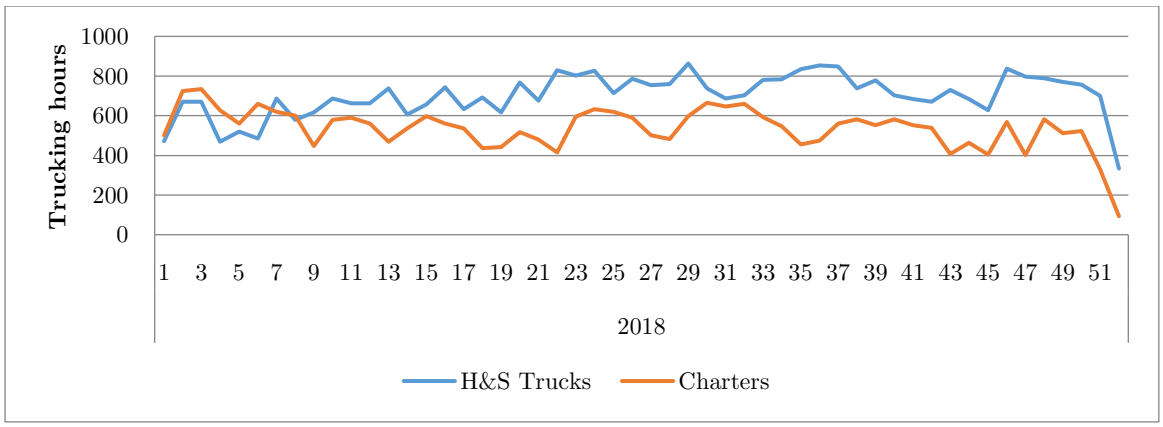


Figure 5.5 Estimated trucking hours for the GBN region aggregated on a weekly level

The historical trucking capacity can also be expressed in trucking units (rather than trucking hours). To do this, information is needed on how many hours a truck is productive per day. Table 5.1 and 5.2 show the average daily productivity of trucks in the BFN and GBN region. Unsurprisingly, the daily productivity of their own trucks is higher than those of charters. It makes sense from an economic point of view to first fully utilize the trucks in your own fleet before utilizing charters. Dividing the trucking capacity in hours by the average daily productivity yields an estimation of the required trucking capacity expressed in trucking units.

Table 5.1 Average daily productivity of trucks in BFN region

	Average daily production during working days	Average daily production during weekends
Own trucks	12.60	7.65
Charters	12.30	5.95
Total	12.54	7.49

Table 5.2 Average daily productivity of trucks in GBN region

	Average daily production during working days	Average daily production during weekends
Own trucks	12.08	8.90
Charters	10.05	6.13
Total	11.69	7.89

Since the forecasted loadings and deliveries are used as a foundation to predict the required trucking capacity, it is interesting to investigate the correlation between the sum of the loadings and deliveries and the historical trucking capacity. Hence, a scatterplot was made that shows the relationships between the total number of loadings and deliveries and the required trucking capacity in hours (Figure 5.6 and 5.8). From these figures it seems reasonable to conclude that there exists a very strong correlation between the daily number of loadings and deliveries and the daily (estimated) trucking capacity. This is further confirmed by computing the Pearson correlation coefficient. This coefficient quantifies the degree to which a relationship between two variables can be described by a line. In other words, it measures the linear association between variables

(Rodgers & Nicewander, 1988). Analysis showed that the Pearson correlation coefficient between the daily trucking capacity in hours and the sum of the number of loadings and deliveries is 0.96 and 0.95 for the BFN and GBN region, respectively. Note that interest in this study lies on forecasting the sub-daily trucking capacity (i.e. for each day, the trucking capacity is forecasted from midnight to noon (AM) and noon to midnight (PM)). It might therefore also be valuable to assess the correlation between the sub-daily number loadings and the deliveries and the sub-daily (estimated) trucking capacity. In Figure 5.7 and 5.9, it can be observed that although the correlation is still strong, it is not as strong as the correlation on a daily aggregation level, especially for the GBN region. This is also what would be expected, since the trucking capacity in AM depends for a significant extent also on the loadings and deliveries on PM and vice versa. It can be concluded that for the GBN region, this is even more so, since the correlation coefficient drops from 0.95 to 0.46.

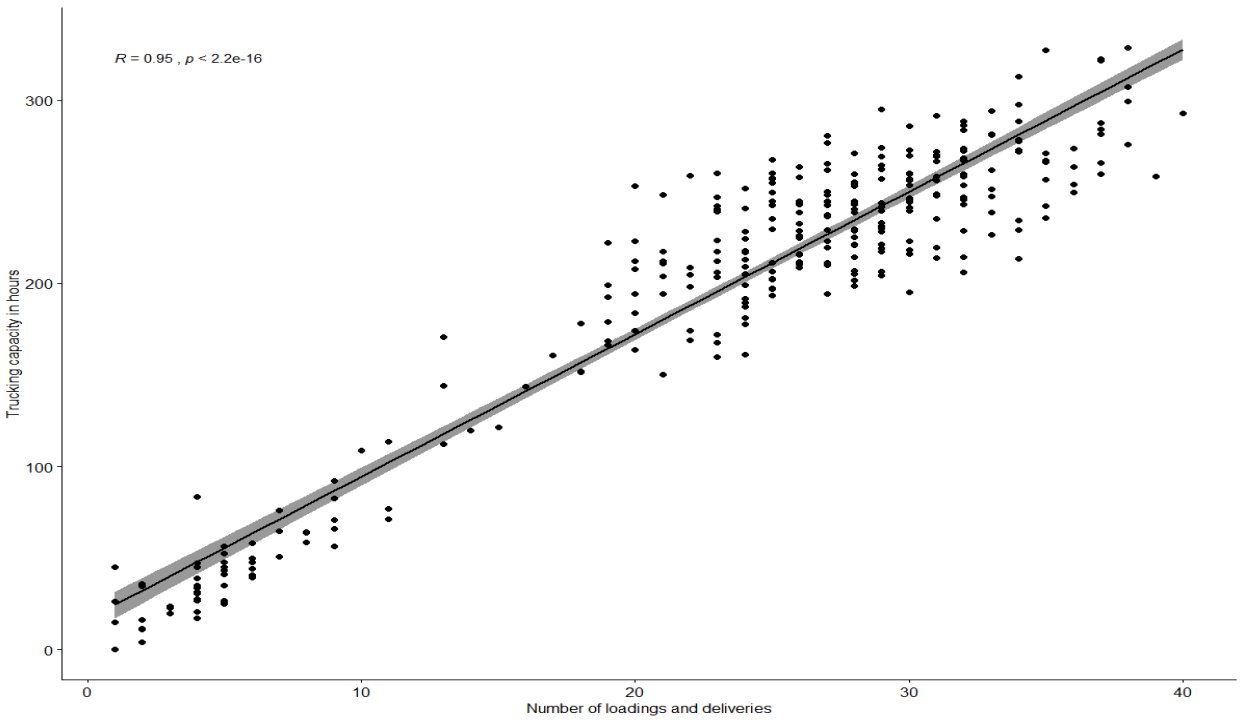


Figure 5.6 Relationship between daily loadings & deliveries and daily trucking capacity GBN region

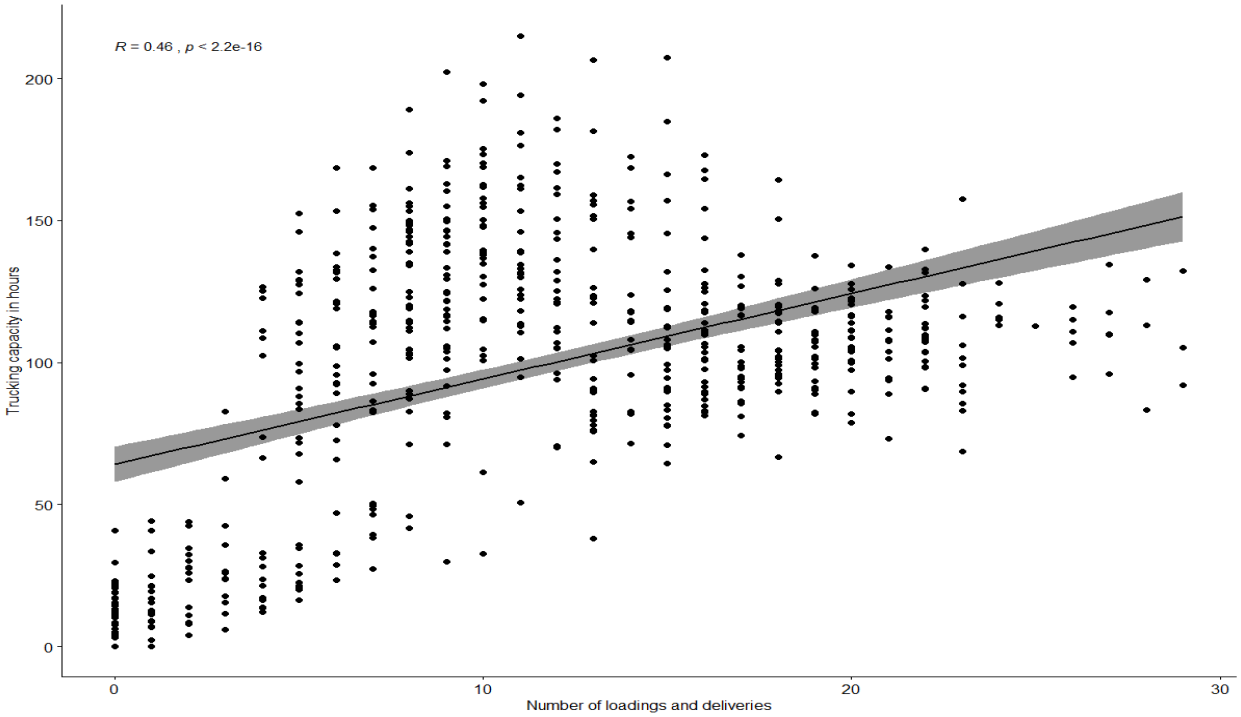


Figure 5.7 Relationship between sub-daily loadings & deliveries and sub-daily trucking capacity GBN

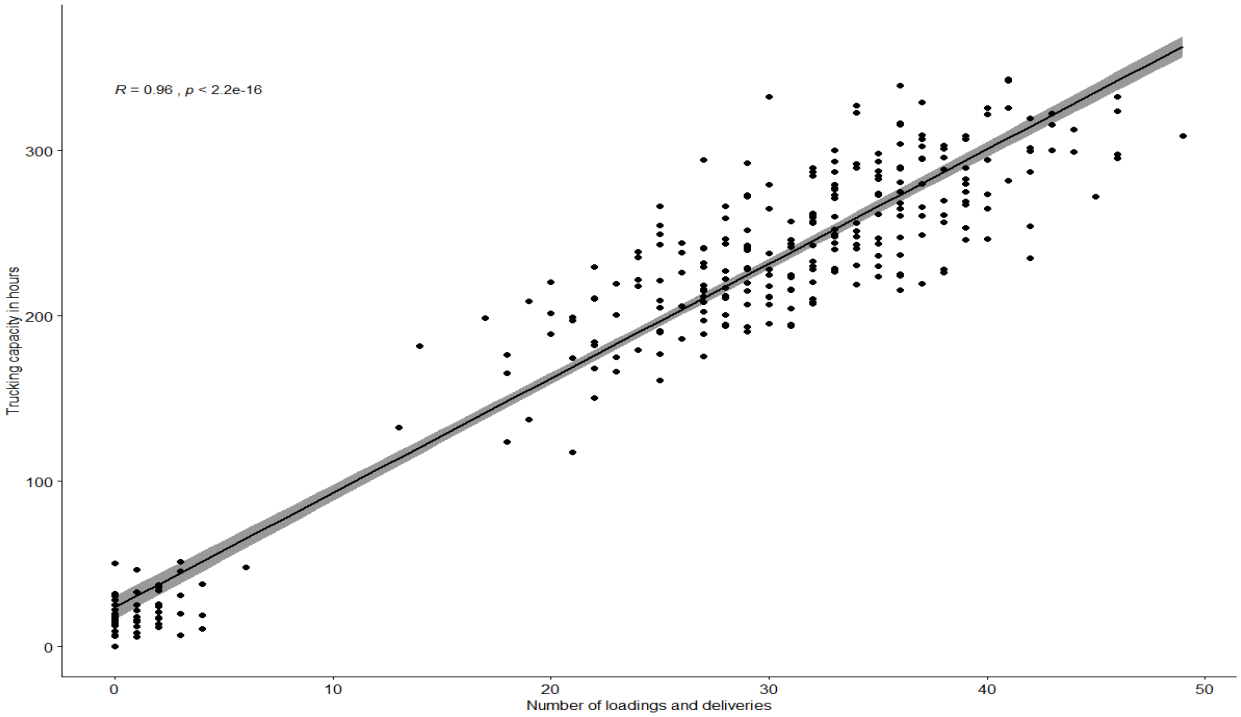


Figure 5.8 Relationship between daily loadings & deliveries and daily trucking capacity BFN region

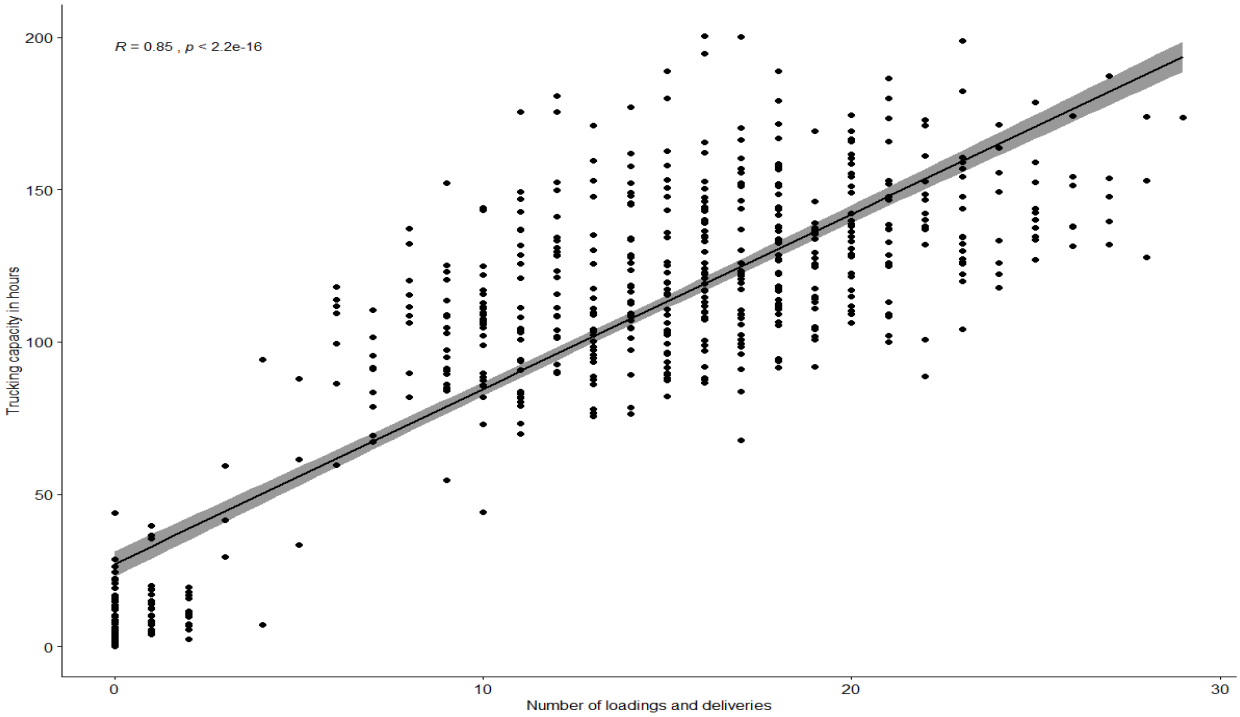


Figure 5.9 Relationship between sub-daily loadings & deliveries and sub-daily trucking capacity BFN region

5.3 Forecasting trucking capacity

5.3.1 Model development

Recall that in Chapter 2 and 3, models were developed to generate a forecast for the expected number of loadings and deliveries. In the previous section, the historical trucking capacity was estimated using both actual data as well as a number of assumptions. Moreover, it was also shown that the trucking capacity is closely related to the number of loadings and deliveries. This relationship is now exploited in order to predict the future trucking need. More specifically, a multiple linear regression model is developed in which the sub-daily trucking capacity in hours is the variable to be forecasted. To avoid confusion with earlier (dependent) variables, the forecasted trucking capacity for period τ as of time T is denoted by $\delta_T(\tau)$. Since the aim is to generate a sub-daily forecast for one week ahead, the time bucket length of τ is 0.5 days. The most important predictor variables in the model are the loadings and deliveries. Since a distinction needs to be made regarding the number of loadings and deliveries that occur in period τ and the number of loadings that do not occur in period τ but do still occur on the same day as period τ, some new notation is introduced. First, let $D(\tau)$ denote the specific day in which period τ falls. Additionally, let $Lo_{D(\tau)}$ and $De_{D(\tau)}$ denote the number of loadings and deliveries that do occur on $D(\tau)$, but not in period τ. Note that when a forecast is going to be generated by this multiple regression model, the loadings and deliveries for future time periods are not known. Hence, the forecast for the number of loadings and deliveries (obtained in Chapter 2 and 3) is used instead.

To get acquainted with this new terminology, consider the following example. On a given Monday, one wants to generate a forecast for the trucking capacity on Wednesday AM ($\tau = 3$) and PM ($\tau = 4$). Furthermore, suppose that the forecast for the expected number of loadings on Wednesday AM and Wednesday PM is 20 and 15, respectively. Recall that this forecast is obtained by the Bayesian algorithm covered in Chapter 3. With this information in mind, it is not particularly difficult to grasp how the new notation works. First, it can be observed that $D(3) = D(4)$, since both denote the Wednesday in the example. Secondly, $Lo_{D(3)}$ is equal to 15 and refers to the forecasted number of loadings for Wednesday PM. In a similar vein, $Lo_{D(4)}$ is

equal to 20 and refers to the forecasted number of loadings for Wednesday AM.

In addition to the abovementioned explanatory variables, six dummy variables are included, representing the day of the week. Moreover, two supplementary dummy variables are added to account for European wide holiday effects such as Christmas and Easter and the part of the day (i.e. AM or PM). Finally, the model is completed by including two lagged variables, representing the trucking capacity of one week and two weeks in the past. As such, the model can be summarized as:

$$\delta_T(\tau) = \beta_0 + \beta_1 * Lo(\tau) + \beta_2 * De(\tau) + \beta_3 * Lo_{D(\tau)} + \beta_4 * De_{D(\tau)} + \beta_5 * \delta(\tau - 14) + \beta_6 * \delta(\tau - 28) + \sum_{k=1}^{8} \beta_{6+k} x_k(\tau) + \varepsilon_T(\tau)$$

Equation 5.1

Where,

$\delta_T(\tau)$	the forecasted trucking capacity in hours for period τ as of time T
$Lo(\tau)$	the number of loadings during period τ
$De(\tau)$	the number of deliveries during period τ
$Lo_{D(\tau)}$	the number of loadings during $D(\tau)$, but not in period τ
$De_{D(\tau)}$	the number of deliveries during $D(\tau)$, but not in period τ
$\delta(\tau - 14)$	the actual trucking capacity in hours in period $\tau - 14$ (i.e. one week ago)
$\delta(\tau - 28)$	the actual trucking capacity in hours in period $\tau - 28$ (i.e. two weeks ago)
$x_1(\tau)$	1 if $D(\tau)$ is a Monday, 0 otherwise
$x_2(\tau)$	1 if $D(\tau)$ is a Tuesday, 0 otherwise
$x_3(\tau)$	1 if $D(\tau)$ is a Wednesday, 0 otherwise
$x_4(\tau)$	1 if $D(\tau)$ is a Thursday, 0 otherwise
$x_5(\tau)$	1 if $D(\tau)$ is a Friday, 0 otherwise
$x_6(\tau)$	1 if $D(\tau)$ is a Saturday, 0 otherwise
$x_7(\tau)$	1 if $D(\tau)$ is a holiday, 0 otherwise
$x_8(\tau)$	1 if period τ falls within AM, 0 otherwise
$\varepsilon_T(\tau)$	error term

Additionally, the coefficients ($\beta_1, \beta_2, \dots, \beta_{14}$) represent the effect of each predictor after taking into account the effects of all the other predictors in the model. In other words, the coefficients measure the marginal effects of the predictor variables. Finally, the intercept, β_0, denotes the predicted value of the trucking capacity when all predictor variables are equal to zero.

5.3.2 Theoretical assumptions

In (multiple) linear regression, one implicitly makes a number of assumptions regarding Equation 5.1. Firstly, it is assumed that the relationship between trucking capacity and the predictor variables satisfy a linear relationship. That is, it is assumed that the model is a reasonable approximation to reality. Secondly,

four assumptions are made regarding the errors of the model (Hyndman & Athanasopoulos, 2018):

- The errors have mean zero; otherwise the forecasts are systematically biased
- The errors are uncorrelated to the predictor variables; otherwise there is more information that should be included in the systematic part of the model
- The errors are normally distributed
- The residuals should be homoscedastic; otherwise p-values are generally underestimated, and coefficient estimates are less precise

5.3.3 Fitting the initial model to the data

The regression coefficients in Equation 5.1 are estimated using least square estimation. The least squares principle provides a way of estimating the coefficients effectively by minimizing the sum of the squared errors. This is called least squares estimation because it gives the least value for the sum of squared errors. To fit the model, the data of 2018 is used as training set. Table 5.3 presents information about the fitted models for both the BFN as well as the GBN region. Particularly, it shows the estimated coefficients for each model, accompanied with its standard error and p-value. The standard error can be interpreted as the standard deviation which would be obtained from repeatedly estimating the coefficients on similar data sets (Hyndman & Athanasopoulos, 2018). That is, it provides a measure of uncertainty in the estimated coefficient. Moreover, the p-value represents the probability that the estimated coefficient would be as large as it is if there was no relationship between trucking capacity and the estimated coefficient. To assess how well the regression models fit the data, the coefficient of determination (R^2) is computed. This goodness of fit statistic represents the square of the correlation between the observed values of the trucking capacity and the predicted values. The R^2 for the BFN region and the GBN region equals 0.92 and 0.89, respectively. The closer the predicted values are to the actual values, the closer R^2 is to 1. It therefore seems reasonable to conclude that the models provide a good fit. Moreover, it can be observed that although for the BFN region all coefficients are significant, this is not the case for the GBN region. More specifically, it seems that certain dummy variables, representing day of the week and holiday effects, appear to be insignificant for this region.

Table 5.3 Estimation coefficients BFN and GBN regions

BFN region	Corresponding explanatory variable	Coefficient	Standard error	p-value	GBN region	Coefficient	Standard error	p-value
β_0	Intercept	4.458	1.713	**0.009**	β_0	9.530	2.063	**< 0,0001**
β_1	$Lo(\tau)$	4.143	0.199	**< 0,0001**	β_1	1.070	0.344	**0.002**
β_2	$De(\tau)$	2.972	0.293	**< 0,0001**	β_2	2.978	0.168	**< 0,0001**
β_3	$Lo_{D(\tau)}$	1.116	0.198	**< 0,0001**	β_3	1.889	0.341	**< 0,0001**
β_4	$De_{D(\tau)}$	1.962	0.294	**< 0,0001**	β_4	3.048	0.174	**< 0,0001**
β_5	$\delta(\tau - 14)$	0.074	0.022	**0.001**	β_5	0.129	0.026	**< 0,0001**
β_6	$\delta(\tau - 28)$	0.045	0.022	**0.045**	β_6	0.121	0.026	**< 0,0001**
β_7	$x_1(\tau)$	30.064	4.741	**< 0,0001**	β_7	16.728	4.124	**< 0,0001**
β_8	$x_2(\tau)$	32.892	5.147	**< 0,0001**	β_8	4.161	4.322	0.336
β_9	$x_3(\tau)$	26.119	4.925	**< 0,0001**	β_9	4.795	4.432	0.280
β_{10}	$x_4(\tau)$	27.316	5.179	**< 0,0001**	β_{10}	8.463	4.211	**0.045**
β_{11}	$x_5(\tau)$	21.355	4.817	**< 0,0001**	β_{11}	6.038	3.671	0.100
β_{12}	$x_6(\tau)$	6.468	2.180	**0.003**	β_{12}	-0.222	2.050	0.914
β_{13}	$x_7(\tau)$	-35.934	5.941	**< 0,0001**	β_{13}	-8.535	4.644	**0.036**
β_{14}	$x_8(\tau)$	-8.536	1.524	**< 0,0001**	β_{14}	-8.904	1.892	**< 0,0001**

5.3.4 Testing the theoretical assumptions

Recall that in an earlier section, four assumptions were made regarding the residuals of the multiple linear regression model. The first assumption stated that the residuals should have mean zero, and this assumption is satisfied by both models. The second assumption required that the residuals are uncorrelated to the predictor variables. This assumption is also satisfied for both models (see Appendix D). Third, it was assumed that the errors are normally distributed. This assumption is satisfied as well by virtue of the Central Limit Theorem. The final assumption specified that the variance of each residual, conditional on the chosen values of the explanatory variables, should be a constant. In other words, it was assumed that the residuals should be homoscedastic. To check for heteroscedasticity, the predicted trucking capacity is fitted against the residuals (Figure 5.10 and 5.11). From these plots it can be observed that for both regions it seems to be the case that, as the fitted value increases, the variance of the residuals also increases. Hence, it seems reasonable to assume that the residuals of both models are heteroskedastic. To confirm the suspicion for heteroscedasticity, a white test was performed. The p-value of the white test for both models was < 0.001. From this can be concluded that there is significant evidence to reject the null hypothesis and conclude that the residuals of both models are indeed heteroskedastic. A closer look at the residuals plot shows that there are essentially two groups of clustered residuals. For the first group, the predicted trucking capacity (i.e. the fitted values) is lower than for the second group of clustered residuals. This can intuitively be explained by the fact that both weekends and weekdays are included in the model. Since the required trucking capacity during weekends is (especially in the BFN region) significantly less than the required capacity during working days, it was to be expected that this pattern would show in Figure 5.10 and 5.11. Since the required trucking capacity is higher during working days, the variance is also higher. It was explained earlier that when the assumption of homoscedasticity is violated, the usual ordinary least square regression coefficients become less efficient and the standard errors tend to be biased. Hence, in the next section, an attempt is made to resolve the presence of unequal variances, by performing a variance-stabilizing Box-Cox transformation.

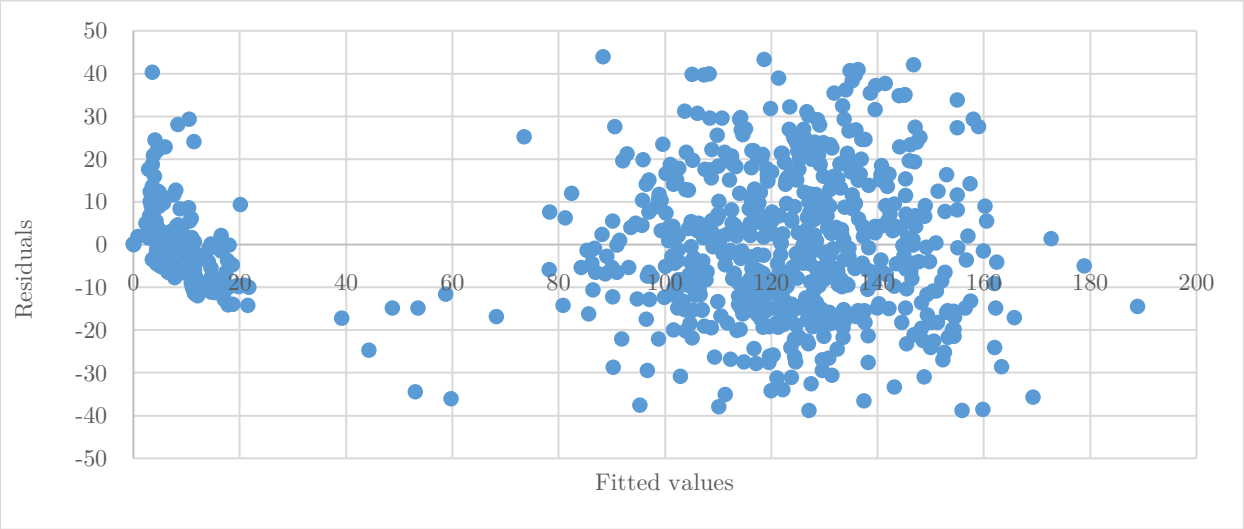


Figure 5.10 Fitted values vs. Residuals plot BFN region

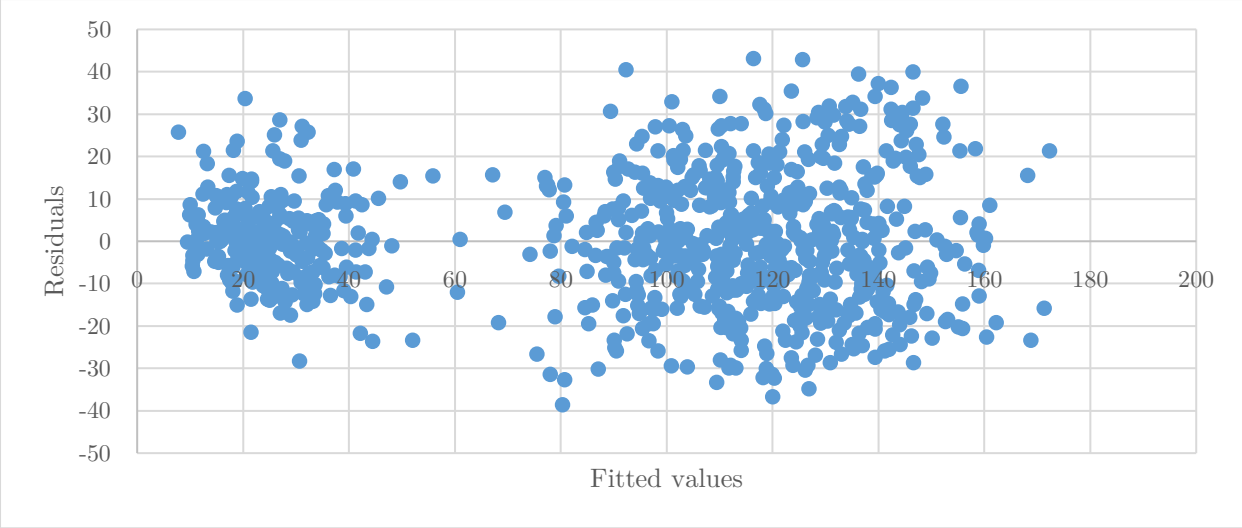


Figure 5.11 Fitted values vs. Residuals plot GBN region

5.3.5 Variance stabilizing transformation to resolve heteroscedasticity

In the previous section, the presence of heteroscedasticity of the residuals was diagnosed for both regression models. In this section, a Box-Cox transformation on the response variable is conducted as a corrective measure for the presence of unequal variance. At the core of the Box-Cox transformation is the power parameter lambda (λ). The value of λ is chosen, using maximum likelihood estimation to obtain the optimal λ that minimizes the sum of squares of error of the transformed data (Sakia,1992). Let the initial observations be denoted by $\delta_1, \dots \delta_T$ and the transformed observations by $w_1, \dots w_T$. The Box-Cox transformation can now be defined as (Sakia, 1992):

$$w_t = \begin{cases} \ln(\delta_t) & if\ \lambda = 0; \\ \dfrac{\delta_t^{\lambda} - 1}{\lambda} & otherwise. \end{cases}$$

Equation 5.2

In words, when the value of lambda equals zero, the Box-Cox transformation is always a natural logarithm. In contrast, when lambda takes any value not equal to zero, a power transformation is used, followed by some scaling (Hyndman & Athanasopoulos, 2018). Note that after the Box-Cox transformation, the data is no longer in their original scale. Consequently, after a forecast has been obtained for the transformed data, one needs to reverse the transformation to generate the forecasts on the original scale (Hyndman & Athanasopoulos, 2018). The reverse Box-Cox transformation is given by:

$$\delta_t = \begin{cases} \exp(w_t) & if\ \lambda = 0; \\ (\lambda w_t + 1)^{\frac{1}{\lambda}} & otherwise. \end{cases}$$

Equation 5.3

Equation 5.2 is now used to perform the Box-Cox transformation on the response variable (i.e. the trucking capacity). The optimal value of λ for the BFN and GBN model were 0.47 and 0.97, respectively. After performing the transformation and re-estimating the regression models, a white test confirmed that the heteroscedasticity was resolved. This can also be observed by visually inspecting the figures in Appendix D.

5.3.6 Re-estimating the models and evaluating the overall model fit

In the previous section it was shown the heteroscedasticity of the residuals was resolved by performing a Box-Cox transformation. Table 5.5 presents information about the re-estimated models for both the BFN as well as the GBN region. Besides resolving the heteroscedasticity, the coefficient of determination (R^2) also

improved for both models. Particularly, it increased from 0.92 to 0.95 and from 0.89 to 0.90 for the BFN and GBN region, respectively. Furthermore, the AIC value also improved for both models, going from 4,842 to 1,029 and from 4,723 to 4,149, for the BFN and GBN region, respectively. If one considers the p-values of the estimated coefficients for the BFN model, it can be concluded that all coefficients are significant, apart from β_6, which represents the lagged value of the trucking capacity 28 periods (i.e. two weeks) in the past. Similarly, for the GBN region, all coefficients are significant, except for β_8, β_9, and β_{12}, representing the dummy variables for Tuesday, Wednesday, and Saturday, respectively. Although the abovementioned variables are insignificant, they will not be excluded from the models. The intuition behind this is twofold. To start with, statistical significance does not always indicate predictive value. More importantly, eventually H&S Foodtrans wants to generate forecasts for all their regions (not only for the BFN and GBN regions). It is therefore more convenient and intuitive to have the same explanatory variables for each region.

Table 5.4 Re-estimated coefficients BFN and GBN region

BFN region	Corresponding explanatory variable	Coefficient	Standard error	p-value	GBN region	Coefficient	Standard error	p-value
β_0	Intercept	-0.345	0.185	0.063	β_0	9.158	1.514	**< 0,0001**
β_1	$Lo(\tau)$	0.358	0.021	**< 0,0001**	β_1	0.822	0.252	**0.001**
β_2	$De(\tau)$	0.280	0.032	**< 0,0001**	β_2	2.197	0.124	**< 0,0001**
β_3	$Lo_{D(\tau)}$	0.116	0.021	**< 0,0001**	β_3	1.415	0.250	**< 0,0001**
β_4	$De_{D(\tau)}$	0.186	0.032	**< 0,0001**	β_4	2.220	0.127	**< 0,0001**
β_5	$\delta(\tau - 14)$	0.007	0.002	**0.003**	β_5	0.094	0.019	**< 0,0001**
β_6	$\delta(\tau - 28)$	0.004	0.002	0.113	β_6	0.088	0.019	**< 0,0001**
β_7	$x_1(\tau)$	10.416	0.512	**< 0,0001**	β_7	14.242	3.027	**< 0,0001**
β_8	$x_2(\tau)$	10.532	0.555	**< 0,0001**	β_8	5.070	3.172	0.110
β_9	$x_3(\tau)$	9.956	0.531	**< 0,0001**	β_9	5.512	3.253	0.091
β_{10}	$x_4(\tau)$	9.989	0.559	**< 0,0001**	β_{10}	8.235	3.091	**0.008**
β_{11}	$x_5(\tau)$	9.659	0.520	**< 0,0001**	β_{11}	6.546	2.695	**0.015**
β_{12}	$x_6(\tau)$	3.924	0.235	**< 0,0001**	β_{12}	-0.154	1.505	0.919
β_{13}	$x_7(\tau)$	-7.254	0.641	**< 0,0001**	β_{13}	-6.230	3.409	**0.048**
β_{14}	$x_8(\tau)$	-0.636	0.164	**< 0,0001**	β_{14}	-6.730	1.389	**< 0,0001**

5.3.7 Summary of the complete forecasting method for predicting trucking capacity

Recall that with regards to trucking capacity, the goal of this research was to design a forecasting method that predicts the sub-daily amount of trucking hours needed for one week ahead. At this point, all separate steps to generate such a forecast have been discussed. Particularly, Chapter 2 discussed in detail how an initial forecast of the expected number of loadings can be generated. Thereupon, Chapter 3 examined how this initial forecast can be adjusted based on the advance demand information. Finally, this chapter has shown how the adjusted forecasts for the expected number of loadings and deliveries are used as explanatory variables in a multiple linear regression model. In the remainder of this section, the integrated forecasting method of how to obtain the final forecast for the required trucking capacity in a given region is summarized:

Step 1: Generate the initial τ step ahead forecast for the expected number of loadings $\forall \boldsymbol{\tau} \in \{\mathbf{1}, \mathbf{2}, \dots, \mathbf{14}\}$

In the first step, the initial (one week ahead) forecast is generated for the sub-daily number of loadings and deliveries. Note that to generate theses initial forecasts, all models developed in Chapter 2 can be employed. Which forecasting model is actually used in this step, depends on the preferences of the practitioners at H&S Foodtrans. It was shown that the accuracy of the various models in this step were relatively similar, whereas they did differ substantially in the level of complexity. Hence, for simplicity reasons, it can be decided to use the simple mean method for every region in this step. Another possibility is to select the best forecasting

model per series, independent of the level of complexity.

Step 2: Adjust the initial forecasts of the loadings and deliveries by applying the Bayesian technique $\forall \boldsymbol{\tau} \in \{\mathbf{1}, \mathbf{2}, \dots, \mathbf{14}\}$

In the second step, the initial forecasts as obtained in step 1 are adjusted to incorporate the advance demand information. The Bayesian estimate of the expected number of loadings and deliveries is used to derive this adjusted forecast. Note that the Bayesian technique is applied separately to the forecasted loadings and the forecasted deliveries. The steps of the algorithm that should be followed to ultimately result in the expected number of loadings and deliveries for future period τ are outlined in Chapter 3. Also note that this algorithm is followed separately for all $\tau \in \{1,2, \dots, 14\}$. The output of this step is the adjusted one week ahead forecast for the sub-daily number of loadings and deliveries in a given region. This adjusted forecast for the expected loadings and deliveries, at the most current time period T for τ time units in the future, is denoted by $\tilde{Y}_T^{Lo}(\tau)$ and $\tilde{Y}_T^{De}(\tau)$, respectively.

Step 3: Run the multiple linear regression model to generate the forecast the required trucking capacity

In the final step, the forecasted number of loadings and deliveries (derived in step 2) are used as explanatory variables in a multiple linear regression model (Equation 5.1). In particular, the first two explanatory variables represent the forecasted loadings $(Lo(\tau))$ and deliveries $(De(\tau))$ in period τ. Furthermore, the consecutive two variables refer to the forecasted loadings $(Lo_{D(\tau)})$ and deliveries $(De_{D(\tau)})$ that are expected to occur on day $D(\tau)$, but not during period τ. In addition to the forecasted loadings and deliveries, the regression model also encompasses the following independent variables; (1) lagged values (i.e. the trucking capacity one week and two weeks ago), and (2) dummy variables representing day of the week, part of the day, and holiday effects.

5.3.8 Real world application and results

In order to assess the accuracy of the total forecasting model as summarized in the previous section, the model is tested on the same test set which was used before to evaluate the accuracy of the Bayesian technique (see Section 3.4.). In short, the test set contains the data of the BFN region ranging from 27/04/2019 to 25/05/2019. The accuracy of the forecasting model proposed by this research is contrasted with a benchmark method that is currently being employed in certain regions within H&S Foodtrans (although not in a systematic manner). This benchmark model deeply resembles the "simple mean method" that was discussed earlier in the context of forecasting loadings and deliveries. In particular, the benchmark model simply computes the average historical trucking capacity over the last three months for a certain period, and this number is then used as the prediction for future periods. For example, the average hours of trucking capacity for Monday AM over the last three months is calculated and constitutes the forecast for the next occurrence of Monday AM. Formally, the benchmark model can be written as:

$$\delta_T(\tau) = \frac{\sum_{k=1}^{12} \delta_T(\tau - (14 * k))}{12} \qquad \textit{Equation 5.4}$$

Table 5.6 presents the accuracy of both the proposed forecasting model as well as the benchmark model. Note that he trucking capacity differs substantially for working days compared to weekends, especially in the BFN region. For this reason, the accuracy of predicting the trucking capacity also differs substantially for working days as compared to weekends. Consequently, it makes more sense to consider the accuracy of working days and weekends separately, as depicted in Table 5.6. Starting with the most astonishing conclusion, it can be observed that the MAE and MAPE of the benchmark model during workings days were 34.43 and 85%, respectively. In contrast, the MAE and MAPE of the proposed forecasting model during working days were 20.39 and 26%, respectively. Hence, on average, the absolute error of the benchmark model during working days in the BFN region is 34.43 trucking hours. The absolute error of the proposed

forecasting model, on the other hand is only 20.39 trucking hours, improving on the benchmark model by 41%. A second observation is that the MAPE during weekends is relatively high (110% for the proposed model). As already mentioned, this is mostly due to the fact that the trucking capacity in the BFN region is very low during the weekends and therefore the error is percentage wise relatively high. This is not troublesome, however, since the error in absolute terms is only 2.58 hours. If one recalls that the average daily production of trucks during weekends in the BFN region is 7.49, this means that the error expressed in number of trucks is only 0.34. In a similar way, the MAE during working days can be expressed in number of trucks. Recall that the average daily production of trucks during working days in the BFN region is 12.54. As such, the MAE expressed in number of trucks is 1.62.

Table 5.5 Accuracy of proposed forecasting model vs. benchmark

	MAE working days (in hours)	MAPE working days	MAE weekends (in hours)	MAPE weekends
Proposed Forecasting Model	20.39	26%	2.58	110%
Benchmark Model	34.43	85%	2.76	5128%

5.4 Synthesis of findings regarding forecasting trucking capacity

In order to be able to plan their drayage operation in a more proactive fashion, H&S is in need for a forecasting model that predicts the number of loadings and deliveries and the corresponding trucking capacity. This chapter attempted to develop a forecasting model that could serve this purpose. As a first step, the historical trucking capacity per region was estimated based on actual data as well as theoretical assumptions. Thereupon, a multiple linear regression model was developed that uses the forecasted loadings and deliveries and other explanatory variables to predict the required trucking capacity. After applying the proposed model to a real-world test case and comparing the accuracy with the current forecasting approach occasionally used by H&S, the conclusion was drawn that the proposed model significantly outperforms the benchmark model. Particularly, the MAE during working days improved by 41%.

6 Connecting the dots

Proactive planning of drayage operations and efficient empty tank container repositioning are vital factors in achieving the operational excellence necessary to survive in the low-margin industry of intermodal logistics service providers. This research aimed to contribute to this problem area by proposing a forecasting methodology that predicts the number of loadings and deliveries and the corresponding required tank container and trucking capacity. This chapter circles back to these two operational challenges faced by H&S Foodtrans (and intermodal LSPs in general) and particularize (1) how the proposed forecasting methodology might help in addressing these issues and (2) the magnitude of savings that can be expected. Finally, since the forecasting methodology proposed in this research is currently being implemented at H&S and Den Hartogh, the final section is devoted to providing more details on this project.

6.1 Forecasting and the proactive planning of drayage operations

The current planning of the drayage operation within H&S Foodtrans occurs in a relatively ad hoc fashion. Due to their busy schedules, planners are often too carried away by the everyday humdrum to be engaged with estimating the number of trucking hours needed for tomorrow, let alone further in the future. This often results in last-minute charter requests which are undesirable for at least three reasons. First, attempts to book charters at the very last moment might fail due to the fact that no charters are available anymore. As a consequence, certain orders might have to be cancelled or delayed, reducing the on-time performance of H&S towards their clients. Second, the quality in terms of the drivers' qualifications and capabilities is lower for charters that are still available on such a short notice (i.e. the high performing charters are fully booked at an earlier stage). Third, it might sometimes also occur that a premium is asked by charters for last-minute requests. In addition to booking charters longer in advance, the proactive planning of drayage operations also includes smoothing the workload (i.e. the number of loadings and deliveries) throughout the day and week. A more equally divided workload has the potential to substantially reduce the costs associated with drayage operations as it increases utilization of own trucks and decreases the total number of trucks needed since planners can make more optimal combinations.

Firstly, regarding the wish to book charters longer in advance, this research contributes by presenting a forecast methodology that predicts the sub-daily required trucking capacity for one week ahead. This forecast can be used to gain insight into the expected trucking hours needed for the coming week. By using this information, charters can be booked longer in advance, thereby increasing the on-time delivery towards clients and reducing costs. The proposed forecasting methodology for predicting the required trucking capacity was tested in the BFN region. As discussed in greater detail in section 5.3.8, the mean absolute error during working days was shown to be 20.39 hours, improving the current forecast used by H&S Foodtrans by 41%. This level of accuracy indicates that charters can be booked at an earlier stage than the current practice.

The second component of planning the drayage operation more proactively involves smoothing the workload throughout the day and week, avoiding high peaks and deep troughs in the workload. This can partly be assessed by actively searching for more work for time periods in which (based on the forecast) a trough in workload is expected. In practice, this "searching for work" means consulting other LSPs, discussing if they might be interested in selling certain orders. At present, this happens occasionally, but this pursuit might be intensified now that H&S has access to a more accurate forecast. Moreover, this forecast can also be used by account managers to create a more balanced workload. In practice this would involve a change in the business process flow at H&S Foodtrans. The current business process flow works as follows. First, the

account manager receives an order from a client. Thereupon, the transport management software performs some technical checks on the order (e.g. is the contract of the client still valid). When the order is accepted by the software, it is transferred to the planning department. Next, the planning department plans and schedules the orders in such way that goods are delivered at the right time and place. More specifically, the planning department assigns a tank container and trucking unit to an order. This planning process is performed in a sequential manner. First a tank container is assigned to an order and only when this has been done, the trucking unit is scheduled. The part of the planning department that assigns the tank container is called the Multi Material Planning (MMP), and the part that plans the trucking unit is referred to as the Truck Capacity Planning (TCP). The alteration in the current business process flow that needs to be made to proactively smooth the workload changes the role of account managers. Instead of unconditionally accepting all order from clients, their new role involves actively balancing the orders throughout the day and week. In order to flourish in their new role, account managers require a forecast for the expected number of loadings and deliveries and corresponding trucking capacity. To illustrate, consider the situation in which on Monday a client places an order for Friday AM. Instead of simply accepting this order, the account manager now first considers the operational forecast of the number of loadings and deliveries and the corresponding trucking capacity. Assume, for the sake of this example, that according to this forecast, the number of loadings and deliveries and corresponding trucking capacity is significantly higher at Friday AM than PM. In the light of this information, the account manager might now ask the client whether the order might also be carried out on Friday PM instead of AM. This alteration in the business process flow of H&S Foodtrans is visualized in Figure 6.1 and 6.2.

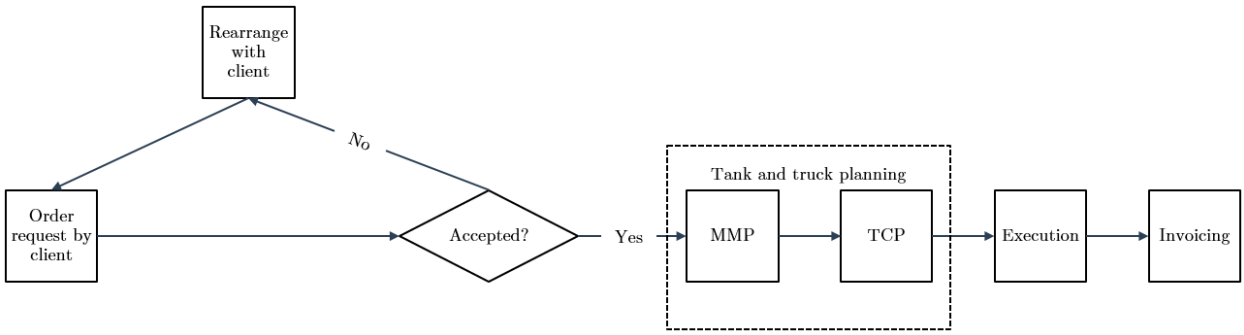


Figure 6.1 Current business process flow at H&S Foodtrans

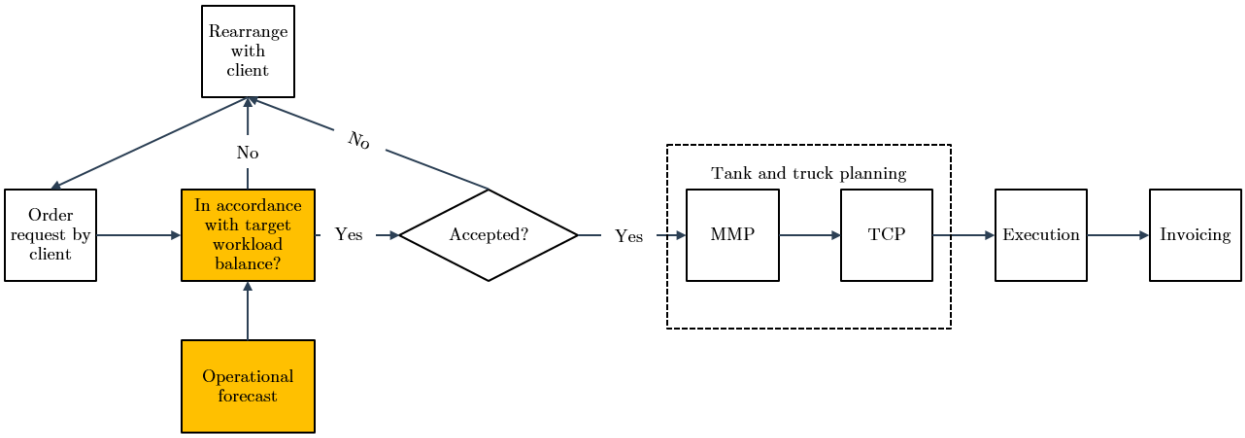


Figure 6.2 New business process flow at H&S Foodtrans

Above it has been argued that a more balanced workload has the potential to significantly reduce the costs of drayage operations. To quantify the potential savings that might be achieved by creating a more balanced workload, an attempt has been made to quantify the potential savings related to a better balance throughout the day.

Recall that one of the major benefits of having a more balanced workload includes the expected decrease of the total number of trucks needed since planners can make optimal combinations. Consider, for instance, the situation in which there are 20 loadings and deliveries during both AM and PM as opposed to the situation in which there are 35 loadings and deliveries during AM and only 5 during PM. Although the total number of loadings and deliveries is the same in both situations, it is expected that less trucks are needed in the first situation as more efficient combinations can be made. This observation is confirmed by Figure 6.3. This Figure presents a scatter plot of the daily balance between AM and PM and the number of trucks per action that were used. Note that an action in this context denotes either a loading or a delivery. The first important conclusion that can be drawn on the basis of this figure is that in AM more loadings and deliveries are performed than in PM. Moreover, as expected, it seems that there exists a significant positive correlation between the imbalance between AM and PM and the number of trucks per action (i.e. Pearson correlation coefficient and corresponding p-value are equal to 0.41 and 8.924e-12, respectively). Note however, that this figure includes all the planning regions at H&S Foodtrans. It might very well be possible that specific regions show a different pattern. Particularly, if there are regions that also experience the inverted situation in which the percentage of loadings and deliveries that are carried out on AM is significantly less than during PM, it is expected that the graph would be approximately u-shaped. That is, the number of trucks per action would, on average, be smallest if the number of loadings during AM is 50%. The further one moves away from this perfect balance in any direction, the number of trucks per action are expected to increase.

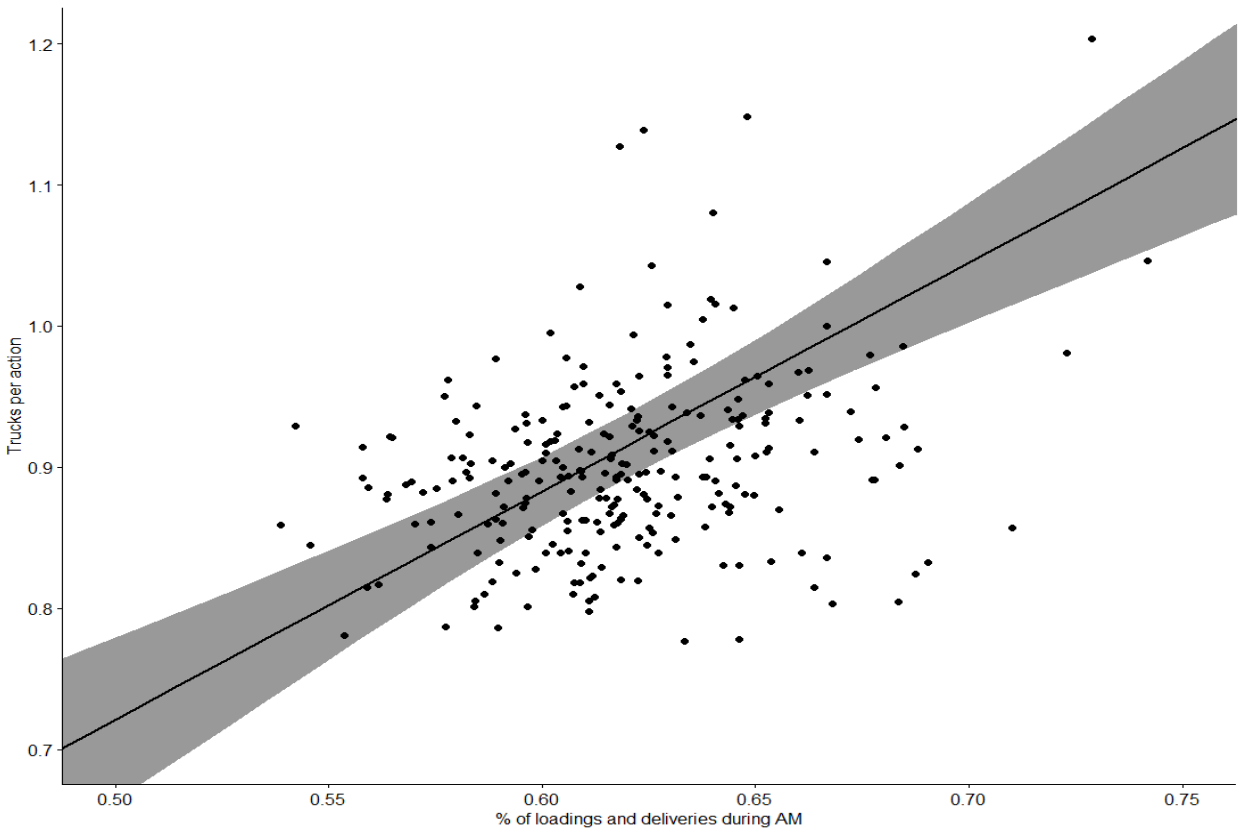


Figure 6.3 Relationship between the (im)balance throughout the day and the number of trucks per action

Considering H&S' European network as a whole, it seems to be the case that the greater the imbalance between AM and PM, the more trucks per action are needed. Put differently, a more balanced workload is associated with less trucks per action. This observation is used to estimate the potential savings from improving the workload balance. First, the slope of the regression line depicted in Figure 6.3 is computed and turns out to be 1.619. The interpretation of this slope is that a 1% decrease in the percentage of loadings and deliveries during AM is associated with a $1\% * 1.619 = 0.01619$ decrease in the number of trucks per action. In other words, if we were to improve the balance between the loadings and deliveries throughout the day, this would, on average, lead to a decrease of 0.001619 trucks per action. Similar calculations can be made for larger percentage improvements in the balance between the loadings and deliveries throughout the day (see column two in Table 6.1). Subsequently, if one multiplies the average decrease in trucks per action by the number of actions per year, the yearly savings in terms of trucks can be obtained (column three in Table 6.1). The decrease in the number of trucks needed per year can also be expressed in a percentage if one considers that the total number of trucks employed throughout the most recent year was equal to 56,000 (column six in Table 6.1). Next, the total costs related to trucking operations in the most recent year are known to be equal to € 27,000,000. Note that not all these costs are a function of the number of trucks used (e.g. fuel costs). For this reason, the conservative estimation, made in consultation with the TCP department, is that 30% of the total yearly trucking costs are directly related to the number of trucks used. These costs savings mainly originate from a reduction in labour costs since drivers are (partly) paid per day, irrespective of their productive hours. Furthermore, if systematically less trucks are needed, H&S might suffice with a smaller trucking fleet, reducing the total fixed costs related to their trucking fleet. As such, the total savings can be approximated by $30\% * €27{,}000{,}000 * \Delta\%\ \text{used trucks}$ (column nine in Table 6.1).

It is important to also emphasize the limitations of the analysis conducted above. The estimation should be considered as a guesstimate, presenting the reader with a rough idea of the magnitude of the potential savings. Further studies on the topic of quantifying the potential benefits of balancing the workload are therefore required to elucidate its potential. Although it should be emphasized that the numbers in Table 6.1 should be interpreted with caution, it seems reasonable to conclude that smoothing the workload throughout the day offers substantial potential for reducing the costs associated with the drayage operations of H&S Foodtrans. In order to be able to utilize this potential, it is essential to have an accurate prediction of the expected number of loadings and deliveries and corresponding trucking capacity in order to support account managers in balancing the workload. The proposed forecasting methodology developed in this study serves this purpose.

Table 6.1 Estimation of potential savings as a result of a more balanced workload throughout the day

% improvement in balance throughout the day	Savings in number of trucks per action	Number of actions per year	Savings in trucks per year	Total number of trucks used per year	% decrease in number of trucks used per year	Total costs associated with trucking operations	% of trucking costs dependent on number of trucks	Total estimated savings from improvement of balance
1%	0.016	60,000	972	56,000	-2%	€ 27,000,000	30%	€ 140,549
2%	0.032	60,000	1,943	56,000	-3%	€ 27,000,000	30%	€ 281,099
3%	0.049	60,000	2,915	56,000	-5%	€ 27,000,000	30%	€ 421,648
4%	0.065	60,000	3,887	56,000	-7%	€ 27,000,000	30%	€ 562,198
5%	0.081	60,000	4,859	56,000	-9%	€ 27,000,000	30%	€ 702,747
6%	0.097	60,000	5,830	56,000	-10%	€ 27,000,000	30%	€ 843,297
7%	0.113	60,000	6,802	56,000	-12%	€ 27,000,000	30%	€ 983,846
8%	0.130	60,000	7,774	56,000	-14%	€ 27,000,000	30%	€ 1,124,396
9%	0.146	60,000	8,745	56,000	-16%	€ 27,000,000	30%	€ 1,264,945
10%	0.162	60,000	9,717	56,000	-17%	€ 27,000,000	30%	€ 1,405,495

6.2 Forecasting and empty container repositioning

Having discussed the first operational challenge with regards to proactive planning of drayage operations, it is now time to turn to the second pressing challenge for LSPs as identified in this thesis. As aforementioned, this second challenge is concerned with efficient empty container repositioning in order to be able to fulfil future demand while minimizing empty container movements. Taking the imbalance in the container flow across regions as a given, the question arises what should be done with empty containers in a certain region to minimize empty container movements while still ensuring that demand can be met. To make efficient repositioning decisions, one needs insights regarding the number of loadings that are expected to occur in a given region at a given time. At present, all repositioning decisions are made on the basis of experience by two very knowledgeable and capable planners at the MMP department at H&S. Despite their competence, solely relying on the experience of these planners has at least two serious drawbacks. Firstly, they have to store a lot of information in their heads and regardless of their competence, they might sometimes still be biased. Secondly, relying exclusively on the experience of two planners is highly risky for a large firm as H&S. If one (let alone two) of these planners drop out for one reason or another, there is no alternative system on which repositioning decisions might be based, since there is currently no forecasting model at hand that predicts the expected required tank container capacity. What makes this issue especially important, are the significant costs that are associated with empty tank container repositioning. It was already mentioned earlier that for land transportation of containers, it is estimated that 40% up to 50% of all container movements are empty container movements (Branch, 2006; Konings & Thijs, 2001). To be more specific with respect to the case in point at H&S Foodtrans: the average costs per order related to empty container movements were approximated to be €350. With over 3,000 orders on an annual basis, this translates to yearly costs of €11,323,586.

The magnitude of these numbers stresses the enormous importance of having a forecast model in place to support (and if necessary temporarily replace) the two experienced planners at H&S Foodtrans. The proposed forecasting methodology for predicting the number of loadings and deliveries and corresponding tank container capacity could serve this purpose. The developed forecasting methodology was tested on the Rotterdam region and it was shown that the accuracy in terms of mean absolute error varied between 0.57 and 2.72, depending on the level of aggregation. In other words, in the Rotterdam region, the tank container capacity expressed in number of tank containers can be forecasted with a mean absolute error varying between 0.57 and 2.72 dependent on the type of tank container (i.e. level of aggregation). Taken together, the conclusion seems warranted that the proposed forecasting methodology provides a useful tool to support the experienced planners at H&S Foodtrans. Although it is difficult to precisely quantify the potential savings that can be brought about by the proposed forecast methodology, it seems self-evident that even a small improvement in empty tank container repositioning, and hence a reduction in empty container movements, can lead to substantial savings due to the magnitude of costs involved with empty container repositioning. Furthermore, application of the devised forecasting methodology also significantly reduces the risky practice of exclusively relying on the (qualitative) knowhow of two planners.

6.3 Implementation at H&S and Den Hartogh

This research has further strengthened H&S Foodtrans in their conviction that forecasting the number of loadings and deliveries and corresponding tank container and trucking capacity is of crucial importance in addressing their operational challenges. To reap the benefits of this study, H&S Foodtrans acknowledged the need to implement the proposed forecasting methodology and integrate it in their planning software. To this end, a collaborative project between H&S, Den Hartogh, and CQM was initiated in order to develop and integrate a forecasting methodology that works for both H&S as well as Den Hartogh.

Den Hartogh is a logistics service provider of bulk liquids and gasses for the chemical industry. Although their operation is closely intertwined with the operations of H&S, they are not their competitors since they ship very different commodities. In contrast, H&S and Den Hartogh collaborate on many topics and even use the same tailor-made planning software (i.e. Transfusion) that has been developed exclusively for them. Furthermore, CQM is a data science consultancy firm that has helped both H&S as well as Den Hartogh in optimizing their operational processes throughout the years. As such, CQM takes the leading and coordinating role in this project.

The goal of this collaborative project is (1) to design a forecasting model capable of accurately predicting the number of loadings and deliveries and corresponding tank container and trucking capacity and (2) integrate this model in the planning software at H&S and Den Hartogh. The findings of this study in general, and the proposed forecasting methodology in particular, are used as foundational stepping stones to develop an accurate model applicable to both H&S and Den Hartogh. Specifically, it will be tested whether the proposed forecasting methodology is transferable to all other planning regions of H&S Foodtrans and the operations of Den Hartogh.

7 Conclusions and recommendations

In this final chapter, an answer is provided to the main research question by synthesizing the key insights and results of this thesis. Furthermore, the limitations of the research are discussed together with potential directions for future research.

7.1 Conclusions

As the required capacity is a derivative of the number of loadings and deliveries, this research began by forecasting the expected number of loadings and deliveries in a certain region from historical data. Since the time series considered in this thesis exhibit multiple seasonal patterns, a number of forecasting models were implemented that are well equipped to deal with problems involving such high frequency time series data with multiple levels of seasonality. Subsequently, these models were evaluated, and it was concluded that, in general, the models with the highest predictive accuracy were the simple mean method, dynamic harmonic, regression and ANNs. On the basis of the remarkable performance of the simple mean method, the conclusion seems warranted that a lot of the variation in the data can be explained by the weekly and daily seasonal components and their interaction effects. Moreover the striking performance of the simple mean method also adds to the growing corpus of research showing that complex models do not always generate more accurate forecasts (Green & Armstrong, 2015).

The second step in answering the main research question involved adjusting this initial forecast by exploiting advance demand information (i.e. demand for the future that is already known at present). A Bayesian algorithm was developed that explores and analyses the possibility of using the expected number of orders for a future period as the variable to be estimated. Thereupon, the Bayesian estimate of the expected number of orders is used to derive the adjusted forecast. During a one-month test period, the Bayesian algorithm was applied to the BFN and Rotterdam loadings series and it was shown that it increased the accuracy of the initial forecast by 65% and 28%, respectively. It was also acknowledged, however, that Bayesian models are criticized on the grounds that their complexity requires a "specialist" to understand and implement them. For this reason, a relatively simple model that combines the initial forecast with an inflator algorithm was developed that can serve as alternative for the Bayesian technique. Although this combined forecast did not outperform the more complex Bayesian adjustment, the difference in predictive accuracy was shown to be relatively small. One should also be mindful that another reason for preferring the Bayesian technique over the combined forecast might be that the Bayesian algorithm produces an entire probability distribution, whereas the combined forecast only generated point forecasts. As a result, the Bayesian technique allows for the computation of prediction intervals, whereas this is not the case for the combined forecast. The next steps in this research involved investigating how this forecast can be used to make an accurate prediction of the tank container and trucking capacity.

With regards to the tank container capacity, it was argued that the forecasted loadings can be disaggregated by type of tank container. A top-down method based on forecasted proportions was employed to generate a final forecast of the tank container capacity. In this method, one first generates a preliminary forecast per tank container type from historical data. To obtain these forecasts, an artificial neural network was developed for each tank container type. These forecasts are not used directly. Instead, the preliminary forecasts are only used to calculate the forecast proportions of each tank container type. Finally, the forecasted loadings (derived from Bayesian algorithm) can be disaggregated down the hierarchy based on these forecasted proportions. Similarly as the Bayesian algorithm, the accuracy of the proposed forecasting methodology for predicting tank container capacity was assessed during a one-month test period. Based on this test set it was

concluded that the tank container capacity can be forecasted with reasonable accuracy. Particularly, the mean absolute error expressed in number of tank containers depends on the level of aggregation and varies between 0.57 and 2.72. It is therefore believed that the proposed forecasting methodology for predicting tank container capacity provides a valuable addition to the knowledge of the experienced planners at the MMP department in making repositioning decisions for empty tank containers. Moreover, it also serves as a back-up on which less experienced planners might rely in case the two experienced planners are suddenly unavailable.

Regarding trucking capacity, the first step was to estimate the historical trucking capacity per region based on actual data as well as theoretical assumptions. Thereupon, a multiple linear regression model was developed that uses the forecasted loadings and deliveries and other explanatory variables to predict the required trucking capacity. After applying the proposed model to the same one-month test period and comparing the accuracy with the current forecasting approach occasionally used by H&S, it was revealed that the proposed methodology improves the predictive accuracy by 41%. Particularly, the mean absolute error during working days was shown to be 20.39 hours (i.e. approximately 1.62 trucks). This result demonstrates that the proposed forecasting methodology provides a valuable starting point for efforts aimed at planning the drayage operations in a more proactive manner.

In general, the findings of this research indicate that the proposed forecasting methodology accurately predicts the number of loadings and deliveries and the corresponding tank container and trucking capacity. As such, it is believed that the devised forecast methodology presents a powerful tool that can assist planners at H&S Foodtrans in (1) making more efficient tank container repositioning decisions, and (2) planning the drayage operations in a more proactive fashion. Moreover, this study also adds to the growing body of research indicating that complex models do not always generate more accurate forecasts.

7.2 Limitations and directions for future research

This chapter is concluded by a discussion of the implication of the findings to future research into this area. A number of limitations regarding this study are highlighted and it is discussed how these limitations might relate to fruitful directions for future research in general, and the implementation project that is currently being conducted.

The first limitation is concerned with the observation that the models developed are trained (and evaluated) based on the realized demand. For example, a data set containing the realized daily number of loadings of 2017 and 2018 was used to train various models. This realized demand does not, however, represent direct customer demand. It might, for instance, be the case that a client places an order for a certain day, but that due to capacity constraints, this order cannot be accepted. In this example, the request of the client is not included in the data set used to train the models. Pro-active rescheduling of orders by planners to balance the workload throughout the week and unforeseen delays and disruptions during transportation distort the observation of actual demand in a similar way. This limitation might be addressed by future research that focusses on the direct rather than the realized demand. One should be mindful that at present this type of data is simply not available. However, in collaboration with their software developer, H&S is currently making an attempt to gather this data.

Along the same line, the data that was used to predict the required capacity per tank container type, consists of the historical allocation of tank container types. Although this data does capture the expert knowledge of the planners at H&S, it does not take into account the dynamic nature of the tank container types. In other words, the historical allocation of tank type i to order j does not necessarily mean that this order represents a demand *only* for tank type i. It could, for instance, be the case that this order was also allowed to be shipped

by another type of tank container. A different modelling choice could thus involve clustering various order types, and interpret these clusters as representing demand for a certain type of tank container. The advantage of this method would be that one would measure the demand for tank containers more directly. Moreover, it also ensures that the forecast is not "biased" by tank allocation choices of the past. Although certainly an interesting field for further work, one should also be mindful of a number of drawbacks related to this method. First, the degree of difficulty to disentangle these complexities regarding which order types represent a demand for which tank types should not be underestimated. The number of parameters that determine which tank types are best suited to be allocated to a certain order are so numerous that it is difficult to fit them into a set of business rules. Second, this method does not capture the expert knowledge of the planners at H&S. For these reasons, H&S decided that they want to stick to forecasting the demand for tank containers based on historical allocation.

Another limitation lies the fact that the (historical) trucking capacity was estimated based on both actual data as well as theoretical assumptions. Although H&S does record the exact sequence of actions for each order, as well as the actual time it takes to perform the loadings and delivery actions, there is no actual data on the other trucking actions. Hence, assumptions were made related to the trucking hours required for a pick-up action, drop action, cleaning action, and the trucking hours needed to drive to the next location. Nevertheless, despite the theoretical assumptions, it is expected that the estimation of the (historical) trucking capacity is fairly robust. The theoretical assumptions were derived in consultation with the planners at H&S Foodtrans that have years of practical experience in dealing with these matters. Moreover, in order to gain even more validation for the theoretical assumptions, they were compared to Den Harogh, an industry partner of H&S. Independently, Den Hartogh arrived at very similar assumptions.

A final limitation that needs to be considered is concerned with the size of the Bayesian test set. In Chapter 3, it was explained that, in order to effectively evaluate the Bayesian algorithm, one needs the *most recent* number of advanced orders in future period τ as of time T for all future periods τ that one wants to include in the forecast. Since this advance order information is continuously overwritten in the data warehouse of H&S, this data could not be retrieved for historical records. Hence, to surpass this complication, the required data for the algorithm is manually retrieved from the data warehouse for a total of one month. There seems to be no doubt that the Bayesian algorithm significantly improves the initial forecast. However, given the small sample size of the Bayesian test set, caution must be exercised in dealing with absolute errors in isolation. Ideally, the promising findings of this study regarding the predictive accuracy of the forecasting models should be replicated in a study with a larger test set. Luckily, results regarding this topic are soon to be expected since the proposed forecasting methodology is currently being implemented in consultation with a data science consultant called CQM. Part of this project involves testing the accuracy of the proposed forecasting methodology on a larger test set including multiple planning regions.

In addition to the recommendations discussed above, a final interesting direction for further research might lie in the topic of collaborative forecasting. Instead of using only local information, there might be a great deal of untapped potential in using the forecasts of supply chain partners to enhance one's own predictive model. What makes this area of research particularly interesting for H&S Foodtrans is that a small number of clients are responsible for a large part of the orders. Specifically, over the last six months, the largest 5% of the clients accounted for 55% of the orders. The prospect of harnessing this information, and collaborating with the largest clients seems to be worthwhile exploring. For more details regarding this topic, the reader is referred to Appendix E.

References

Abuizam, R., & Thomopoulos, N. T. (2006). Adjusting An Existing Forecasting Model When Some Future Demands Are Known In Advance: A Bayesian Technique. *Journal of Business & Economics Research (JBER)*, *4*(7).

Armstrong, J. S. (2001). Combining forecasts. In *Principles of forecasting* (pp. 417-439). Springer, Boston, MA.

Athanasopoulos, G., Ahmed, R. A., & Hyndman, R. J. (2009). Hierarchical forecasts for Australian domestic tourism. *International Journal of Forecasting*, *25*(1), 146-166.

Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, *20*(4), 451-468.

Bloomfield, P. (2004). *Fourier analysis of time series: an introduction*. John Wiley & Sons.

Braekers, K., Janssens, G. K., & Caris, A. (2011). Challenges in managing empty container movements at multiple planning levels. *Transport Reviews*, *31*(6), 681-708.

Branch, A. E. (2006). *Export practice and management*. Cengage Learning EMEA.

Chianca, C. V., Ticona, A., & Penna, T. J. P. (2005). Fourier-detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, *357*(3-4), 447-454.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: a seasonal-trend decomposition. *Journal of official statistics*, *6*(1), 3-73.

Crainic, T. G., & Kim, K. H. (2007). Intermodal transportation. *Handbooks in operations research and management science*, *14*, 467-537.

Dangerfield, B. J., & Morris, J. S. (1992). Top-down or bottom-up: Aggregate versus disaggregate extrapolations. *International Journal of Forecasting*, *8*(2), 233-241.

De Livera, A. M., Hyndman, R. J., & Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, *106*(496), 1513-1527.

Dunn, D. M., Williams, W. H., & DeChaine, T. L. (1976). Aggregate versus subaggregate models in local area forecasting. *Journal of the American Statistical Association*, *71*(353), 68-71.

Fite, J. T., Don Taylor, G., Usher, J. S., English, J. R., & Roberts, J. N. (2002). Forecasting freight demand using economic indices. *International Journal of Physical Distribution & Logistics Management*, *32*(4), 299-308.

Fliedner, E. B., & Mabert, V. A. (1992). Constrained forecasting: some implementation guidelines. *Decision Sciences*, *23*(5), 1143-1161.

Funke, J., & Kopfer, H. (2016). A model for a multi-size inland container transportation problem. *Transportation Research Part E: Logistics and Transportation Review*, *89*, 70-85.

Gallego, G., & Özer, Ö. (2001). Integrating replenishment decisions with advance demand information. *Management science*, *47*(10), 1344-1360.

Granger, C. W., & Pesaran, M. H. (2000). Economic and statistical measures of forecast accuracy. Journal of Forecasting, 19(7), 537-560.

Green, K. C., & Armstrong, J. S. (2015). Simple versus complex forecasting: The evidence. Journal of Business Research, 68(8), 1678-1685.

Gross, C. W., & Sohl, J. E. (1990). Disaggregation methods to expedite product line forecasting. Journal of Forecasting, 9(3), 233-254.

Guerrero, V. M., & Elizondo, J. A. (1997). Forecasting a cumulative variable using its partially accumulated data. *Management Science*, *43*(6), 879-889.

Hariharan, R., & Zipkin, P. (1995). Customer-order information, leadtimes, and inventories. *Management Science*, *41*(10), 1599-1607.

Hippert, H. S., Pedreira, C. E., & Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on power systems*, *16*(1), 44-55.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, *22*(4), 679-688.

International Tank Container Organisation. (2011). *Tank Containers: A Sustainable Solution for Bulk Liquid Transport*.

Karaesmen, F., Buzacott, J. A., & Dallery, Y. (2002). Integrating advance order information in make-to-stock production systems. *Iie Transactions*, *34*(8), 649-662.

Konings, R., & Thijs, R. (2001). Foldable containers: a new perspective on reducing container-repositioning costs. *European journal of transport and infrastructure research*, *1*(4).

Lee Rodgers, J., & Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, *42*(1), 59-66.

Lütkepohl, H. (1984). Forecasting contemporaneously aggregated vector ARMA processes. *Journal of Business & Economic Statistics*, *2*(3), 201-214.

Macharis, C., & Bontekoning, Y. M. (2004). Opportunities for OR in intermodal freight transport research: A review. *European Journal of operational research*, *153*(2), 400-416.

MacKay, D. J. (1992). A practical Bayesian framework for backpropagation networks. *Neural computation*, *4*(3), 448-472.

McLeavey, D. W., & Narasimhan, S. L. (1985). *Production planning and inventory control*. Allyn & Bacon.

Negnevitsky, M. (2005). *Artificial intelligence: a guide to intelligent systems*. Pearson education.

Neusser, K. (2016). *Time series econometrics*. Switzerland: Springer.

Portugal, M. S. (1995). Neural networks versus time series methods: a forecasting exercise. *Revista Brasileira de Economia*, *49*(4), 611-630.

Sakia, R. M. (1992). The Box-Cox transformation technique: a review. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *41*(2), 169-178.

Tan, T. (2008). Using imperfect advance demand information in forecasting. *IMA Journal of Management Mathematics*, *19*(2), 163-173.

Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, *54*(8), 799-805.

Utley, J. S., & May, J. G. (2010). The use of advance order data in demand forecasting. *Operations Management Research*, *3*(1-2), 33-42.

Zellner, A., & Tobias, J. (1998). *A note on aggregation, disaggregation and forecasting performance* (No. 1557-2016-132780).

Appendices

A. Organizational structure of H&S Foodtrans

H&S Group B.V. is an encapsulating business entity that governs all aspects of providing logistic services in the (intermodal) liquid food industry and consists of five business units: coldstores, logistic services, cleaning, transport, and Foodtrans. Although these business units are organized under the same industrial group, they operate independently of each other. The organizational structure is visualized in Figure A.1.

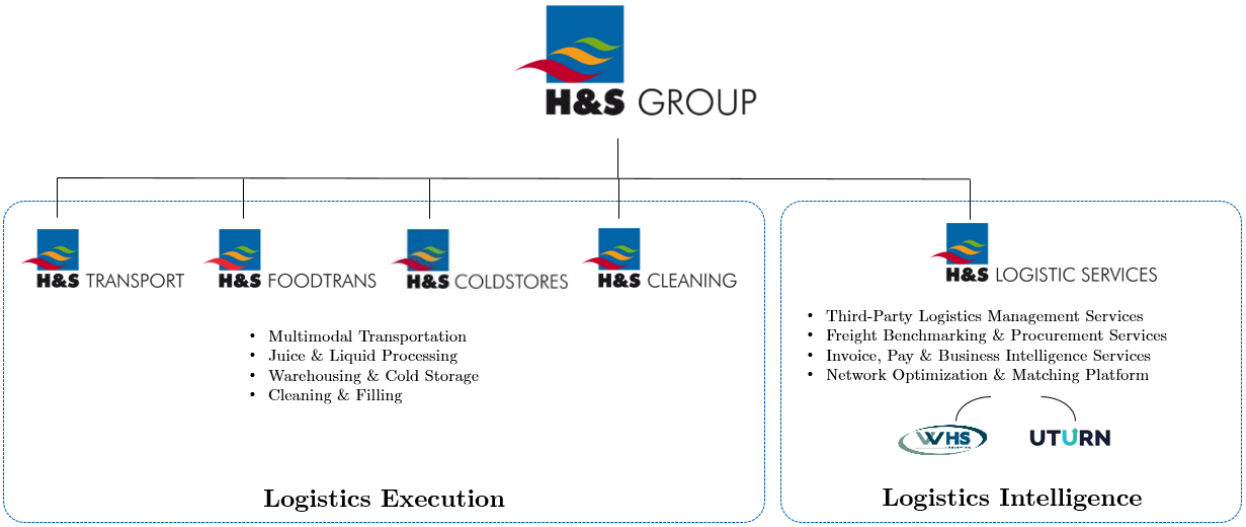


Figure A.1 Organizational structure H&S Group

B. Seasonality in the three planning regions

Figure B.1 presents a heat map of the number of loadings and deliveries in 2018 for the three planning regions considered in this research. It thus visualizes some of the conclusions that were drawn in in section 2.1.2.

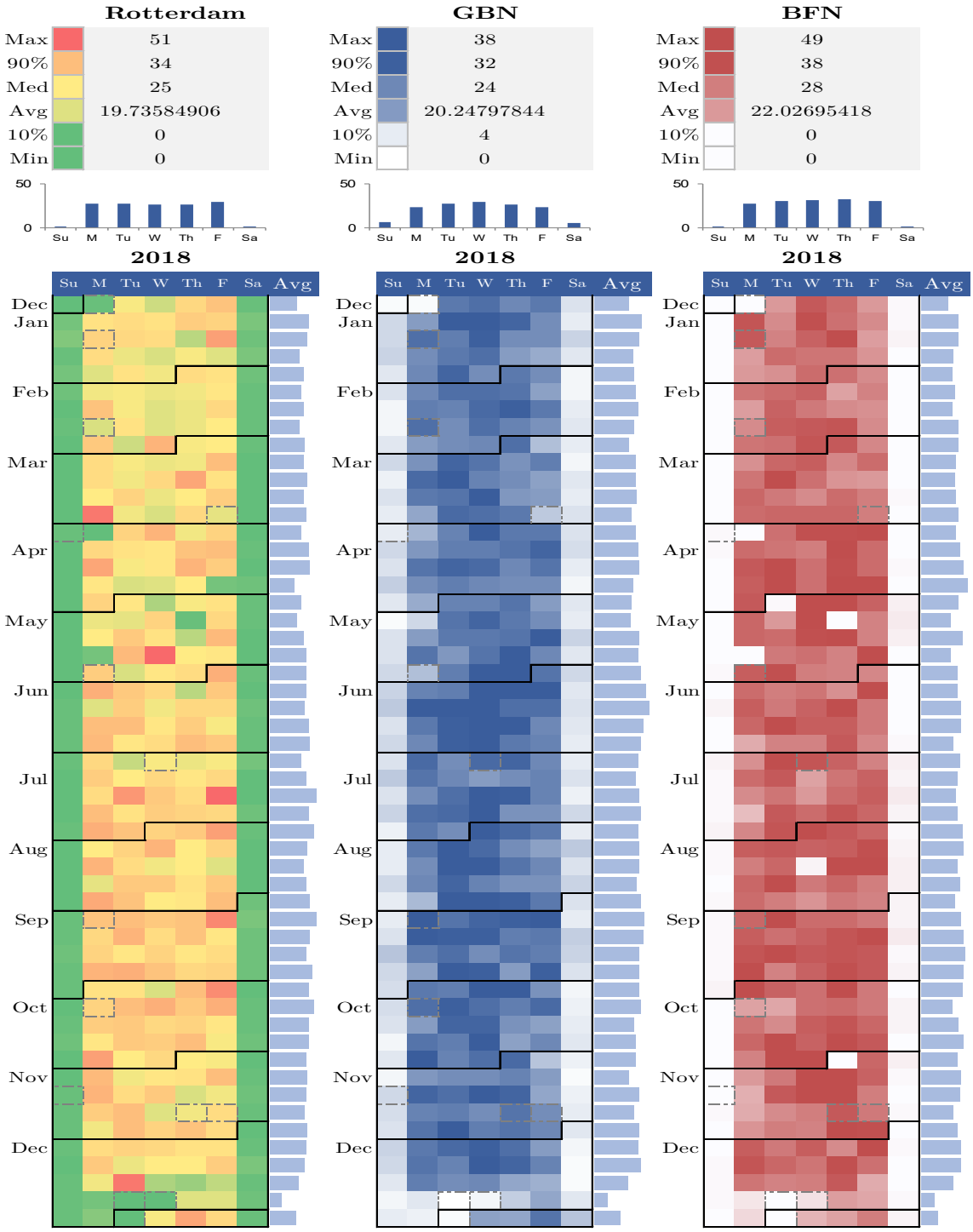


Figure B.1 Heat map of loadings and deliveries in three planning regions

C. Parameters of forecasting models Chapter 2

In Chapter 2, various forecasting models were implemented to predict the number of loadings and deliveries based on historical data. Table C.1 and C.2 summarize the relevant parameters of the implemented models for each region. Some of the notation used in these tables might demand further clarification which is briefly given below:

- The triplet (E,T,S) refers to the three components: error, trend, and seasonality. For example, the model $ETS(M, M_d, M)$ denotes an exponential smoothing model with multiplicative errors, a damped multiplicative trend, and multiplicative seasonality. In particular, the Trend component might either be non-existent (N), additive (A), additive damped (A_d), multiplicative (M), or multiplicative damped (M_d). In a similar vein, the seasonal component can either be non-existent, additive, or multiplicative. Finally, the error component can also be additive or multiplicative.
- If one combines differencing with autoregression and a moving average model, we obtain a non-seasonal ARIMA model, formally referred to as an $ARIMA(p, d, q)$ model. In this model p, d, and q denote the order of the autoregressive part, degree of first differencing involved, and the order of the moving average part, respectively.
- Another forecasting model that was discussed in Chapter 2 is the so-called TBATS model. Recall that this model can be summarized as:

  $TBATS(omega, \{p, q\}, phi, \{\langle m1, k1\rangle \dots, \langle mJ, kJ\rangle\})$.

  In this model summary, $omega$ refers to the Box-Cox parameter, $\{p, q\}$ refers to the $ARMA(p, q)$ process that is used to model the error term, and phi is the damping parameter. Finally, $m1, \dots mJ$ and $k1, \dots kJ$ denote the seasonal periods used in the model and the corresponding number of Fourier terms used for each seasonality (De Livera, Hyndman & Snyder, 2011).

Table C.1 Parameters of the models derived in Chapter 2 (1/2)

Time series	STL + exponential smoothing	Double exponential smoothing
Rotterdam loadings	STL + ETS(A,N,N) Smoothing parameters: - alpha = 0.1274 Initial states: - l = 14.9948	Smoothing parameters: - alpha = 0.005902684 - beta = 2.990411e-08 - gamma1 = 0.004007208 - gamma2 = 0.03343798 Initial states: - l = 4.146563 - b = -0.001917039
Rotterdam deliveries	STL + ETS(A,N,N) Smoothing parameters: - alpha = 0.0648 Initial states: - l = 0.692	Smoothing parameters: - alpha = 0.005902684 - beta = 2.990411e-08 - gamma1 = 0.004007208 - gamma2 = 0.03343798 Initial states: - l = 4.146563 - b = -0.001917039
GBN loadings	STL + ETS(A,N,N) Smoothing parameters: - alpha = 0.0954 Initial states: - l = 2.0351	N.A.
GBN deliveries	STL + ETS(A,Ad,N) Smoothing parameters: - alpha = 0.1229 - beta = 1e-04 - phi = 0.9006 Initial states: - l = 11.7098 - b = -1.2565	N.A.
BFN loadings	STL + ETS(A,Ad,N) Smoothing parameters: - alpha = 0.0156 - beta = 3e-04 - phi = 0.8 Initial states: - l = -1.6133 - b = 0.8235	N.A.
BFN deliveries	STL + ETS(A,N,N)	N.A.

	Smoothing parameters: - alpha = 0.0982 Initial states: - l = 1.5294	

Table C.2 Parameters of the models derived in Chapter 2 (2/2)

Time series	Dynamic harmonic regression	TBATS
Rotterdam loadings	No ARIMA process (i.e. only harmonic regression) Coefficients in regression: β_0 = 16.83388 α_1 = -5.22196 γ_1 = -9.01731 α_2 = -7.95307 γ_2 = -5.84468 α_3 = -3.02102 γ_3 = -1.10197 β_1 = 0.821381 θ_1 = -0.35204 β_2 = 0.216445 θ_2 = 0.238127 β_3 = 0.039054 θ_3 = 0.143011 β_4 = -0.37924 θ_4 = -0.20833 β_5 = -0.35288 θ_5 = -0.03142 β_6 = 0.365335 θ_6 = -0.53416 β_7 = -0.07897 θ_7 = -0.08361 β_8 = -0.15524 θ_8 = -0.35334 β_9 = -0.24063 θ_9 = 0.39141 β_{10} = -0.39574 θ_{10} = -0.23709 β_{11} = -0.01476 θ_{11} = -0.02547 β_{12} = -0.22664 θ_{12} = -0.44389 β_{13} = 0.091205 θ_{13} = -0.03815 β_{14} = 0.392093 θ_{14} = 0.405504 δ_1 = -0.2934 ρ_1 = -1.52611	$TBATS(1, \{0,0\}, -, \{\langle 7,3\rangle, \langle 30.43,6\rangle, \langle 365,6\rangle\})$
Rotterdam deliveries	ARIMA process of error term: - ARIMA(7,1,1) Coefficients in regression: α_1 = -0.29079	$TBATS(1, \{0,1\}, -, \{\langle 7,3\rangle, \langle 30.43,8\rangle, \langle 365,6\rangle\})$

	γ_1 = -1.78283 α_2 = -0.6826 γ_2 = -0.47984 α_3 = -0.52048 γ_3 = 0.106741 β_1 = -0.04813 θ_1 = -0.29114 β_2 = -0.09096 θ_2 = 0.135749 β_3 = -0.06254 θ_3 = 0.016879 β_4 = -0.1354 θ_4 = -0.2116 β_5 = 0.142476 θ_5 = 0.026193 β_6 = -0.002 θ_6 = -0.27746 β_7 = -0.22876 θ_7 = 0.093645 β_8 = 0.051027 θ_8 = -0.23875 β_9 = 0.04334 θ_9 = -0.06118 β_{10} = 0.094146 θ_{10} = 0.076091 β_{11} = -0.06128 θ_{11} = -0.15755 β_{12} = -0.09699 θ_{12} = -0.04398 β_{13} = -0.02329 θ_{13} = 0.105538 β_{14} = -0.06904 θ_{14} = 0.091826 δ_1 = 0.154341 ρ_1 = -0.23439	
GBN loadings	ARIMA process of error term: - ARIMA(14,1,3) Coefficients in regression: α_1 = -0.739189755 β_1 = -0.1247967 θ_1 = -1.617247663 β_2 = -0.154413419 θ_2 = -0.63855242 β_3 = -0.000909611 θ_3 = -0.232786289 β_4 = 0.447434405 θ_4 = 0.081944404 β_5 = 0.316866519 θ_5 = 0.170879928 β_6 = 0.372163577 θ_6 = 0.214817916 δ_1 = -0.358004898 ρ_1 = -0.130267426	$TBATS(1, \{5,0\}, 0.829, \{\langle 2,1\rangle, \langle 14,6\rangle, \langle 730,6\rangle\})$
GBN	No ARIMA process (i.e. only harmonic regression)	

deliveries	Coefficients in regression: β_0 = 7.367359 α_1 = -2.78467 β_1 = -0.532 θ_1 = -4.11258 β_2 = -0.63852 θ_2 = -2.03871 β_3 = -0.12129 θ_3 = -0.70705 β_4 = 0.244642 θ_4 = 0.710667 β_5 = 0.292088 θ_5 = 0.878287 β_6 = 0.732602 θ_6 = 1.444247 δ_1 = -0.56784 ρ_1 = -0.130267426	$TBATS(1, \{5,0\}, -, \{\langle 2,1\rangle, \langle 14,6\rangle, \langle 730,6\rangle\})$
BFN loadings	ARIMA process of error term: - ARIMA(14,0,0) Coefficients in regression: β_0 = 7.444660687 α_1 = -1.141346389 β_1 = -1.748421354 θ_1 = -4.962391034 β_2 = -1.25740938 θ_2 = -3.260662924 β_3 = -0.156547467 θ_3 = -0.678790224 β_4 = 0.580834555 θ_4 = 0.532095871 β_5 = 0.983183495 θ_5 = 0.969852802 β_6 = 0.873286196 θ_6 = 1.139582744 δ_1 = -0.413085869 ρ_1 = -0.376567143	$TBATS(1, \{0,0\}, -, \{\langle 2,1\rangle, \langle 14,6\rangle, \langle 730,1\rangle\})$
BFN deliveries	ARIMA process of error term: - ARIMA(14,0,2) Coefficients in regression: β_0 = 3.296612115 α_1 = -1.056122575 β_1 = -0.232600642 θ_1 = -2.542331465 β_2 = -0.343993995 θ_2 = -0.861537147 β_3 = -0.245414195 θ_3 = -0.192374847 β_4 = 0.152382827 θ_4 = 0.195908773 β_5 = 0.312623034 θ_5 = 0.575201726 β_6 = 0.184665836	$TBATS(1, \{5,0\}, -, \{\langle 2,1\rangle, \langle 14,6\rangle, \langle 730,5\rangle\})$

	θ_6 = 1.017555698 δ_1 = -0.109966088 ρ_1 = -0.244754861	

D. Theoretical assumptions multiple linear regression model

Table D.1 confirms that the errors are uncorrelated to the predictor variables. This was important to check since if this is proves not to be the case, there is more information that should be included in the systematic part of the model.

Table D.1 Correlation of the explanatory variables to the residuals

	Correlation BFN region	Correlation GBN region
$Lo(\tau)$	-6.45416E-15	-6.33903E-16
$De(\tau)$	-6.95861E-16	3.70684E-15
$Lo_{D(\tau)}$	-5.48185E-15	1.99549E-15
$De_{D(\tau)}$	-5.4537E-15	-1.92108E-15
$\delta(\tau - 14)$	-1.55267E-15	-1.94362E-15
$\delta(\tau - 28)$	-2.05473E-15	1.06473E-15
$x_1(\tau)$	1.27819E-15	2.25425E-15
$x_2(\tau)$	1.01182E-14	-2.31537E-15
$x_3(\tau)$	9.42481E-15	-1.40521E-15
$x_4(\tau)$	9.86995E-15	4.65364E-16
$x_5(\tau)$	3.14295E-15	1.23475E-15
$x_6(\tau)$	-8.77957E-15	6.42293E-15
$x_7(\tau)$	5.72758E-15	9.17448E-16
$x_8(\tau)$	-2.56064E-15	2.9992E-15

Another important theoretical assumption that was made when deriving the multiple linear regression model, is that the residuals should be homoscedastic; otherwise p-values are generally underestimated, and coefficient estimates are less precise. Initially it was shown that the residuals were in fact heteroskedastic. For this reason, a Box-Cox transformation was performed to resolve this heteroscedasticity. By visually inspecting Figure D.1 and D.2, it seems that the residuals look more homoscedastic. To confirm this notion, a white test was again conducted, which now resulted in insignificant p-values for both models. Hence, it can be concluded that the heteroscedasticity in both models is resolved, using the Box-Cox transformation.

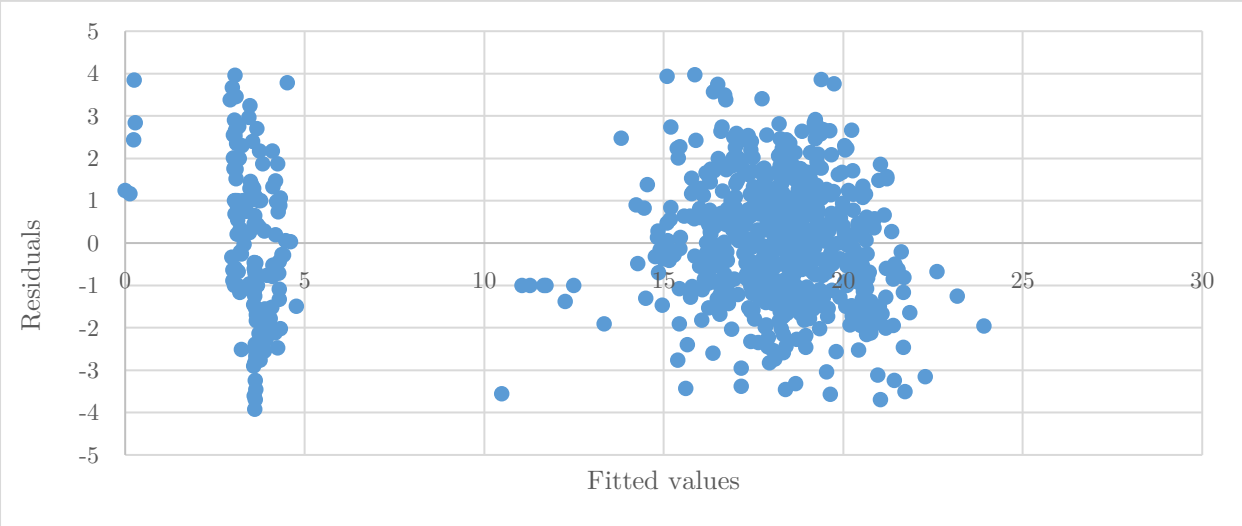


Figure D.1 Fitted values vs. Residuals plot BFN region after Box Cox transformation

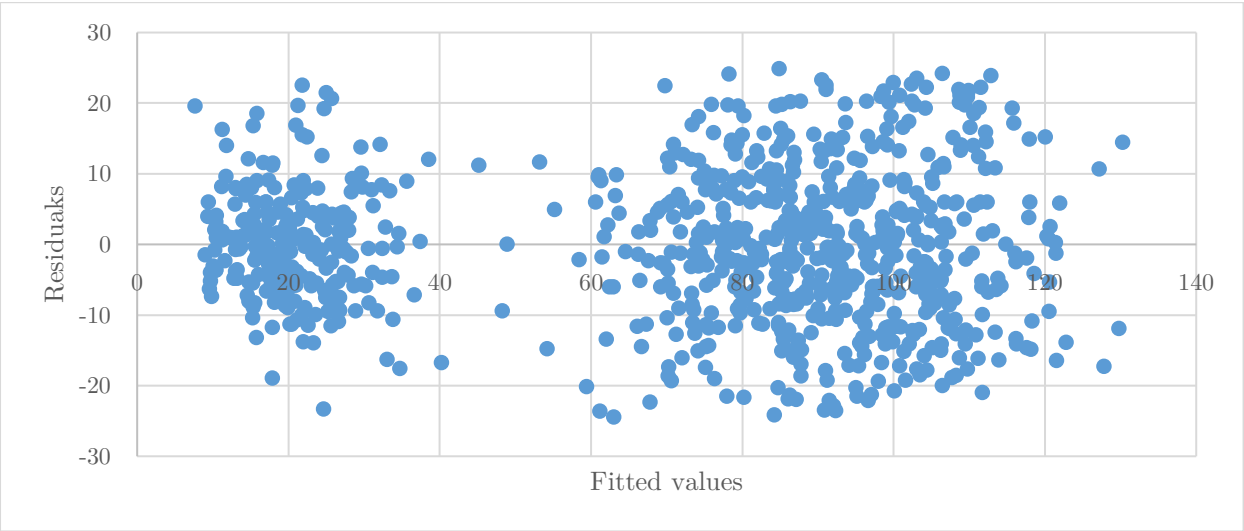


Figure D.2 Fitted values vs. Residuals plot GBN region after Box-Cox transformation

E. Contribution of clients to the number of orders

Figure E.1 represents a Pareto chart, where individual values are represented in descending order by bars, and the cumulative total is represented by the line. The implication of this figure is that a small number of clients is responsible for a large part of the orders. Specifically, over the last six months, the largest 5% of the clients accounted for 55% of the orders and so on.

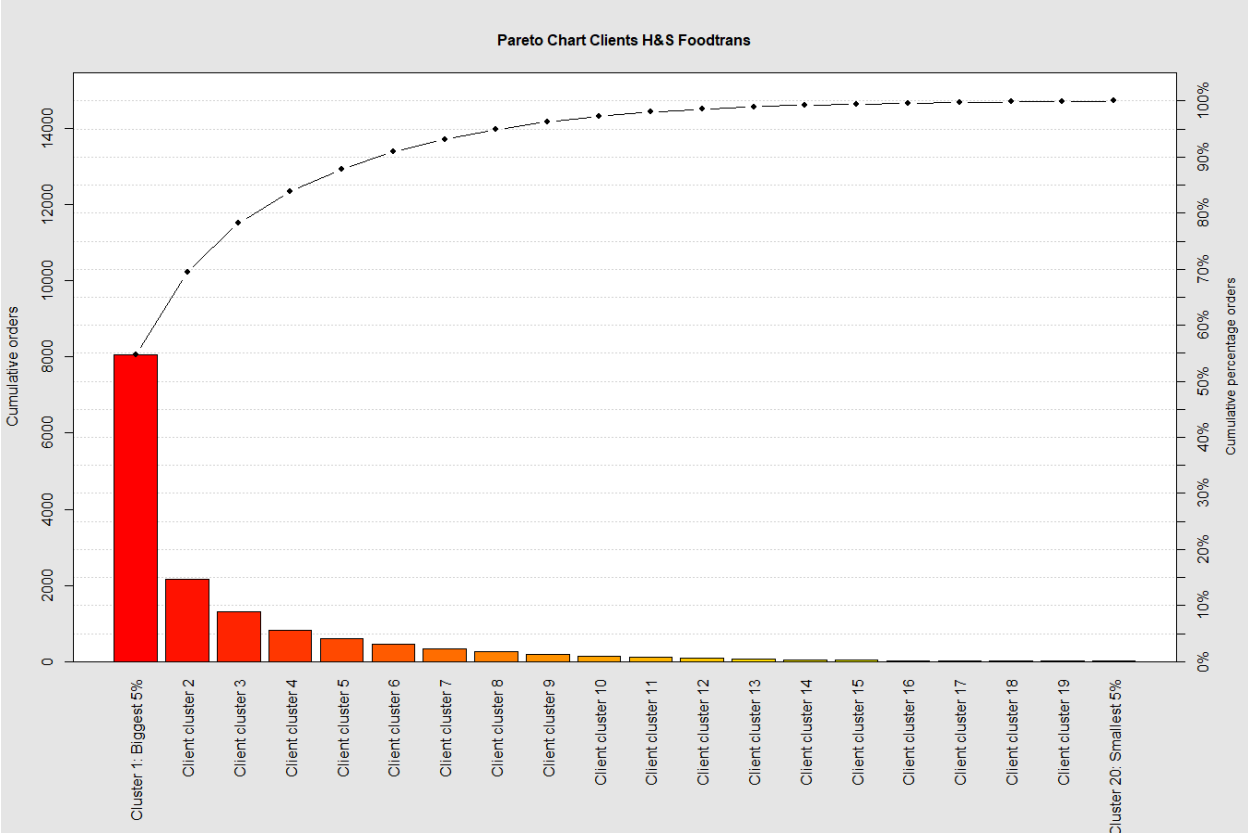


Figure E.1 Pareto chart clients H&S