A comparison of forecasting methods applied to distribution planning in retail operations

Maria Francisca Fonseca Gonçalves Arantes

Master's Dissertation

FEUP Supervisor: Prof. Mário Amorim Lopes

FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Mestrado Integrado em Engenharia e Gestão Industrial

2020-07-15

Abstract

The food retail business is becoming increasingly competitive over the past few years. Thus, operational efficiency is key to thrive in the market.

Supply chain management is challenging for every retail company. Moreover, distribution is one of the most costly areas of the supply chain. Hence, the aim of this dissertation is to provide tools to support the distribution planning of goods from the warehouse to the stores.

In particular, one of the improvement opportunities found was the estimation of the number of pallets to be shipped to each store at the end of the day. This estimation is crucial as it is the input to the route optimization software. Errors in the estimates might lead to last minute changes in the organization of the expedition and for the need of extra trucks to perform the delivery, forcing the company to incur extra costs.

Therefore, two approaches were followed to enhance the estimates. The first was more practical and simple, based on the volume of each product and each store's demand. The idea is to construct pallets according to the product request made by each store, and the volume of each product requested.

On the other hand, a more empirical approach was tested, resorting to forecasting methods for time series. Statistical forecasting methods like Holt-Winters, ARIMA and TBATS were tested, together with machine learning algorithms such as Random Forest and Neural Networks. As it is possible to include explanatory variables in machine learning models, a multivariate model was also built and tested.

As a result, the first methodology proposed proved to be a viable solution to this problem, obtaining the best performance.

Regarding the second approach followed, statistical methods had a similar performance between them, with Random Forest model having the best performance among the forecasting methods for time series. Explanatory variables were useful in improving forecast accuracy. Moreover, LSTM Neural Networks were not able to enhance the forecast accuracy when compared to Random Forest.

Key Words: Retail, Distribution, Forecasting, Holt-Winters, ARIMA, TBATS, Random Forest, Long Short Term Memory Neural Networks

Resumo

Ao longo dos últimos anos, a indústria do retalho alimentar tem vindo a ficar cada vez mais competitiva. Assim, ser eficiente a nível operacional é essencial para ter sucesso no mercado.

A gestão da cadeia de abastecimento é particularmente difícil para qualquer retalhista. Por sua vez, a área da distribuição constitui uma das áreas com maiores custos associados da cadeia de abastecimento. Assim, o objetivo desta dissertação é fornecer as ferramentas de suporte necessárias para a realização do plano de distribuição dos produtos do armazém até às lojas.

Uma das oportunidades de melhoria encontradas foi a estimativa do número de paletes a serem expedidas ao final de cada dia. Esta estimativa é crucial uma vez que é necessária como variável de entrada do software de otimização de rotas. Erros nas estimativas podem resultar em mudanças de última hora na organização da expedição e na necessidade de recorrer a um maior número de camiões para realizar as entregas, fazendo a empresa incorrer em custos extra.

Para isso, duas abordagens foram seguidas. A primeira constitui uma abordagem mais prática e simples baseada no volume de cada produto e no pedido feito por cada loja. A ideia é construir paletes de acordo com os produtos presentes no pedido de cada loja e no seu respetivo volume.

Por outro lado, foi também testada uma metodologia mais teórica recorrendo a métodos de previsão para séries temporais. Métodos estatísticos de previsão como Holt-Winters, ARIMA e TBATS foram testados, bem como métodos de machine learning, tais como o algoritmo Random Forest e as Redes Neuronais. Sendo possível incluir variáveis explicativas nos modelos de machine learning, modelos multivariados foram também construídos e testados.

Foi possível concluir que a primeira metodologia proposta constituí uma opção viável para solucionar este problema, tendo sido o método com o melhor desempenho.

Quanto à segunda metodologia proposta, os modelos estatísticos de previsão apresentaram resultados bastante semelhantes entre eles, com o modelo Random Forest a destacar-se como o modelo com a melhor prestação entre os modelos de previsão para séries temporais analisados. As variáveis explicativas adicionadas ao modelo melhoraram os resultados da previsão. Finalmente, as Redes neuronais LSTM não foram capazes de melhorar os resultados obtidos pelo modelo Random Forest.

Palavras-chave: Retalho, Distribuição, Previsão, Holt-Winters, ARIMA, TBATS, Random Forest, Long Short Term Memory Neural Networks

Acknowledgements

I would like to thank the Kaizen Institute for the opportunity to develop this project. My team for the motivation and fast integration. My supervisor Miguel Hespanhol for the support, the high standards and the trust placed. And Rui Neto for the availability and support given in this project.

Prof. Mário Amorim Lopes for the guidance and availability.

My friends for being with me along the way.

My parents and sisters for the support in all the important moments and the opportunities given.

And finally, I thank Miguel for listening and for the support in the difficult moments.

"It's difficult to make predictions, especially about the future"

Niels Bohr

Contents

Acronyms and Symbols

AIC	Akaike Information Criterion
ANN	Artificial neural networks
ARIMA	Auto-Regressive Integrated Moving Average
ARMA	Auto-Regressive Moving Average
APED	Associação Portuguesa de Empresas de Distribuição
DC	Distribution Center
KPI	Key Performance Indicator
LSTM	Long Short Term Memory
NN	Neural Networks
NRMSE	Normalized Root Mean Squared Error
RF	Random Forest
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
SARIMA	Seasonal Auto-Regressive Integrated Moving Average
SKU	Stock Keeping Unit
SVM	Support Vector Machines
TBATS	Trigonometric seasonality Box-Cox transformation ARMA errors Trend and Seasonal components
wMAPE	Weighted Mean Absolute Percentage Error
3PL	Third-party Logistics provider

List of Figures

List of Tables

Chapter 1

Introduction

Supply chain management in the food retail business poses numerous challenges. Distribution planning, in particular, is a difficult task, being crucial to operations and one of the most expensive areas of the supply chain.

The aim of this proposal is to support the distribution planning of a food retail company, hereafter mentioned as Company A. In order to achieve that, forecasting methods for time series together with a methodology based on volumetry will be implemented to predict the daily number of pallets shipped to each store location by one of their North region warehouses.

It this chapter, a short presentation of the Kaizen Institute and Its principles is made. Afterwards the project's background and motivation are presented as well as a brief introduction to the food distribution industry in Portugal and Company A. Finally, the goals and methodology adopted for the project are presented, finishing with the dissertation structure.

1.1 The Kaizen Institute

The Kaizen Institute is a consulting company founded in 1985 by Masaaki Imai. Shortly after, in 1986, Imai publishes his book *Kaizen: The Key to Japan's Competitive Success*, where he describes the *Toyota Production System*, a methodology aiming to achieve the elimination of all waste while maximizing efficiency (Toyota Motor Corporation (2020)).

It was in the twenty-first century when the Toyota Motor Company surpassed General Motors to become the top automotive manufacturer in the world that the Lean and Kaizen philosophies emerged, due to their key role in Toyota's success.

The word Kaizen comes from the junction of the Japanese words "Kai" and "Zen" that means change and better, originating the translation continuous improvement.

The Kaizen Institute is dedicated to implement and build a sustainable culture of continuous improvement within its clients.

It is present in Portugal since 1999, and currently has offices spread across the five continents.

1.2 The Kaizen Philosophy

The Kaizen Business model of continuous improvement stands upon five guiding principles presented next:

Value creation for the client: Value is the difference between the utility perceived by the client and the price to pay for a product or a service. Nowadays, clients are only willing to pay for value-adding activities. Thus, organizations should concentrate their resources on activities that aggregate and generate value to the client while eliminating any others.

Gemba orientation: Gemba means the real place, representing the place to make improvement (Coimbra (2013)) that could be either the shop floor or an office, for example. The methodology focuses on going to Gemba as it is the best source of information, the best place to engage the workforce, where change takes place and where problems and improvement opportunities can be identified.

Muda elimination: Muda is the Japanese word for waste, and it includes the non-value added activities a client is not willing to pay for. The elimination of Muda is an opportunity to reduce costs while increasing the company's value and efficiency. There are seven types of Muda: defects, people waiting, people moving, too much processing, material waiting, material moving and overproduction (Imai (2005)). In distribution activities some of these wastes can be more thoroughly identified, for example:

- Errors: misinformation given to drivers about their routes, cargo wrongly loaded on the truck, that translate into poor service quality, corrections, extra miles and extra costs.
- Over-processing: complicated processes, inefficient placement of cargo or inexistent standard processes may lead to additional efforts and poor resource utilization.
- People, vehicles and cargo waiting: This occurs when trucks are waiting to be loaded or unloaded, leading to extra costs and long processing times.
- People, vehicles and cargo moving: Due to poor routing optimization or non-value-added movement of vehicles and goods.

Involvement of people: Kaizen highlights the importance of developing people through teamwork (Coimbra (2013)), as its sustainability relies upon people's performance. Workers' motivation and training are of extreme importance as well as a spirit of mutual respect and trust across all levels of the organization. One can not delegate Kaizen (Imai (2005)), everyone in the company from top management to the employees needs to be fully engaged and aligned with the continuous improvement mindset. Moreover, this philosophy blames the process and not the people, while seeing problems as improvement opportunities.

Visual standards: Visual elements implementation allows for a better perception of the process, increasing productivity and process transparency; besides, problems can be more rapidly identified and tackled.

For a company to improve, change must occur. People are naturally resistant to change, so change management is crucial for Kaizen's success. A paradigm is a preconceived idea of a situation or the way one reacts. In continuous improvement one is constantly questioning old paradigms and creating new approaches and processes aiming at improving efficiency. Standardization, worker's motivation, and training are essential to sustain change over time.

1.3 The food distribution and retail business in Portugal

The food distribution market in Portugal is a rather competitive market, with a high sales volume and great potential growth.

In the years of 2018 and 2019, almost 300 supermarkets have opened, mostly in the form of smaller proximity stores (Diário de Notícias (2019)). According to APED, the Portuguese association of distribution companies, Portugal has a ratio of stores per inhabitant lower than other European Union countries (Diário de Notícias (2019)).

The main players in the market are the Portuguese companies Sonae SGPS, that through Sonae MC owns Continente hypermarkets and supermarkets, and Jerónimo Martins owner of the supermarket chain Pingo Doce and the wholesaler Recheio. In the past few years, hard discounters like Lidl and Aldi have entered the Portuguese market. More recently Mercadona, market leader in Spain, has opened its first store in Portugal. Other international retailers like Auchan, Dia, Intermarché and Supercor are also present in the Portuguese territory.

In 2018, hypermarkets represented 26,3% of the sales, supermarkets 49,8%, and discounters 15,2%, whereas other points of sale accounted for 8,6% (Pedro Curvelo (2019)).

Some of the strategies followed by companies include the use of client cards, with benefits for its holders and promotional campaigns, rather effective on Portuguese consumers (Pedro Curvelo (2019)). Other retailers focus their efforts on differentiation, providing higher quality products, and/or offering a wide variety of articles in their stores.

1.4 Company A

The company mentioned through this dissertation is a Portuguese company with 225 years of history, dedicated to food distribution and specialized retail, being the first one in their core business.

The company is organized by Business Units, having a divisional structure. Furthermore, they have operations in three different countries. In Portugal, they operate 42 wholesale stores and 461 supermarkets, being one of the most competitive players in the food distribution market.

As strategic goals, the company aspires to become the leader in the markets in which operates, create strong and responsible chains and brands, and to grow steadily in terms of sales and profitability in each one of their business units.

Concerning their corporate responsibility, the company has five main cornerstones: promotion of health through nutrition, respect for the environment, encouragement of a responsible consumption, support to surrounding communities and self-establishment as an employer of reference.

The company is dedicated to providing quality and differentiated products at competitive prices, while maintaining a high service level. They focus on a quality, trustworthy and affordable private label. Additionally, they have a widespread network of stores, that tend to be located in residential areas.

Finally, the company is devoted to continuous improvement and innovation, focusing on constantly increasing productivity and efficiency in their operations.

1.5 Project background and motivation

The food retail business is becoming increasingly competitive over the past few years. Hence, operational efficiency is crucial to succeed in this market.

Nowadays, most retailers operate their own distribution center (DC), rather than having suppliers responsible for the delivery of their individual products directly to the stores. The broad majority of products flow from supplier to the store through this channel (Hbner et al. (2013)). Distribution centers offer the opportunity to bundle products from multiple suppliers, thus reducing the number of trucks arriving in stores and facilitating a more frequent product delivery (Sternbeck and Kuhn (2014)).

On the other hand, the competitiveness of the market also requires a higher consumer orientation (Hbner et al. (2013)). Currently, consumers are better informed, price sensitive, and more demanding, expecting higher service levels, fresh products and greater shopping convenience (Hbner et al. (2013), Sternbeck and Kuhn (2014)). Moreover, there is an increased demand for differentiated articles and product diversity at the stores.

Therefore, supply chain management in food distribution and retail is of most importance and poses numerous challenges. Retailers must fulfill consumers' needs while increasing profitability and reducing operational costs.

Supply chain management comprises four main areas: procurement, warehouse, distribution and sales. In retail, distribution planning is of most importance representing, in many companies, one to two-thirds of the total logistics costs (Hosseini et al. (2014)).

Brick-and-mortar stores still represent the majority of retailers sales, despite the growing trend of online sales (Martins et al. (2018)). Besides, to ensure proximity to customers, retailers are opening more and more stores in new locations, with different characteristics such as a smaller store size (Martins et al. (2018)). These new stores must be included in the distribution planning. Moreover, there is a pressure to increase the delivery frequency while supplying smaller shipments (Martins et al. (2018)). These changing market dynamics are currently challenging distribution planning.

The focus of this project will be on improving the distribution planning of Company A. The analysis will be concentrated in their distribution center in the North of Portugal, which serves almost 200 stores spread over the North and Center of the country.

1.6 Problem statement and main objectives

The distribution plan from Company A's DC to the stores requires the use of a route optimization software. As an input, the software requires an estimate of the number of pallets that each store will receive, once the execution of the pallets is still ongoing.

Hence, improving this forecast is crucial. Flaws in the prediction of the number of pallets force the reorganization of trucks and their routes at the last minute, possibly generating the need for more trucks to deliver the goods. A more accurate prediction of the number of pallets that will be shipped would potentially reduce the transportation costs incurred by the company.

The main goal of the project is to find the best method to be implemented in warehouse operations to estimate the number of pallets that each store is going to receive. In order to achieve that two approaches are proposed. The first one is rather practical, consisting of a prediction method based on volumetry. The number of pallets to be shipped to each store is estimated according to each store's order that contains the quantity of each SKU requested by each store, and the volume of each SKU. On the second one, a more theoretical perspective was followed, testing several forecasting methods for time series.

These methodologies will be implemented in the non-perishable items warehouse, one of the five warehouses in the North of Portugal.

1.7 Dissertation structure

This dissertation is organized as follows. Chapter 2 gives a theoretical overview of time series and the forecasting methods used, and provides a literature review on distribution planning and forecasting methods applied to retail, namely in predicting demand. Later, in Chapter 3 the warehouse operations are described along with the implementation of the first methodology proposed. The analysis of the data set is presented in Chapter 4. Finally, Chapter 5 reports the implementation of the different forecasting methods for time series as well as the results obtained. The main conclusions of this dissertation are presented in Chapter 6 in addition to further work suggestions.

Chapter 2

Theoretical background and literature review

2.1 Considerations on time series

In forecasting, it is supposed that the future will behave like the past therefore time series analysis is extremely important.

A time series is a sequence of data points collected at constant time intervals. It can be decomposed into several parts:

- Trend: a long-term increase or decrease in the data, for example, if the dependent variable is on average growing over time;
- Seasonality: a variation according to specific fixed frequency pattern. For example, the dependent variable might tend to increase in certain weekdays, or months;
- Residuals: the difference between an observation and its predicted value at each time step.

Stationarity is an important characteristic to take into consideration when dealing with time series. A time series is stationary if its statistical properties such as mean, variance and autocorrelation remain constant over time. Most of the time series models work on the assumption that the time series provided is stationary. Therefore there are several methods to transform a non-stationary time series into a stationary one:

- De-trending: remove the trend component of the time series. The most widely used methods to estimate trend are aggregation, taking the average value for a particular time period; Smoothing, taking rolling averages and Polynomial Fitting, to fit a regression model;
- Differencing: the most commonly used method consisting of calculating the difference between values with a particular time lag;
- Decomposing: trend and seasonality are modeled separately and removed from the model.

2.2 Holt-Winters exponential smoothing method

Exponential smoothing was proposed in the late 1950s. Forecasts produced using exponential smoothing methods are weighted averages of past observations, with the weights decaying exponentially as the observations get older (Hyndman and Athanasopoulos (2018)).

The simplest of the exponential smoothing methods is simple exponential smoothing, applicable for forecasting data with no clear trend or seasonality. Here, the only component included is the level.

In 1957, Holt extended simple exponential smoothing to allow the forecasting of data with a trend. Both trend and level are included as components. The forecasts generated by Holt's linear method show a constant trend indefinitely into the future. Empirical evidence suggested that these methods tend to over-forecast, especially for longer forecast horizons. So a parameter that "dampens" the trend to a flat line some time in the future was introduced (Hyndman and Athanasopoulos (2018)).

Holt and Winters extended Holt's method to capture seasonality. Here, level, trend and seasonal components are included in the model and represented through three smoothing equations, with corresponding smoothing parameters α, β and γ. The Holt-Winters method has two variations, the additive model and the multiplicative one. The additive method is favored when the seasonal variations are approximately constant through the series, while the multiplicative method is preferred when the seasonal variations are changing proportionally to the level of the series (Hyndman and Athanasopoulos (2018)). The components for the additive and multiplicative models are represented in Appendix A.

2.3 Seasonal ARIMA model

In order to understand the Seasonal ARIMA (SARIMA) model, one should start by introducing the ARMA model. The AR stands for the auto-regressive part and consists of a linear combination of previous values of the series. On the other hand, the MA stands for the moving average part and uses past forecast errors in a regression-like model. The ARMA model is not applicable to non-stationary time series. To do that differencing is applied, in what is called the integration part of ARIMA (Hyndman and Athanasopoulos (2018)).

Since its introduction in the 1970s, the Box-Jenkins ARIMA has become one of the most popular methods for time series forecasting (Alon et al. (2001)).

In terms of notation, the ARIMA model is described as ARIMA(p,d,q) where p is the auto-regression order, d the difference order and q the moving average order.

The ARIMA model can be described according to Equation 1 from Appendix B.

Finally, SARIMA is an evolution of the ARIMA model that can incorporate seasonality.

In terms of notation, the SARIMA model is described as SARIMA(p,d,q)(P,D,Q)m where p is the trend auto-regression order, d the trend difference order and q the trend moving average order. Regarding the seasonal part, P is the seasonal auto-regressive order, D the seasonal difference order, Q the seasonal moving average order and m the number of time steps for a single seasonal period.

The SARIMA model can be described according to Equation 3 from Appendix B.

2.4 TBATS model

The TBATS model was presented by de Livera et al. (2011) when modeling time series with complex seasonal patterns. TBATS stands for Trigonometric seasonality, Box-Cox transformation, ARMA errors, Trend and Seasonal components.

This model has its origin in the exponential smoothing methods and tries to overcome its limitations in modeling complex seasonality. Namely, the overparameterization and the inability to accommodate both non-integer periods and dual calendar effects (de Livera et al. (2011)).

Thus, a trigonometric representation of seasonal components based on Fourier series is presented. This framework is able to model complex seasonality such as linear and non-linear time series as well as single and multiple seasonality, non-integer seasonality and dual calendar effects (de Livera et al. (2011)).

The forecast value results of a combination of level, trend, the ARMA process for residuals and the seasonal components. TBATS model equations can be found in Figure 48 from Appendix C.

2.5 Random Forest

Random Forest (RF) is an ensemble method, combining the result of multiple predictions delivered by several decision trees.

This machine learning algorithm can be used for both regression and classification problems. Some of its advantages include the fact it is simple to use, has few parameters to tune, its computationally efficient and avoids over-fitting, an issue for decision trees. Besides, it can be applied to a variety of prediction problems and has been applied with success to multiple practical problems due to its flexibility dealing with small sample sizes, high-dimensional feature spaces and complex data structures (Tyralis and Papacharalampous (2017)). Moreover, it can be applied to time series forecasting.

It is important to mention that Random Forest is a bagging technique instead of a boosting technique, as trees are run in parallel and there is no interaction between them while they are being built (Chakure (2020)).

To build each tree only a random set of training data points is considered as well as a random portion of the total features. These random selections originated the name Random Forest and are the reason why the model prevents over-fitting. The number of features to consider at each decision node is limited to some percentage of the total and constitutes a hyper parameter that can be tuned. Then each tree chooses the best features among the ones available to split the data. This process allows for the use of all potential features guaranteeing the model does not depend too much on an individual feature.

The RF regression algorithm, applied in time series forecasting, makes a prediction based on the average of the results of each decision tree. However, in classification problems the forecast outcome is planted based on majority voting rule.

2.6 Artificial neural networks

Neural networks (NN) are multi-layer networks of neurons used for both classification and regression problems. Its architecture is based on the human nervous system, in which each neuron has a set of both input and output connections.

Neural networks have an input layer, where information is received, and an output layer, where the solution of the problem is shown. Between them, one or several hidden layers of neurons exist.

Information flows from the input layer to the output layer through the connections between neurons. Inputs in every neuron have a weight associated. The output of each neuron is calculated using an activation function to the weighted sum of its inputs. This output is then passed on to another neuron. The weights of the network are computed by a learning algorithm, for example through the backpropagation algorithm.

Furthermore, an important aspect of Neural Networks is choosing the right architecture. This includes defining the number of layers, the number of units in each layer, the connections between units, the transfer function between the middle and output layers, a training algorithm and so on (Zhang and Qi (2005)).

2.6.1 Recurrent and Long Short Term Memory Neural Networks

A Recurrent Neural Network (RNN) is a neural network capable of capturing time or sequence-dependent behavior. It is widely used in text processing and for making predictions. The output of a recurrent neural network is dependent on all previous predictions and the information learned from them. This occurs since the output of NN layer is fed back as an input of the same NN layer for the next time step.

RNNs are able to work with short-term dependencies, remembering things for a small duration of time. However, they are unable to deal with long-term dependencies, when there is a need to recall something that occurred several steps before.

Long Short Term Memory Networks (LSTM) are a type of RNNs that overcome this shortcoming.

A LSTM network is composed of memory blocks named cells. In LSTMs information flows through a mechanism known as cell states. A cell state has three dependencies: the previous cell state, containing memory information after the previous time step; the previous hidden state, the output of the previous cell and the input information at the current step (Srivastava (2017)).

LSTM networks are able to modify the cell state. These operations are performed at the cells through a mechanism called gate. There are three different gates: the forget gate, the input gate and the output gate. The forget gate takes information regarding the output of the previous cell and the previous cell state, and decides which part of the information to keep and which part of the information to forget. The input gate is where new information is added to the cell. At last, the output gate selects relevant information from the current cell state and produces an output (Srivastava (2017)).

2.7 Literature review on forecasting

Forecasting is the process of predicting the future based on past information. It plays a key role in decision making as it tries to foresee how uncontrollable variables, relevant for decision making, will behave in the future. Often companies have to plan ahead as decisions rarely have immediate effects. In order to do this properly and improve decision making, companies must use forecasting to evaluate future conditions, anticipate change and reduce uncertainty.

Forecasting has numerous applications in distinct fields of operation, namely in predicting stock prices, commodity prices, product demand, the weather and so on.

2.7.1 Application of forecasting in retail

A retailers' formula to success is to offer the right product, in the right place and for the right price, at the right time. Thus, forecasting has a key role in organizational performance as it can potentially improve and support many aspects of the retail supply chain. Forecasting can be helpful at a retailers' strategic level, particularly on decisions regarding store and warehouse location and size, the distribution network or marketing, for example, when defining the target market segment and positioning (Ma et al. (2018)).

At a tactical level, forecasting is useful when predicting the demand of each store for a product. This influences decisions concerning workforce planning and scheduling, and distribution planning, both fleet and routing. An accurate product demand forecast is crucial for achieving a high service level. Moreover, it should be integrated with the inventory management and distribution systems (Ma et al. (2018)).

Lastly, at an operational level forecasting helps planning and organizing daily activities such as purchasing, distribution and labor force management (Ma et al. (2018)).

Forecasting avoids excessive inventory and so, it reduces inventory holding costs and extra costs due to obsolete items. Furthermore, distribution can be more efficient and effective.

Retail forecasting depends on the level of aggregation. The aggregation can be in product units, such as SKU, brand and category, location, for example, store or market level, and time units, daily, weekly, monthly and so on.

There are several aspects that impact store aggregated sales such as store location, presence of competitive retailers, consumer demographics, promotions (held by the store or by competitors), weather, seasons, holidays and local events (Ma et al. (2018)).

2.7.2 Statistical forecasting methods

Forecasting methods can be qualitative or quantitative, being the focus of this dissertation the quantitative ones. Qualitative methods are usually used when there is no data available, so they are made based on intuition, experience or researches.

The traditional forecasting approach, through the use of statistical forecasting methods like Holt-Winters Exponential Smoothing, ARIMA models or regression is still widely used.

The traditional methods are essentially linear methods and to use them one must specify the model form without necessarily capturing the complexity existent in the data. They have the advantage of being relatively easy to implement and interpret, while at the same time being able to perform well (Chu and Zhang (2003)).

To overcome some limitations of Holt-Winters and Arima models, authors are proposing different methodologies.

For instance, de Livera et al. (2011) proposed a different approach to handle time series with complex seasonal patterns. It is suggested a trigonometric formulation capable of dealing with seasonal complex patterns, identifying and extracting seasonal components that otherwise would not appear in the time series plot. The advantage of this framework is the ability to model linear and non-linear time series as well as single and multiple seasonality, non-integer seasonality and dual calendar effects. As a result, the model outperformed the traditional exponential smoothing methods incapable of modeling both non-integer period and dual calendar effects.

Additionally, Bratina and Faganel (2009) modeled the primary demand for beer in Slovenia using ARMAX, which proved to be superior to ARMA methods. The demand for beer increases during Summer and special occasions like New Year. Thus, these effects were modeled with outside temperature and using dummy variables, respectively.

Likewise, Arunraj et al. (2016) demonstrated that the use of external variables enhances the performance of SARIMA model when forecasting the daily sales of banana. The SARIMAX model is implemented fitting the time series using SARIMA and a multiple linear regression model, where the errors of SARIMA model are modeled by independent variables like holidays, promotions and months. In this case, the external variables were important to explain outlying data.

2.7.3 Machine learning models applied to time series forecasting

The application of machine learning methods in forecasting for time series is seen as an alternative to the traditional forecasting methods.

Machine learning algorithms are able to learn from the data, identify patterns and trends and make an inference about the future. They are able to simulate a wide variety of non-linear behaviors and approximate almost any function. Moreover, the model is built through data mining, so few assumptions about its form need to be made (Chu and Zhang (2003)). The flexibility and adaptability of these non-linear methods is what makes them an interesting forecasting tool.

However, machine learning methods are computationally more demanding and difficult than statistical ones, and require a greater dependence on computer science, special software and in-house expertise to be implemented (Makridakis et al. (2018) and Alon et al. (2001)).

The use of ensemble methods like Random Forest and Gradient Boosting constitutes a viable option for time series forecasting. Jain et al. (2015) applied the Extreme Gradient Boosting algorithm to the prediction of sales for the retail outlets of a pharmaceutical company. The inputs of the model were both temporal features, previous sales data, and economical features like promotions, holidays, location and presence of competitors. Its performance proved to be superior when compared to other algorithms like Linear Regression and Random Forest.

Besides, Tyralis and Papacharalampous (2017) also applied Random Forest to time series forecasting providing a framework for variable selection. They concluded that Random Forest achieved better results when using a small number of recently lagged predictor variables.

Furthermore, Neural Networks' application to time series forecasting is also being broadly explored. Zhang and Qi (2005) studied the effect of data preprocessing on neural networks forecasting performance. They found that Neural Networks were not able to capture trend and seasonality properly and that detrending and deseasonalization were the most effective preprocessing method, reducing significantly the forecasting error. Moreover, they compared the performance of Neural Networks with the seasonal ARIMA model, NN outperforming seasonal ARIMA.

Nevertheless, there are mixed findings in the literature regarding whether or not data preprocessing improves Neural Networks forecasting accuracy and what processes should be applied. Some studies concluded that Neural Networks are able to model seasonality directly and others the opposite. Modeling seasonality appears as a particularly difficult issue for every forecasting model.

2.7.4 Comparison between statistical and machine learning models applied to time series forecasting

Several studies comparing the performance of Neural Networks and other machine learning models with the traditional models have been conducted and yet findings are not consensual in determining the best model.

Chu and Zhang (2003) compared the forecasting accuracy of aggregate retail sales obtained by several linear and non-linear models. Three linear models were tested: ARIMA, regression with dummy variables and regression with trigonometric variables. At the same time, the non-linear version of these models was implemented through Neural Networks. As retail sales show clear seasonal variations, the aim was to access whether or not the use of seasonal dummy variables or trigonometric variables improved the forecasting accuracy. As a result, dummy variables were found to improve the forecasting accuracy of regression models, although the models are not robust nor consistent. The trigonometric variables were not helpful. The overall best model was the neural network built on deseasonalized time series data, without the use of dummy nor trigonometric variables.

Furthermore, Alon et al. (2001) compared the performance of artificial neural networks, Holt-Winters model, ARIMA model and multivariate regression in predicting the US aggregate retail sales, a time series with distinct trend and seasonality. As a conclusion, ANN performed better than the traditional methods, followed by ARIMA and exponential smoothing. Neural networks were able to capture the dynamic trend and seasonal patterns and their interactions, being able to outperform the traditional models when macroeconomic conditions were volatile. When macroeconomic conditions are stable ARIMA and Holt-Winters models proved to be a viable option.

Alternatively, Di Pillo et al. (2016) applied support vector machines (SVM) to forecast sales under the effect of promotions. Their forecasting used 13 input attributes, 9 calendar attributes and 4 problem specific attributes. One of them a boolean variable representing whether or not the product was under promotion on a particular day. As a benchmark, statistical models like ARIMA, exponential smoothing and Holt-Winters were used. The result obtained was that SVM provided an applicable alternative to the statistical methods obtaining lower forecasting errors. SVM was able to capture part of the variability caused by promotions whereas traditional methods were not. Moreover, the article finds SVM a fair competitor to neural networks in forecasting time series.

Contrarily, Makridakis et al. (2018) compared the out-of-sample accuracy of ten machine learning methods with that of eight traditional statistical models. They observed that the statistical models outperformed the machine learning ones for all the forecasting horizons tested. Additionally, they demanded greater computational requirements than statistical methods. Once again, it is suggested to deseasonalize data before using a machine learning algorithm. Furthermore, they believed that the full potential of machine learning was not yet achieved and that further research is needed to enhance the models.

When estimating the aggregate retail sales in South Africa, Aye et al. (2013) used 17 models with seasonal dummy variables and 9 full seasonal models. The models with seasonal dummy variables outperformed the full seasonal ones. Moreover, non-linear models obtained better results than the linear ones. However, the authors highlighted the importance of including several models as each model captures different characteristics of the time series for different forecasting horizons as well as for different economic cycles.

Finally, Pavlyshenko (2019) states that sales prediction is rather a regression problem than a time series problem, and that better results can usually be achieved applying regression instead of time series methods. Some drawbacks of time series algorithms are that they rely on a long period of historical data to be able to capture seasonality. And this might not always be available. Besides, exogenous factors need to be taken into consideration as they impact sales significantly.

To sum up, traditional forecasting models like ARIMA and Holt-Winters are still commonly used by companies and produce viable outcomes. At the same time, machine learning algorithms are becoming a credible alternative to these methods.

2.8 Operations at the distribution center

The operation at the warehouse can be classified as JIT (just in time) or Stock. The Stock operation represents the common warehouse where products are kept until they are part of an order from a store. Then, they are picked according to the store order and a shipment is constructed. On the other side, the JIT operation does not have stock or has a considerably low storage time. It works as a cross-docking operation where products are received from the suppliers in one day, the shipments to the stores are built on the same day and finally, they are dispatched to their final destination on that same day or the following.

Warehouses that handle perishable products such as fruit and vegetables, fish and fresh items work as JIT operation. These products have short shelf lives and are perishable thus they demand a more frequent delivery to the store and very short storage time.

Additionally, a non-perishable products warehouse can operate with the JIT principle. The main advantage here is the reduction of inventory at store level. Normally, suppliers with high service levels and flexible and frequent deliveries are chosen to work with the JIT operation. Fast-moving consumer goods are also the type of products processed within this principle.

Alternatively, frozen and non-perishable products warehouses can work as a Stock operation.

Distribution centers, working either as Stock or JIT operation, allow for more frequent deliveries to stores as they take advantage of economies of scale for both inbound and outbound vehicles (Martins et al. (2018)). Moreover, transportation costs are reduced as a consequence of the enhanced vehicle utilization.

When compared to Stock, cross-docking also allows for a cost reduction in both inventory holding costs and warehousing costs, shorter delivery lead times, enhanced customer service, faster inventory turnover and reduces the risk of loss or damage (Van Belle et al. (2012)). On the other side, it is a complex process that requires the full coordination of suppliers, the DC and the outbound vehicles.

Regarding the transportation of goods from the DC to the stores, retailers have some strategic decisions to make. For instance, whether to own their own fleet or to outsource transportation to a third-party logistics provider (3PL).

Currently, the consumer goods industry relies on a high share of outsourced transportation services (Günther and Seiler (2009)). The high demand seasonality, high number of promotions, products with shortening life cycles and constant product innovations are some of the characteristics of the industry. This combined with the above-mentioned current market dynamics like the demand for a faster replenishment, higher service levels, smaller shipments and more frequent deliveries urges the need for flexible and reliable transportation process (Günther and Seiler (2009)). 3PL can offer a more flexible cost structure while lowering the total transportation expenses. Even so, retailers want to manage and control how distribution is executed, due to its major influence on their performance (Martins et al. (2018)).

Chapter 3

Description of the operation and current practices

The focus of this project will be on the operation of the warehouses in the North of Portugal, responsible to meet the demand of almost 200 stores in the North and Center regions of the country. For that purpose, this food retailer possesses five warehouses in the North of the country characterized by the type of products they hold and the storage conditions required: non-perishable items, fruit and vegetables, fish, fresh items and frozen items.

As mentioned in the introductory chapter, consumers are demanding higher service levels and a great variety of products available at the stores. Consequently, there is a large pressure on the supply chain that has to be capable to respond efficiently for consumers needs to be fulfilled. Figure 1 shows how this supply chain is organized.

Figure 1: Supply chain organization

Usually, goods are shipped to the distribution center by suppliers where they are stored or immediately prepared to be shipped to the stores. At the store, products are replenished at the shelves and sold to the final consumer.

3.1 The JIT Operation

The non-perishable items warehouse includes JIT and Stock operations. The JIT Operation can be divided into four different steps: reception, execution, filming and expedition, represented in Figure 2. One of the main differences from the Stock operation is the execution mode, in this case picking by line. In the JIT layout, each store has a predetermined position on the ground. Then, pickers go around the layout distributing each product to the corresponding store according to the order made by each store.

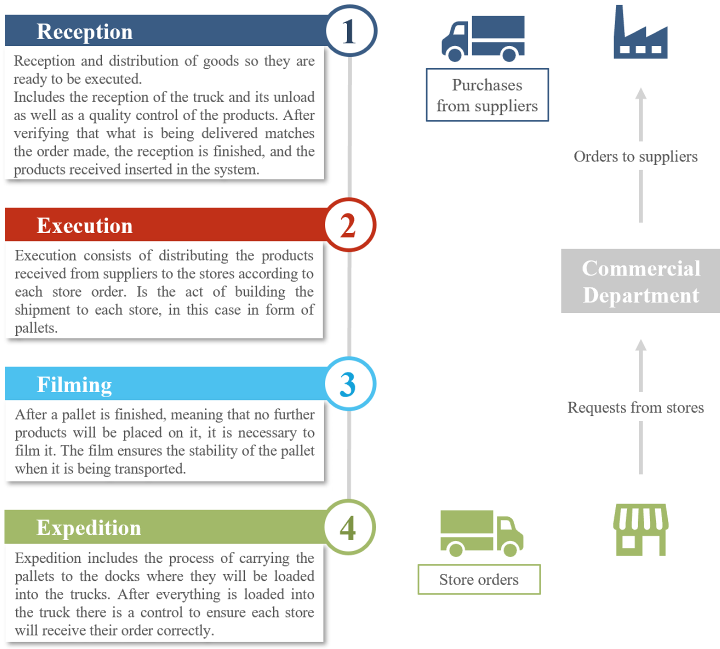


Figure 2: Description of JIT warehouse operations

3.2 The Stock Operation

A scheme of the operation is portrayed in Figure 3. The Stock operation is composed of five steps, most of them common to the JIT operation with the exception of the "Storage and lowering" step. In Stock, inventory is held in racks with multiple levels of height. When goods are received from suppliers, they are stored on the higher levels of the racks. The ground level of the racks is the picking zone, where pickers collect the products to be shipped to a store. In this case, the execution mode is picking by store, as pickers are building a pallet that is destined to a single store according to the order of that store. They move through the layout to each product's location, collect the right amount of the product, and continue this process until they finish fulfilling the order.

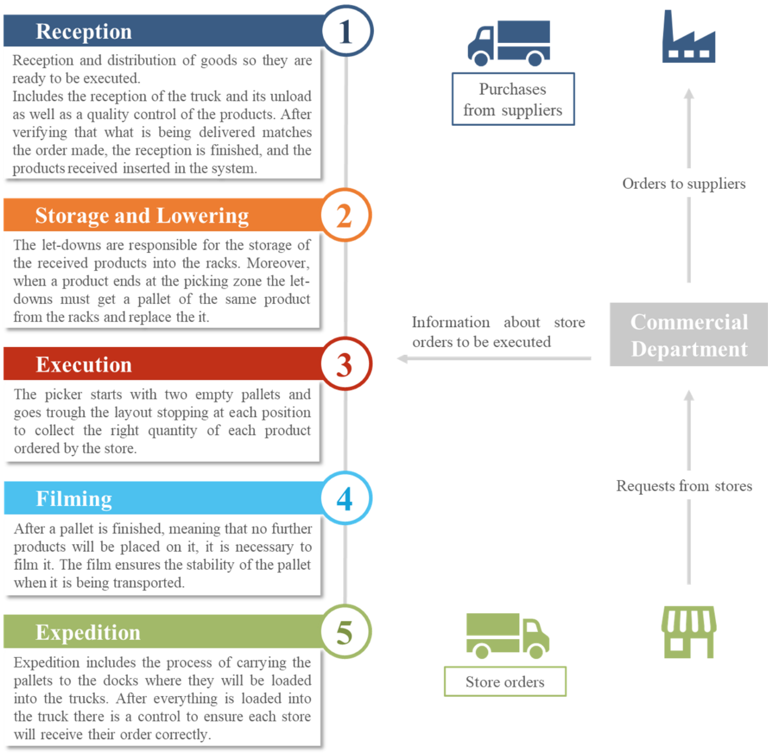


Figure 3: Description of Stock warehouse operations

3.3 Data analysis at the warehouse level

To get an overview of the warehouse operation and its scale, a first analysis of the total number of pallets shipped per day was performed, and it is represented in Figure 4. From this plot, one can conclude that there are some variations of the daily number of pallets shipped throughout the year, being the average number around 4000 pallets shipped per day.

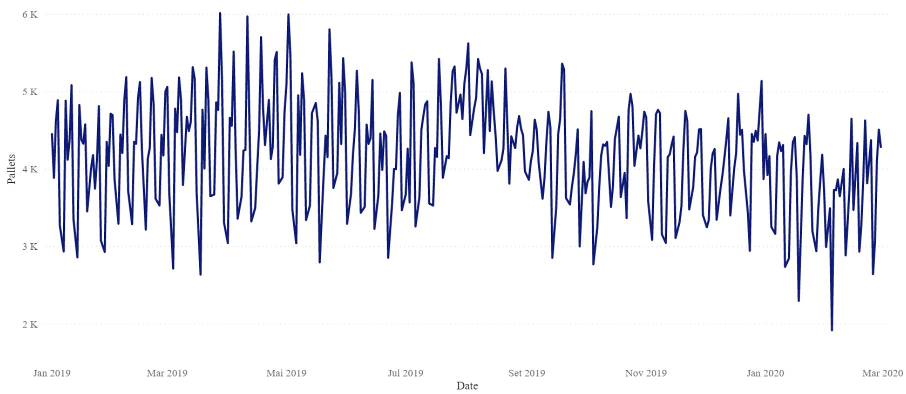


Figure 4: Total number of pallets distributed by the warehouse per day between January 2019 and March 2020

Figure 5 confirms this variation, with the average number of pallets shipped per day and per store having its peak in August. The months of March until May are also a period with a higher number of pallets shipped. After its peak in August, the expedition of pallets decreases and remains lower for the last quarter of the year.

A similar analysis was conducted calculating the average number of pallets shipped per store per weekday (Figure 6). It is shown that Mondays (2) and Saturdays (7) are the weakest days, with a considerable difference to the strongest days Thursdays (5) and Fridays (6). This difference might be explained by the fact that sales are expected to increase during the weekend, so it is natural that shipments to stores reach their peak on the days before it.

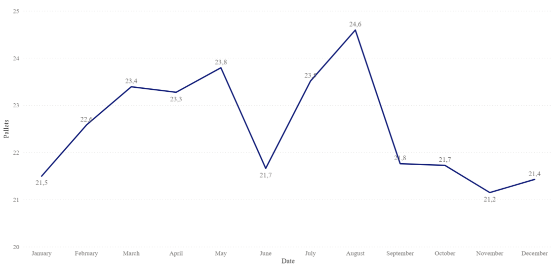


Figure 5: Average number of pallets daily shipped per store grouped by month in 2019

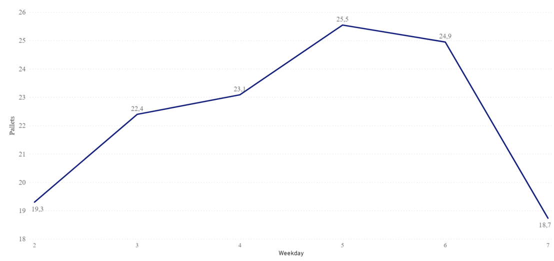


Figure 6: Average number of pallets daily shipped per store grouped by weekday in 2019

Figure 7 and Figure 8 show the coefficient of variation aggregated by month and by weekday, respectively. The coefficient of variation has slight monthly variations, being September the month with the most stable number of pallets shipped per day. On the other side, December is the month with more fluctuations in shipment size. Moreover, the coefficient of variation remains rather constant for weekdays, having its peak on Saturdays. It is relevant to mention that there is a significant dispersion of the number of pallets shipped around the mean.

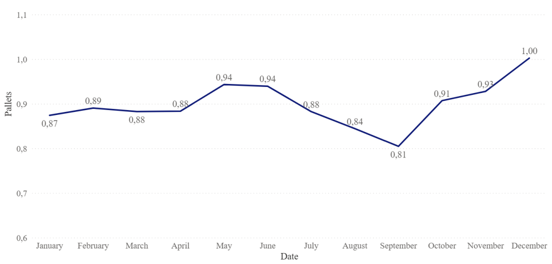


Figure 7: CV of the number of pallets daily shipped per store grouped by month in 2019

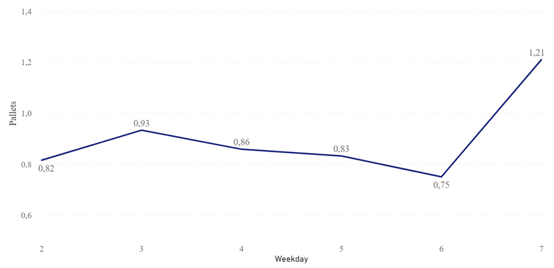


Figure 8: CV of the number of pallets daily shipped per store grouped by weekday in 2019

The top 20 stores with the highest average number of pallets shipped per day are represented in Figure 49 from Appendix E. As expected, the top 3 stores represent hypermarket stores. Moreover, there are 5 wholesaler stores in the top 20 (2007, 2002, 2028, 2037, 2003). It is possible to conclude that the first six stores have a much higher number of pallets shipped than the others. As the number of pallets shipped to each store reflects its demand, these stores are crucial to the business and seem to represent a significant portion of sales. The top 6 stores are responsible for approximately 13,5% of the total number of pallets shipped.

Finally, an ABC-XYZ analysis was performed. Usually, this analysis is performed to determine what to forecast. In this case, it was used to find the stores with the highest volume of pallets shipped, and mostly to assess if their demand pattern was irregular or, on the contrary, if it was rather constant. The result of this analysis is presented in Table 1.

It is shown that most of the stores are type X (82%), meaning they have a constant pattern in the number of pallets shipped per day. This leads to higher forecast accuracy. Y stores, on the other side, have stronger fluctuations in the number of pallets shipped per day, being more difficult to predict than X stores. Finally, Z stores have irregular patterns leading to low forecast accuracy. Fortunately, there is only one type Z store. One should prioritize the forecast of the stores with the highest volume of pallets shipped and with the most constant patterns, as they represent a significant part of the business and they are easier to forecast, bringing remarkable results. A scatter plot of the stores organized by groups is represented in Figure 50 from Appendix F.

Table 1: ABC-XYZ analysis

	X	Y	Z
A	14%	6%	0%
B	23%	7%	0%
C	45%	4%	1%

3.4 Transportation between distribution center and stores

In Company A distribution is outsourced to a specialized company. The distribution planning and execution is currently a joint task between the two companies.

There are two types of vehicles responsible for delivering products to the stores. The first is a tractor, also known as tractor-trailer truck, where the frontal part of the truck decouples from the trailer (the part where the goods are loaded into) and usually has a capacity for 33 pallets. And the second one a rigid, both parts of the truck are inseparable and has a cargo capacity that can vary between 12 and 22 pallets. Some stores are limited to rigid vehicles as they don't have the capacity to unload large vehicles.

Regarding the delivery routes, vehicles can go directly from the distribution center to the store, they can stop to load at different warehouses and then move on to the store or, finally, they can carry goods from a single warehouse and deliver them to multiple stores. Sometimes a combination of the last two delivery methods happens, and the same vehicle delivers products from different warehouses to multiple stores.

Moreover, transportation and routing are planned taking into account the delivery window established with each store.

Presently, the Company uses a route optimization software that runs daily and determines the routes for the following day. It takes into consideration the restrictions and possibilities mentioned above. The main input given to the software is the number of pallets each store is going to receive from each of the warehouses. Furthermore, it is relevant to mention the use of backhauling to bring products from suppliers to distribution centers. It avoids trucks being empty on their return trips, thus saving in transportation costs charged by suppliers. This is not taken into consideration by the route optimization software, being programmed afterwards.

Lastly, there are some exceptions regarding distribution, with some suppliers delivering their products directly to the stores, representing the minority of the cases.

3.5 Distribution planning in the non-perishable items warehouse

The focus of this dissertation will be on improving the distribution planning of the non-perishable items warehouse. It is the largest warehouse, with the highest number of pallets shipped per day.

This warehouse is divided into Stock and JIT operations, each of them having two different layouts - one for non-food products and other for food products. A pallet is constructed with products exclusively from one of the layouts, being impossible to mix products from different layouts.

Figure 9 represents a timeline of the relevant moments for distribution planning. First, the request made by each store becomes available on the day before being executed at the warehouse. The expedition of this same request occurs during the day after being executed.

During the execution of pallets is necessary to plan and schedule the fleet and the drivers for the next day. In order to do so, the transportation department needs an estimate of the number of pallets each store is going to receive from each warehouse. This estimate will be inserted in the route optimization program, determining the number of pallets each truck will carry and the stores it will visit. Transportation planning is performed at 11:00, so by this time all warehouses should have sent their estimates. Since the software takes about two hours to run, it is impossible to delay this planning.

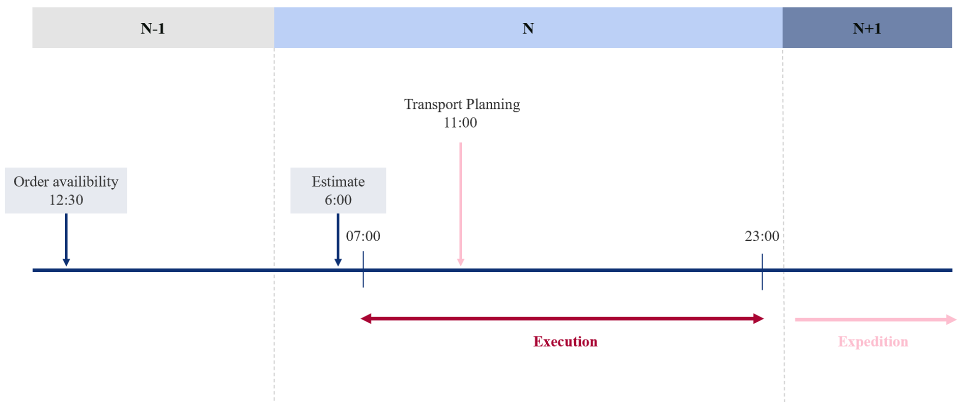


Figure 9: Timeline of distribution planning operations

The accuracy of this estimate is crucial, as errors result in route and cargo reorganization upon departure time.

The methodology that had been used to perform the estimate at the warehouse was using the total number of boxes assigned to each store in the request, and then calculate the number of pallets based on a ratio of boxes per pallet, defined for each warehouse.

The first methodology proposed, as previously mentioned, is based on volumetry. Given the request from each store, in the form of quantity of a given SKU per store, pallets are built according to the volume of each SKU and the total volume defined for a pallet. The aim of this procedure was to be simple and yet effective, thus it was implemented using Excel spreadsheets. The employees responsible to perform this estimate work in the administration department at each warehouse.

There are several parameters to be provided to perform the estimate. Namely, product measurements, available for the products with higher rotation, and a standard volume, to be used when the product in question is not measured. Moreover, it it necessary to define a pallet volume. The base area of a pallet is considered to be constant and equal to 120*80 cm. On the other side, pallet's height is dependent on the layout considered since product characteristics such as shape, size and weight differ across layouts. The percentage of air in a pallet is also considered. Finally, a percentage of zeros is defined. Zeros are suppliers' failures to deliver their scheduled products. This parameter is only considered in the JIT operation, where products arrive on the day or a day before they are executed. In Stock, this percentage is not relevant as the request from the stores is made knowing what products are available. The percentage of zeros is applied to reduce the total request.

These parameters were initially defined based on the staff's expertise. For example, the percentage of zeros was considered as an average supplier failure rate and it remained constant during every weekday. Moreover, the height of each pallet was set based on what is usually observed at the warehouse.

It is relevant to mention that, if an order of a product (or several products from the same supplier) placed by a store has a relevant size, a full pallet or half pallet can be delivered by the supplier. These pallets are allocated directly from the unloading dock, or the rack if it is in Stock operation, to the store in question, without being executed. These orders are registered differently by the system, being possible to distinguish them from the aforementioned regular orders, where pallets are built combining several SKUs.

So, in order to perform a daily forecast, the estimation tool needs to run daily with the following inputs: information downloaded from the information system containing the quantity of each SKU ordered by each store and the number of complete pallets or half pallets to be allocated directly to each store.

This information is organized in the system by warehouse. Besides, each product is executed in a single and predetermined layout, also available in the system. Consequently, the forecasting tool is able to compute the number of pallets per layout and per store, providing an accurate vision of reality. The total number of pallets to be shipped by the warehouse to each store is then calculated aggregating the number of pallets per layout.

Having these parameters and inputs, the model is ready to deliver its output. To sum up, the model computes the total number of pallets following the steps described below:

1. Takes the information regarding the quantity of each SKU ordered by each store and its respective volume, and computes the total volume to be delivered per store and per layout;
2. Applies the percentage of zeros to the total volume per layout per store;
3. Calculates the number of pallets per layout, using the pallet volume parameterized;
4. At the same time, it computes the total number of complete or half pallets to be allocated directly to each store;
5. And finally, the model calculates the total number of pallets to be shipped to each store by the warehouse adding the number of pallets per layout and the number of complete pallets.

3.5.1 Implementation of the forecasting tool

In the non-perishable items warehouse, the forecasting tool was implemented in the beginning of March. The forecasting tool was also implemented in two other warehouses and it is currently being implemented in another one. The forecasting tool is standard for all warehouses, with only a few particularities of each warehouse being added to the tool when needed.

The implementation of the methodology occurred in a similar way for all warehouses. A thorough explanation of the important steps to be followed in order to perform the estimate was given on-site to the operators responsible to perform it. Moreover, a standard with the key steps to be followed was also left on site. An example of the standard is represented in Appendix H.

During the first weeks, the execution of the estimate was accompanied daily to ensure that there were no errors filling in the inputs on the tool and also to answer any doubts. The performance of the tool is highly dependent on the inputs given to it, so ensuring it was done correctly was key.

The performance of the tool was evaluated daily during the first months. The wMAPE was the performance indicator chosen to overcome the fact that stores have considerably different sizes, and therefore, there are strong variations in the number of pallets shipped daily. wMAPE calculation formula is presented in Equation 4 from Appendix D.

3.5.2 Improvement of the forecasting tool

3.5.2.1 Hypothesis of performing a second estimate

The hypothesis of performing a second estimate was tested. This estimate would be performed at 16:30, and therefore, a second transportation planning would take place. It is important to mention that a first transportation plan is mandatory, thus a second estimate would not be able to replace the first one.

The motivation behind this was the fact that when doing the first estimate the execution of pallets has not started, while if a second estimate was performed there would be information available about the number of pallets already executed, thus uncertainty would be reduced.

When a product is allocated to a pallet, it is marked as executed in the store's request present in the system. Moreover, every time a pallet starts to be built it is registered in a different sheet in the system. So, the second estimate would simulate the construction of pallets with the products that were still not executed, similar to what is done in the first estimate. Then, it would take an extra input containing the information regarding the pallets executed or under execution. The number of pallets to be executed would be added to the ones already executed.

For the warehouse in question, this approach was abandoned, as it did not consistently provide better results when compared to the first estimate. The main reason behind this is that the system can not distinguish the pallets that are already finished from the ones under construction. So, by considering every pallet in the system as finished, the tool would overestimate, as a portion of these pallets would still be under construction, with more products still being added to them. To overcome this, the pallet with the lowest height registered for each store was subtracted from the estimate, if the lowest height was inferior to a certain threshold. This improved the estimate, but it was still not enough to provide results that justified the realization of a second estimate.

3.5.2.2 Analysis of the parameters defined and definition of the parameter update process

The parameters initially defined in the tool were analysed and simulations were performed to assess whether or not results could be improved. The height of each pallet per layout, the standard volume, and the percentage of zeros to be considered were the parameters tested. The percentage of zeros was initially considered constant throughout the week, but it was considered relevant to test if differences between weekdays and layouts should be considered. Two weeks of past estimates were tested, corresponding to 12 days.

Resorting to Excel VBA the parameters could be altered and then Excel would automatically re-run the estimates for those 12 days with the updated parameters, evaluating the performance of the new model in the end. Then the results were compared to the actual results obtained for those twelve days to see whether the performance had been improved.

With Excel Solver a simulation to find the parameters that minimized the forecasting error of those 12 days was performed. The upper and lower limits of each parameter were carefully defined.

These procedures were able to enhance the performance of the tool for the 12 days under analysis. Nevertheless, there is a possibility that these parameters can be overfitted for these days. A similar analysis was performed for the remaining warehouses.

Currently, a process for reviewing and updating the parameters is being drafted. The tool should be maintained autonomously by a defined person, with knowledge of the operation. There should be a continuous improvement of the tool. The plan is to establish which parameters should be updated, the update frequency and how to update them. For example, product measurements should keep being added to the tool. On the other side, parameters such as the percentage of zeros, standard volume and the pallet height to be considered don't need to be updated so frequently.

Moreover, these parameters should be updated based on data or an Excel solver procedure similar to the one mentioned above. For instance, the percentage of zeros could be updated based on an analysis of historical data and the standard volume using the median volume of products.

Chapter 4

Methods

4.1 Definition of the forecasting problem and framework

In what regards the forecasting framework, one should start by determining what to forecast. In this dissertation, forecasting is applied to an operational problem, aiming to support the retailer's daily distribution planning. As mentioned before, the retailer's distribution planning and scheduling is performed resorting to a route optimization software. As an input, the software needs an estimation of the number of pallets each store is going to receive. Therefore, the forecasting goal is to predict the number of pallets that are going to be shipped to each store on the following day.

In this case, a two-day step forecast is used, as by the time the plan for the next day is executed, the distribution planned for the present day is still undergoing. Hence, there is no information regarding the real number of pallets shipped to every store until the end of the day, not in time to be used to predict the shipments of the following day.

The following step of the forecasting framework is data analysis, required to identify patterns in the data.

Finally, the forecasting model is selected and adapted and the forecast system implemented.

4.2 Data preprocessing

The dataset was provided by the retailer's transport department and contains daily information regarding each truck's lap. A lap is composed of trips from one or multiple warehouses to one or multiple stores. Each trip is characterized by several features being the relevant ones for this case the following: warehouse of origin, destination store, number of pallets carried and date.

From this dataset, it is possible to extrapolate the number of pallets shipped to each store by the warehouse under analysis. This information is available for the year 2019 and the first two months of 2020.

The warehouse does not operate on Sundays, thus these days were removed from the data set. Consequently, each week of work is considered to have six days. Moreover, there are only two more days of the year when the warehouse does not operate, Christmas day and New Year's day. For these days the solution was to replace the missing data by the average number of pallets shipped on the homologous days to each store, in this case on 2019's Wednesdays.

Finally, there are some particularities regarding a few stores. Three stores opened during 2019, store 533 opened in March, store 534 in October and store 860 in the end of November. For these stores the data set starts from the point where the pallets shipped stabilized, being the initial large shipments to these stores removed.

On the other hand, nine stores were closed for remodeling during different times of the year for a period of time that can vary from one to two months. For these stores, the solution was either replacing the missing data for the forecasted values or only consider the data set from the point where the store re-opened, again starting from the moment where shipments stabilized. Each store was individually analyzed and the missing values replaced by the best solution found.

4.3 Explanatory variables

Multivariate models for time series forecasting can be built using machine learning algorithms. These models can include both the lagged observations of the output variable and external variables, that would help explaining the results. The explanatory variables used in the implementation phase of the machine learning models are represented in Table 2.

Table 2: External variables considered

Variable	**Description**
Month	Categorical variable representing the month of the year (from 1 to 12)
Weekday	Categorical variable representing the day of the year (from 2 to 7)
Day of month	Categorical variable representing the day of the month (from 1 to 31)
Holiday	Binary variable stating whether or not the following day is a state holiday. As it is pallet distribution that is being modeled the holiday effect on demand could be reflected on the days prior to the holiday
Christmas	Binary variable representing the week before and after Christmas day
Easter	Binary variable representing the week before Easter day
Number of boxes	Numeric variable displaying the number of boxes requested by each store in each store order
Number of pallets daily shipped by the warehouse	Numeric variable to assess the total volume of operations

4.4 Clustering

K-means clustering was performed in order to divide the stores into smaller groups with different behavior.

K-means is a simple partitional method that splits data into K clusters. In clustering, one expects to minimize the distance between points in the same cluster, while maximizing the distance between clusters. Thus, points in the same cluster are more similar to one another and different from points on a separate cluster.

In this case, stores were grouped according to the average number of pallets per order, representing the size of the store, and the average number of orders placed per month, representing the frequency of orders.

As a result, stores were divided into three clusters represented in Figure 10. The number of clusters was optimized minimizing the Davies-Bouldin Index, varying the number of clusters from 2 to 20. As a result, cluster 1 contains the stores with the lowest average number of orders placed per month, in this case, mostly wholesaler stores. Cluster 2 aggregates the top 5 stores with the highest average number of pallets per order, constituted by hypermarket stores. At last, cluster 3 includes mainly supermarket stores, which place an order almost every day.

Clustering was motivated by the fact that there are 190 stores, making it unfeasible to conduct a thorough analysis for each one of them. Thus, a store from each cluster will be chosen to be examined. Henceforward, store 657 will be representing the stores with the highest sales, hypermarket stores, store 2022 the wholesaler stores and store 347 the supermarket stores.

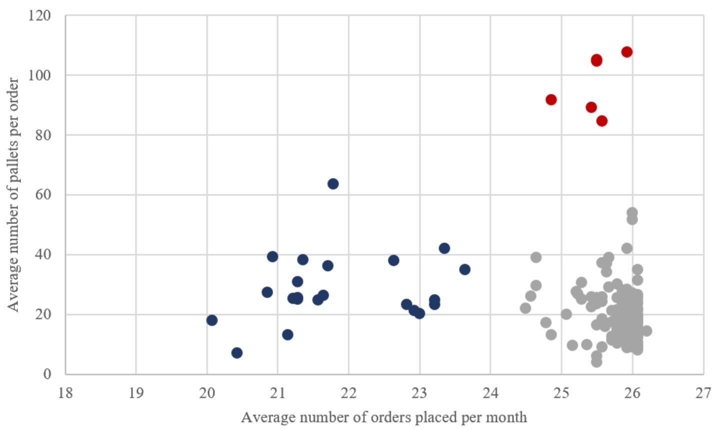


Figure 10: Scatter-plot representing the stores organized by cluster

4.5 Analysis of the store representing hypermarket stores

The initial warehouse analysis showed the aggregated behavior of the pallet shipments made by the warehouse. Therefore, a more thorough analysis was required, as this behavior might not be common to every store.

Figure 11 shows the total number of pallets shipped per day to the hypermarket store. This store receives an average of 107 pallets a day only from the non-perishable items warehouse. The high variation in shipment size is something to be highlighted.

Moreover, the variation of the daily number of pallets shipped to the hypermarket store throughout the year appears to be less pronounced when compared to the aggregated analysis (Figure 4).

The number of boxes shipped to the hypermarket store is presented in Figure 12 and shows a seasonal pattern similar to the number of pallets shipped daily to the same store.

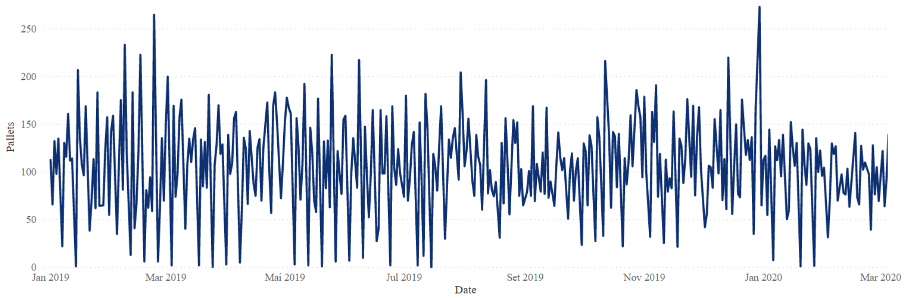


Figure 11: Number of pallets shipped daily to store 657 between January 2019 and March 2020

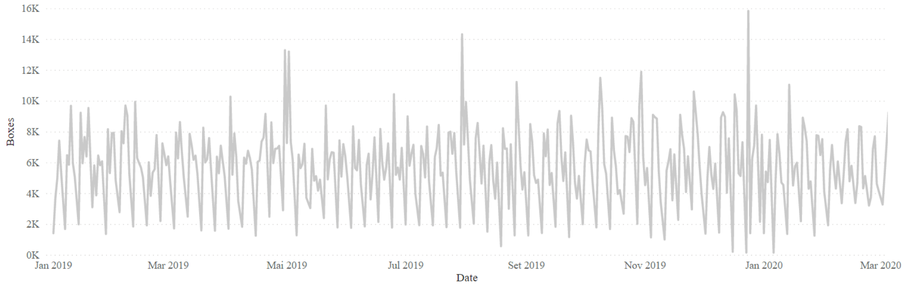


Figure 12: Number of boxes shipped daily to store 657 between January 2019 and March 2020

Table 3 represents the descriptive statistics of the dependent variable number of pallets and the numeric explanatory variables considered for the hypermarket store.

Table 3: Descriptive statistics of the dependent and explanatory variables considered for store 657

	Number of pallets	Number of boxes	Total number of pallets
Minimum	0	0	1920
Median	108	5980	4264
Mean	107	5845	4207
Maximum	273,5	15860	6016
Standard deviation	50,14	2533,11	700,46
Coefficient of variation	0,469	0,433	0,166

Furthermore, weekly seasonality is visible in Figure 13 with Monday being the weakest day. However, it is the day with the highest variation regarding the size of the deliveries. Contrarily to what was noticed in the warehouse operation analysis, Saturday is the day with the highest average number of pallets shipped.

Figure 14 shows the coefficient of variation per weekday. It confirms that Monday is the day with the highest variation in terms of shipment size, whereas for the rest of the week the coefficient of variation shows minor changes.

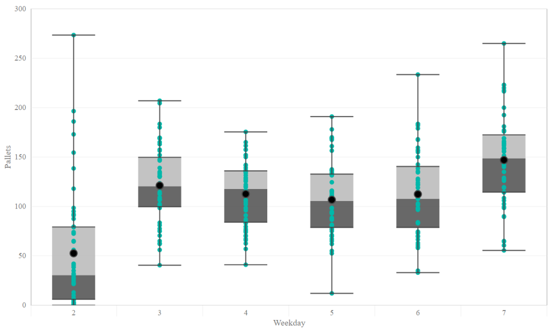


Figure 13: Number of pallets shipped daily to store 657 per weekday in 2019

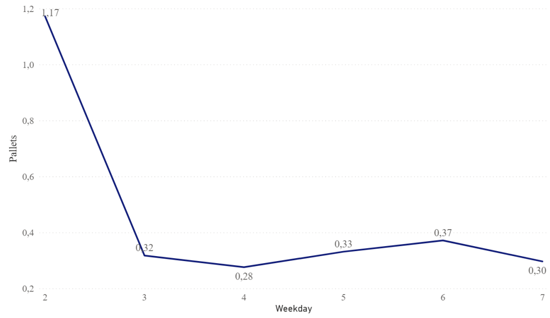


Figure 14: CV of the number of pallets shipped per day to store 657 by weekday in 2019

Additionally, the decomposition of the time series is represented in Figure 15. The seasonal, trend and residual components are shown. There is a visible seasonal pattern that remains constant over time. Moreover, the trend component shows some variation over time, but without a clear pattern. Hence, there is some fluctuation in the number of pallets shipped throughout the year. The residual component shown in the bottom panel is what is left over when the seasonal and trend components have been subtracted from the data.

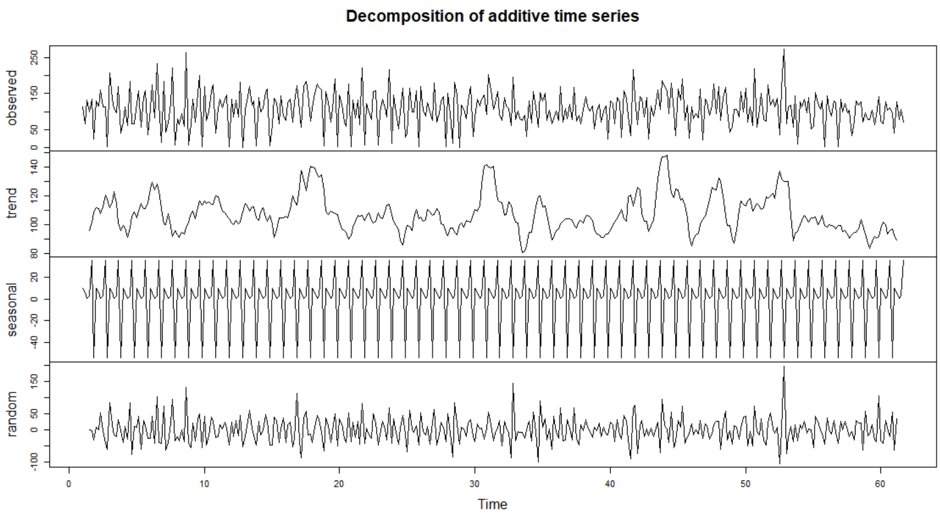


Figure 15: Time series decomposition for store 657

Finally, an autocorrelation plot was designed. Autocorrelation is used to obtain the degree of similarity of a time series with itself, measuring the linear relationship between lagged values of a time series. It is useful to discover seasonal patterns in a time series, as well as to assess if it is suitable for the application of univariate forecasting methods. A relationship between the lagged values and the output needs to exist for them to have predictive power. Figure 51 from Appendix G confirms the existence of weekly seasonality as autocorrelation shows a peak in lag 6, and multiples of 6. An observation of lag 6 represents the homologous observation in the previous week, as a week is considered to have 6 days.

4.6 Analysis of the store representing wholesaler stores

Figure 16 shows the total number of pallets shipped per day to the wholesaler store. This store receives an average of 30,1 pallets a day. A fluctuation in the number of pallets daily shipped to this store through the year is visible, similar to the one in the aggregated analysis, Figure 4. The number of boxes shipped to the wholesaler store is presented in Figure 17 and shows a seasonal pattern similar to the number of pallets shipped daily to the same store.

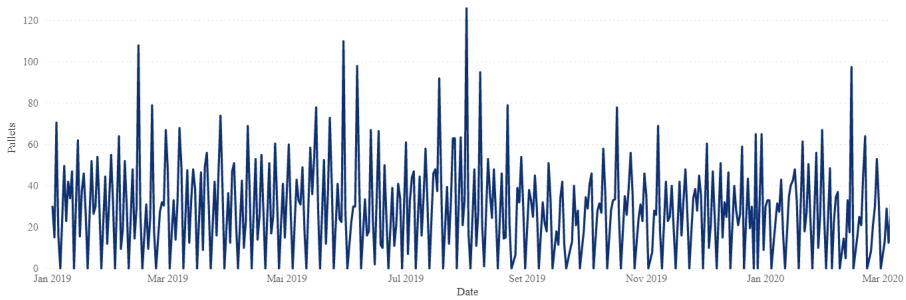


Figure 16: Number of pallets shipped daily to store 2022 between January 2019 and March 2020

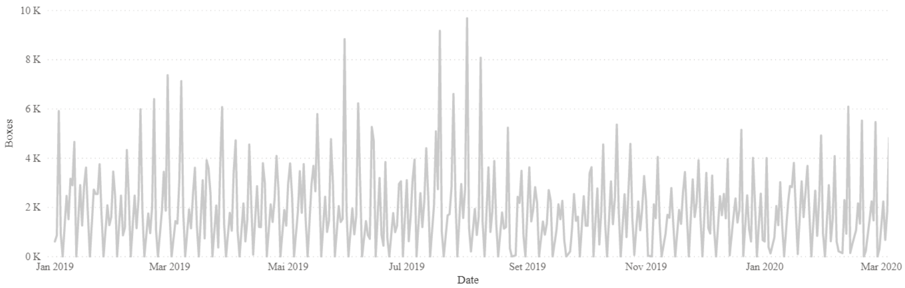


Figure 17: Number of boxes shipped daily to store 2022 between January 2019 and March 2020

The descriptive statistics of the dependent variable number of pallets and the numeric explanatory variables considered for the wholesaler store are represented in Table 4.

It is relevant to mention, that the coefficient of variation of the number of pallets and boxes shipped daily to this store is higher than the one computed for the hypermarket store. At the same time, the number of pallets daily shipped its much smaller.

Additionally, weekly seasonality is represented in Figure 18. Usually, there is no expedition of pallets to this store on Saturdays. Moreover, there is a marked difference between the average number of pallets shipped per day per weekday. Thursday is the strongest day, while Tuesday is the day with the lowest number of pallets shipped. Figure 19 represents the coefficient of variation per weekday, and shows a slight variation of the coefficient.

Table 4: Descriptive statistics of the dependent and explanatory variables for store 2022

	Number of pallets	Number of boxes	Total number of pallets
Minimum	0	0	1920
Median	29,9	1701	4264
Mean	30,1	2006	4207
Maximum	126	9685	6016
Standard deviation	22,20	1683	700,46
Coefficient of variation	0,737	0,839	0,166

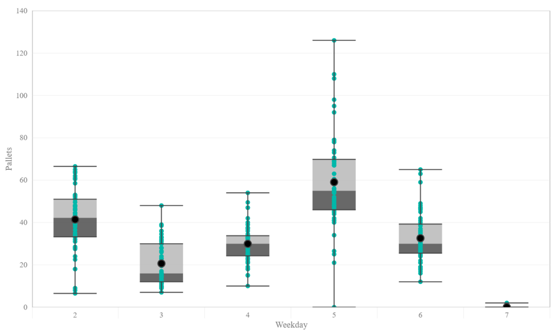


Figure 18: Number of pallets shipped daily to store 2022 per weekday in 2019

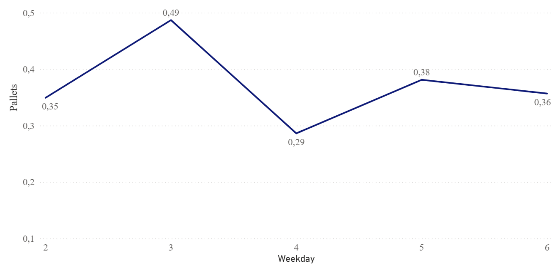


Figure 19: CV of the number of pallets shipped per day to store 2022 by weekday in 2019

Moreover, Figure 20 presents the decomposition of the time series. It shows a slight and irregular trend component, indicating some fluctuations of the number of pallets shipped through the year, as well as a constant and marked seasonal pattern, different from the one observed in the hypermarket store.

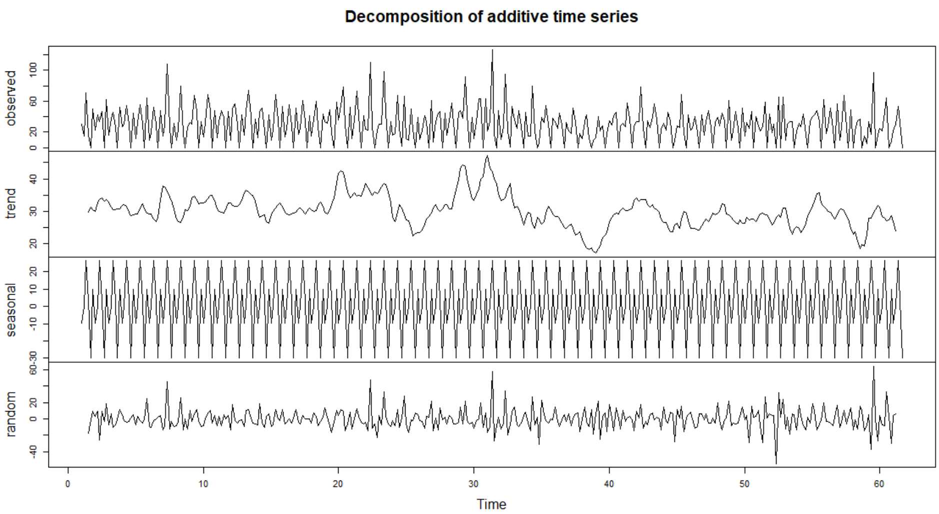


Figure 20: Time series decomposition for store 2022

Lastly, the autocorrelation plot is represented in Figure 52 from Appendix G and also reveals the existence of weekly seasonality as autocorrelation shows a peak in lag 6, and multiples of 6.

4.7 Analysis of the store representing supermarket stores

Figure 21 shows the total number of pallets shipped per day to the supermarket store. This store receives an average of 24,5 pallets a day. Besides, the variation of the daily number of pallets shipped to the supermarket store throughout the year also displays a peek during the month of August, identical to the one in the aggregated analysis, Figure 4. The number of boxes shipped to the supermarket store is presented in Figure 22 and shows a seasonal pattern similar to the number of pallets shipped daily to the same store. It also shows a slight increase in the number of boxes during August.

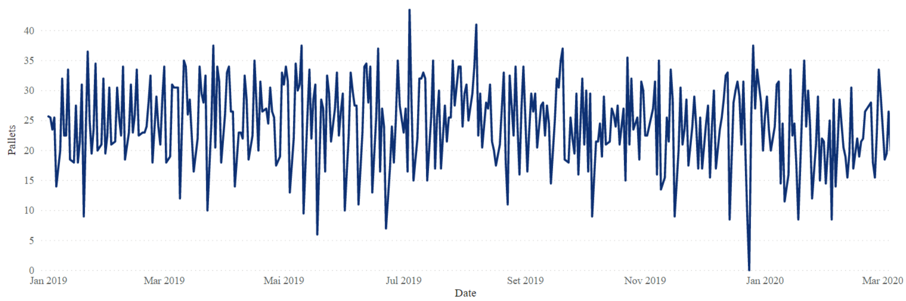


Figure 21: Number of pallets shipped daily to store 347 between January 2019 and March 2020

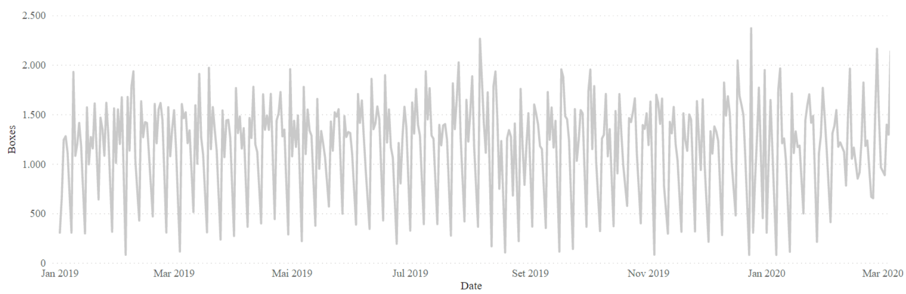


Figure 22: Number of boxes shipped daily to store 347 between January 2019 and March 2020

The descriptive statistics of the dependent variable number of pallets and the numeric explanatory variables considered for the supermarket store are represented in Table 5.

Additionally, weekly seasonality is visible in Figure 23. Saturday is the day with the lowest number of pallets shipped, with Friday being the strongest day, showing a pattern similar to the warehouse aggregated operation. Figure 24 presents the coefficient of variation per weekday. It is rather constant except for a small increase in the coefficient on Saturdays. The coefficient of variation is the lowest for the supermarket store when compared to the other stores analyzed.

Furthermore, Figure 25 presents the decomposition of the time series. It shows a slight and irregular trend component, less pronounced than the trend component computed for the other stores. Besides, it shows a constant and marked seasonal pattern, more similar to the one observed for the hypermarket store.

Table 5: Descriptive statistics of the dependent and explanatory variables considered for store 347

	Number of pallets	Number of boxes	Total number of pallets
Minimum	0	0	1920
Median	25,5	1298	4264
Mean	24,85	1221	4207
Maximum	43,5	2374	6016
Standard deviation	6,86	483,05	700,46
Coefficient of variation	0,276	0,396	0,166

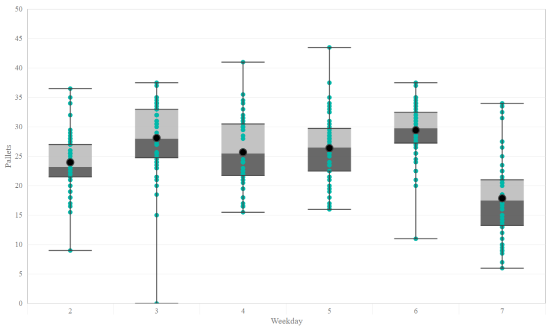


Figure 23: Number of pallets shipped daily to store 347 per weekday in 2019

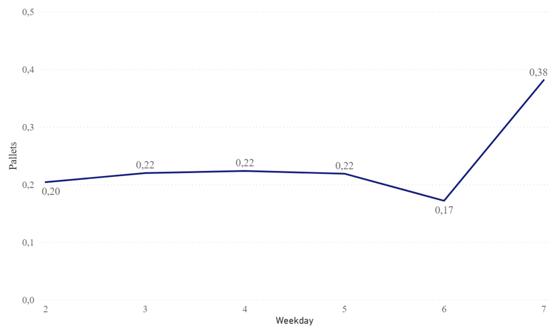


Figure 24: CV of the number of pallets shipped per day to store 347 by weekday in 2019

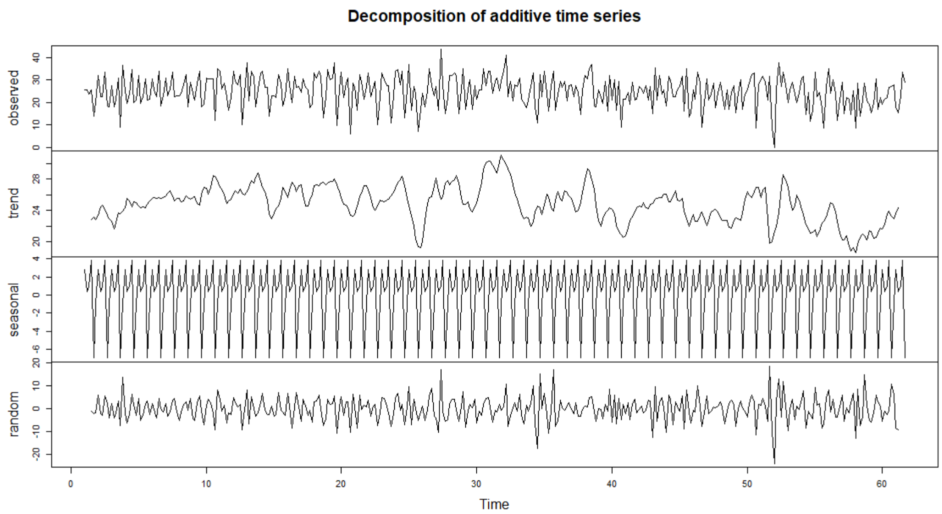


Figure 25: Time series decomposition for store 347

Lastly, the autocorrelation plot is represented in Figure 53 from Appendix G, and also reveals the existence of weekly seasonality as autocorrelation shows a peak in lag 6, and multiples of 6.

Chapter 5

Results

5.1 Performance indicators

It what concerns performance indicators, Normalized Root Mean Squared Error (NRMSE) and weighted Mean Absolute Percentage Error (wMAPE) were chosen to evaluate the performance of the model.

NRMSE is a derivation of RMSE, calculated with its normalization. The normalization method performed was min-max. NRMSE facilitates the comparison between different objects, for example objects with different sizes. In this case, it is particularly important as stores can have significantly different sizes and variations in the number of pallets received. NRMSE makes it possible to compare the performance of the methodologies between stores. RMSE and NRMSE are calculated according to the formulas in Equation 5 and Equation 6 from Appendix D.

On the other side, it was considered relevant to have another performance indicator. WMAPE was selected since it is easy to interpret and overcomes MAPE's inability to be computed when observed values are equal to zero. Its formula of calculation is shown in Equation 4 from Appendix D.

NRMSE tends to penalize large errors, which is an important characteristic as large errors in predicting the number of pallets to be shipped to a store are more disturbing at operational level.

RMSE will also be used when building the machine learning algorithms.

5.2 Implementation of the statistical forecasting methods

As mentioned before, three statistical methods were implemented, Holt-Winters, seasonal Arima and Tbats models. It is important to highlight that these methods are capable of handling time series with seasonality.

The models were developed in RStudio, utilizing R Language and the package forecast. Seasonal Arima was implemented through auto.arima function, Holt-Winters using hw function, and at last, Tbats with tbats function.

These functions return the best model according to the Akaike Information Criterion (AIC). The AIC is used to compare models, rewarding models that have a high goodness-of-fit while penalizing complexity. A lower AIC should avoid over-fitting the model while enabling it to fit the data properly.

The data set was split into a training sample, observations during the year of 2019, and test sample observations during January and February 2020. So, for each store, the test period has a 51-day length.

Three stores were closed during January and February, being removed from the tests. For each one of the remaining stores, the three statistical models were fitted.

In time series forecasting, the order of the observations matters, for instance, the training data needs to be older than the test set. Having said this, the expanding window approach was utilized. This means that the test window is fixed, while more training data is being added to the model. In this case, a 2 step forecast is performed. With the expanding window approach, all data points until 2 days prior to the forecast are used to train the model and produce its output. Every time the model makes a prediction, a new data point is added to the training set and applied to the next prediction.

5.3 Random Forest model construction

Random forest model was built using randomForest package and randomForest function in RStudio. A sliding window approach was chosen, contrarily to what was applied in the statistical method's testing. So, predicted values are computed utilizing the last X observations, where X is the number of lagged observations included in the model.

Machine learning models are suitable for multivariate forecasting models, in opposition to the statistical methods. Thus, besides providing lagged observations as inputs to the model, explanatory variables can also be added.

The Random Forest model was built in three main phases. The first one consisted of deciding the number of lagged weeks of observations to include in the model. This optimization was performed varying the number of weeks from 1 to 34, and running 20 iterations every time a week of data was added to the model. A maximum of 34 weeks was considered due to the fact that every time a lagged week is added to the model as a regressor, a week of training data is lost. Therefore, 34 weeks was considered the limit to guarantee at least four months of training data. The number of weeks that minimized the RMSE was chosen as the first input to the model.

The second phase was feature selection. The explanatory variables mentioned in Chapter 4 were included in the model. The strategy was to include these features one at a time in the model. Then, evaluate whether or not the inclusion of these features improved the accuracy of the model, in this case, given by RMSE. The explanatory variable with the highest predictive power, that when used as an input together with the lagged observations resulted in the model with the lowest RMSE, was added to the model. Variables that did not improve the RMSE of the model were discarded as candidates. In the following iteration, the remaining candidate variables were included in the new model, again one at a time. If any of the candidate variables was able to improve the performance of the model, it was added to it. This process was repeated until there were no more candidate variables available. In this phase, 50 iterations per model were run.

The final phase, after selecting the number of lagged observations and the explanatory variables to include in the model, was parameter tuning. The number of trees to grow and the number of features to consider at each decision node were the two parameters optimized. The number of trees ranged from 100 to 1000, with step 100, and the number of features at each node varied from 2 to 20. It is important to mention that these parameters remained unchanged during the previous phases. The number of trees was set at 1000 and the number of features at each node was considered equal to the square root of the number of inputs. The number of iterations ran per model was also 50.

Following the construction steps mentioned above, a Random Forest model was constructed for the hypermarket store, the wholesaler store and the supermarket store. Then, a Random Forest model containing the optimal number of lagged weeks, explanatory features and tuned parameters found for each store type was run for every store in their respective clusters.

This was motivated by the fact that it was not practical to follow these construction steps for 190 stores. Therefore, clustering was used to group stores with similar behavior. A representative store from each cluster was then chosen as it is believed that shipments to stores with similar behavior and characteristics could be influenced by the same features.

Before the construction of the model an augmented Dickey-Fuller test was performed to each store's time series. It confirmed that the representative stores' time series were stationary so differencing was not applied as a preprocessing step. The results of this test can be found in Figure 58, Figure 59 and Figure 60 from Appendix I. Furthermore, the number of pallets daily shipped to the stores was normalized before being fed to the model.

5.3.1 Construction of the Random Forest model for the hypermarket store

In this section, random forest model construction for the hypermarket store is explained.

Following the framework described above, the first procedure was to determine the number of lagged weeks of observations to include in the model. Figure 26 shows the evolution of the RMSE as the number of weeks included in the model increased. It is visible that RMSE has a minimum when two weeks of lagged values are included in the model, so two is the number of lagged weeks to be fed to the algorithm.

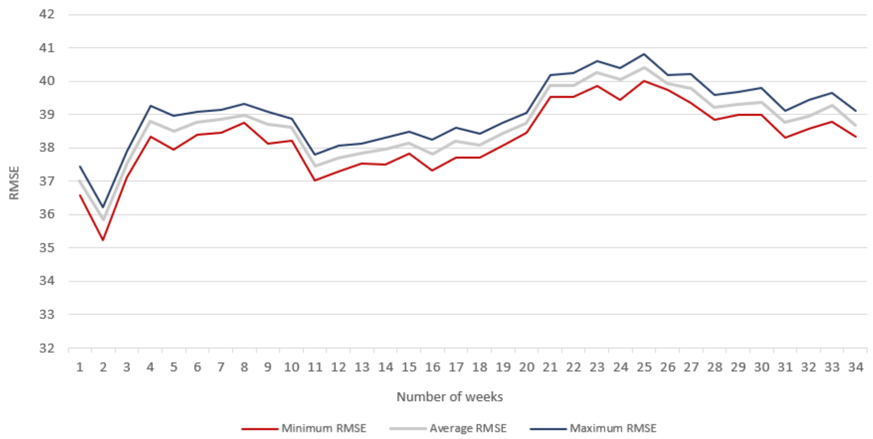


Figure 26: RMSE per number of lagged weeks added to the model for store 657

The next step was feature selection. The first iteration, where all explanatory variables were included in the model one at a time, is represented in Figure 27.

The first variable to enter the model was the total number of pallets shipped by the warehouse, since it was the one with that reduced the RMSE the most. The variables day of month, Easter, holiday, and month were discarded from candidate variables due to the fact their inclusion worsened the performance of the model.

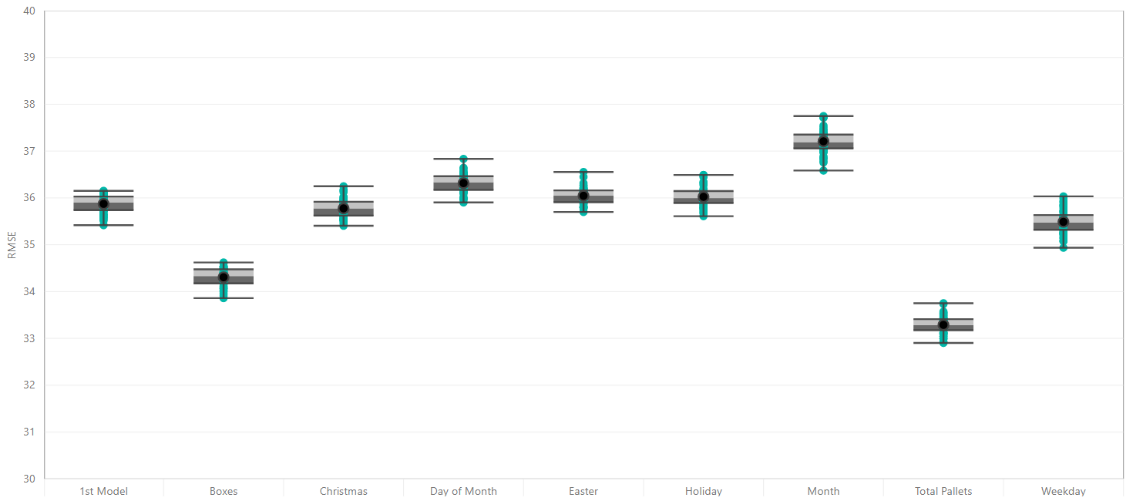
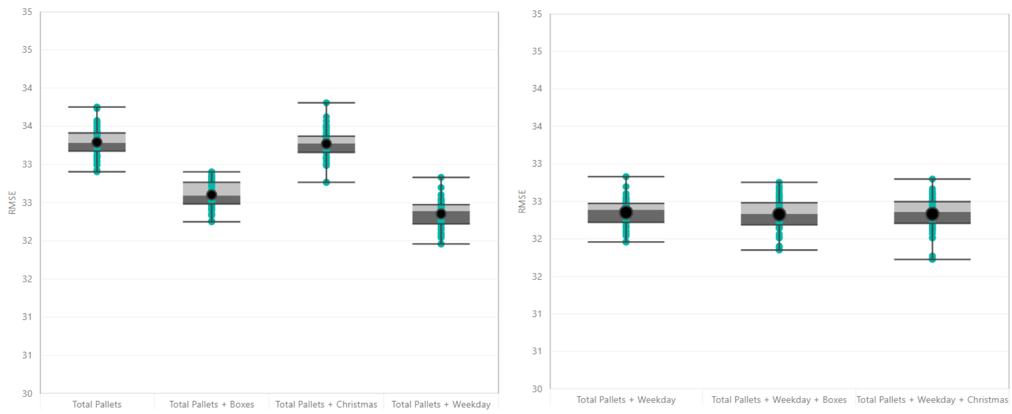


Figure 27: Random Forest feature selection process for store 657, first iteration.

Figure 28 displays the second and third iterations of the feature selection process. In the second iteration, the remaining candidate variables were added to the model, and as a conclusion Weekday was included in the model. No variable was discarded from candidate as all variables improved the performance of the model. Following the same logical procedure in the third and fourth iterations, a final model was reached.



(a) Second Iteration (b) Third Iteration

Figure 28: Random Forest feature selection process for store 657, second and third iterations.

Figure 29 represents the evolution of the RMSE from the initial model, without any explanatory variables, to the final model, after the fourth iteration. The final model for the hypermarket store includes two weeks of lagged observations and the features total number of pallets, weekday, number of boxes and Christmas.

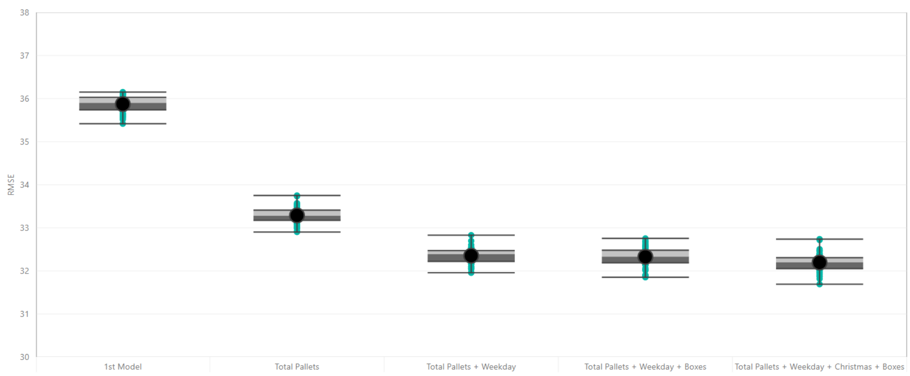


Figure 29: Random Forest final model for store 657

Finally, in the parameter tuning phase, the number of trees to grow was set at 1000 trees and the number of inputs to consider at each split node was defined as 4. Once again these parameters were chosen because they were the ones that minimized the RMSE.

5.3.2 Construction of the Random Forest model for the wholesaler store

Firstly, the number of lagged weeks to include in the model was set at 3 weeks. The evolution of the RMSE as the number of lagged weeks included in the model increases is represented in Figure 30.

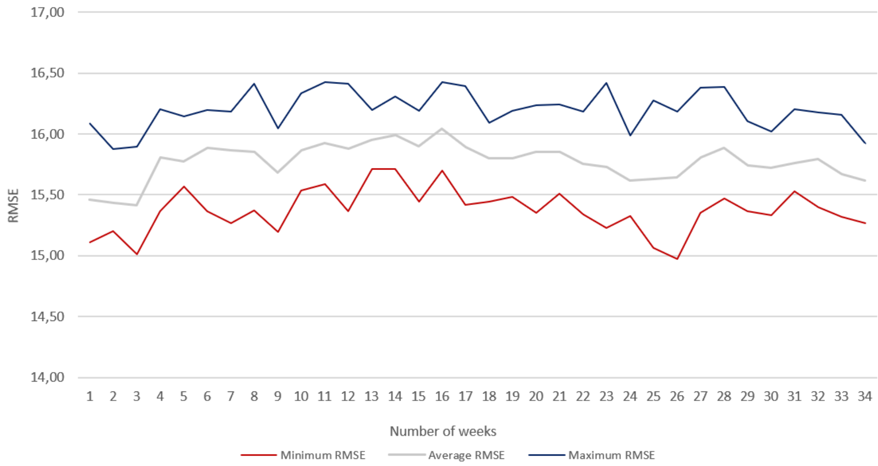


Figure 30: RMSE per number of lagged weeks added to the model for store 2022

Secondly, regarding the feature selection step, the first iteration is shown in Figure 31. The only explanatory variable able to improve the performance of the model was the total number of pallets shipped by the warehouse. Therefore, this variable was the only external variable included in the model.

The final model for the wholesaler store includes 3 weeks of lagged observations and the explanatory variable total number of pallets.

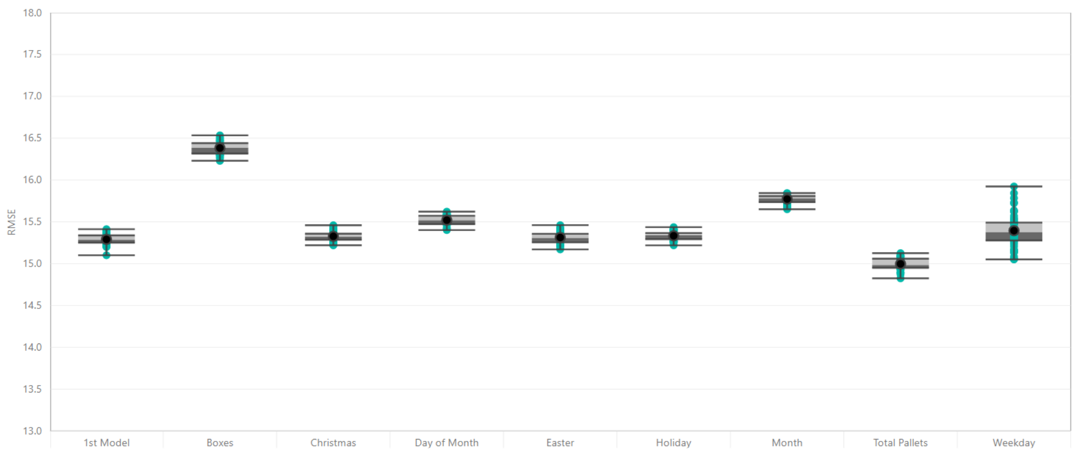


Figure 31: Random Forest feature selection process for store 2022, first iteration

Finally, the optimal number of trees to grow was 900 and the number of inputs to consider at each split node 9.

5.3.3 Construction of the Random Forest model for the supermarket store

At last, a similar process was implemented for the supermarket store as well.

Figure 32 supported the selection of the number of lagged weeks to consider in the model. In this case, a total of 9 weeks, 54 data points, were included in the model.

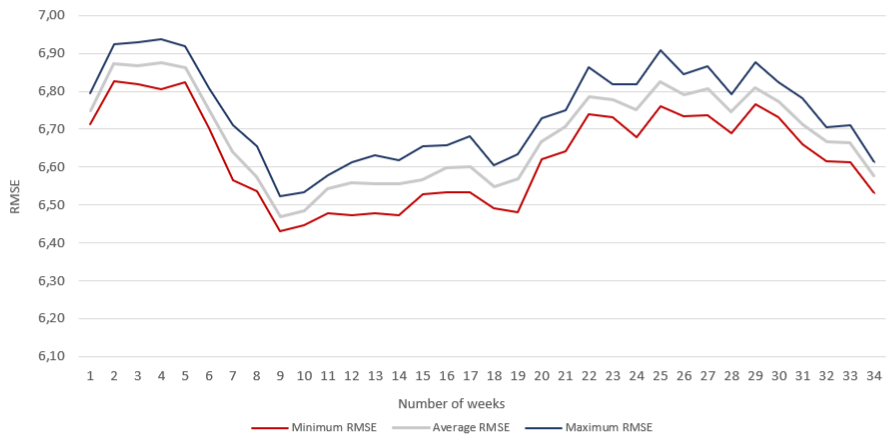


Figure 32: RMSE per number of lagged weeks added to the model for store 347

Regarding feature selection, the results after the first iteration are available in Figure 33. The variable total number of pallets shipped produced the best result, thus entering the model. Besides this variable, only the variables number of boxes and weekday were able to enhance the performance of the model, being the remaining variables discarded as candidates.

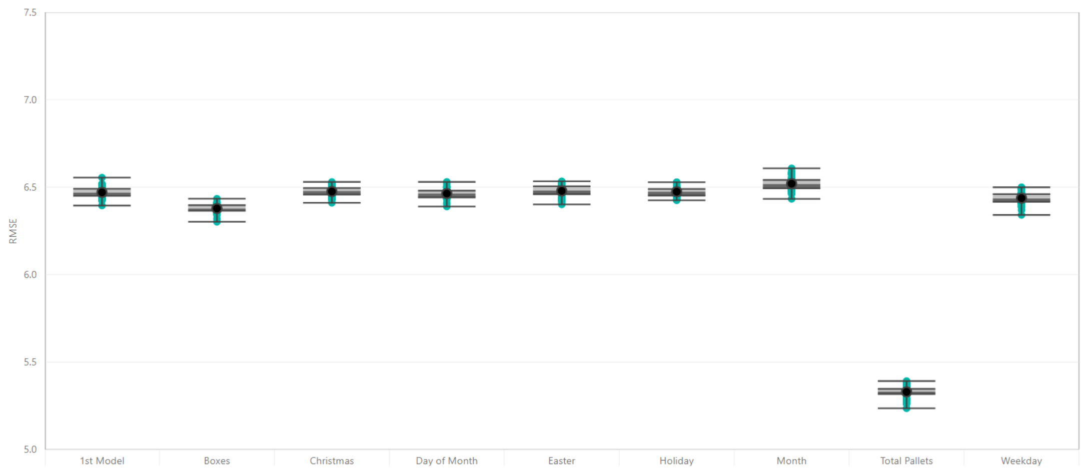


Figure 33: Random Forest feature selection process for store 347, first iteration

In the second iteration, the variables number of boxes and weekday were not able to improve the performance of the current model neither when they were input one at a time nor when both of them were included. These results are available in Figure 34.

Thus, the final model for the supermarket store includes 9 weeks of lagged observations and the explanatory variable total number of pallets.

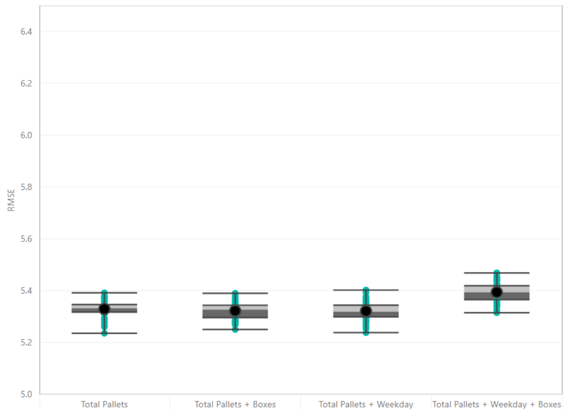


Figure 34: Random Forest feature selection process for store 347, second and third iterations

Finally, after the parameter tuning phase, the optimal number of trees to grow was 700 and the number of inputs to consider at each split node 20.

5.4 LSTM Neural Network Implementation

The Long Short Term Memory Neural Network implementation was performed resorting to package Keras in RStudio.

Similar to the framework followed when constructing the Random Forest model, the first step was to found the ideal number of lagged weeks of observations to include in the model. In the present case, it was considered relevant to vary the number of neurons and epochs while computing the best number of lagged weeks. This decision was motivated by the fact that the architecture of the Neural Networks and the number of training iterations can have a relevant impact on the results obtained. Therefore, the number of neurons considered was 10, 50 and 100 neurons and the number of epochs 100 and 500 epochs.

The best results obtained by the LSTM NN model for the hypermarket store along with a comparison of the Random Forest results obtained also at the first stage of the model construction are represented in Figure 35. Figure 36 and Figure 37 show a similar representation of the results obtained for the wholesaler store and supermarket store, respectively.

Table 6 presents a comparison of the performance of both algorithms for the first stage of model construction, the selection of the number of lagged weeks to include in the model. The average and the standard deviation of the NRMSE obtained per number of weeks included in the model were computed.

For the hypermarket store, it is shown that the LSTM NN model has a worse performance than the Random Forest model, for all lagged weeks of data included in the model. Besides, Random Forest shows a much more stable performance.

For the wholesaler store, Random Forest had a superior performance, on average, than the LSTM NN model. Moreover, it provided less variable results as well.

For the supermarket store, the average performance of both models was similar, with the LSTM NN model having a slightly less stable performance.

Based on these results and on the fact that Neural Networks are harder to train and require a much longer computational time to be trained than Random Forest, a decision was made not to pursue further tests with the Neural Networks model. Random Forest model was considered to be a more indicated option for this particular problem.

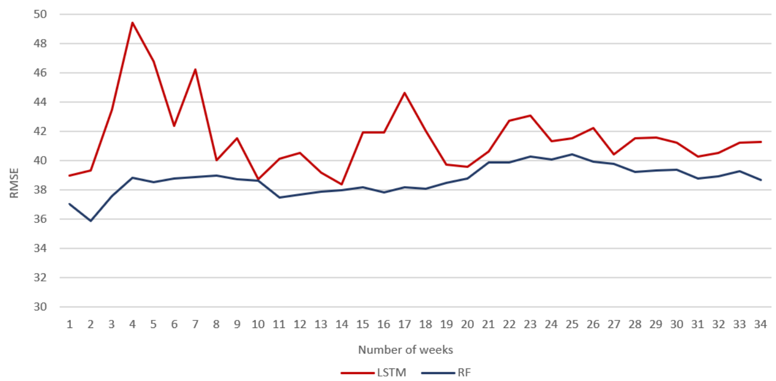


Figure 35: Performance of LSTM NN model and RF model for store 657

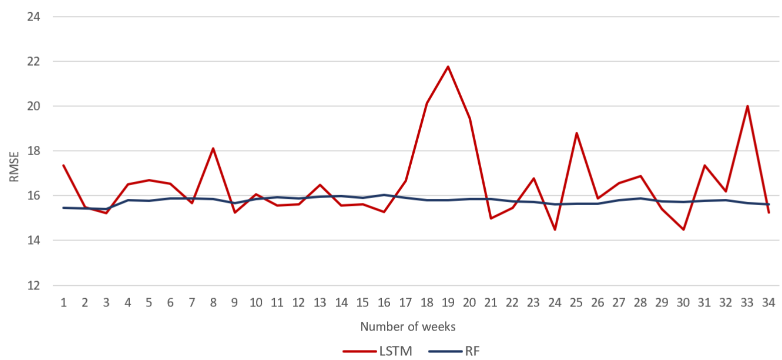


Figure 36: Performance of LSTM NN model and RF model for store 2022

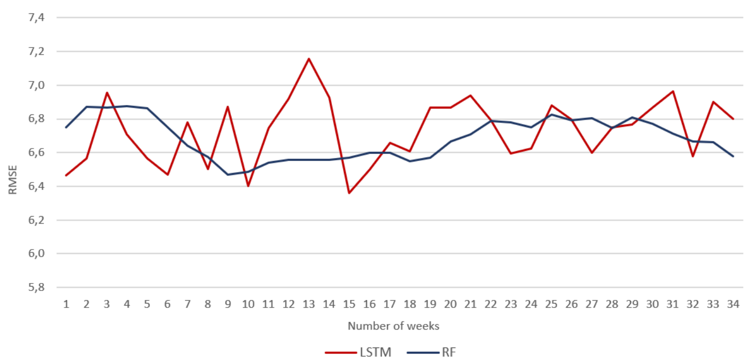


Figure 37: Performance of LSTM NN model and RF model for store 347

Table 6: NRMSE obtained with Random Forest model and LSTM Neural Network model

		Random Forest	**LSTM Neural Network**
Hypermarket	Average	25,5%	27,5%
	Std dev	0,64%	1,53%
Wholesaler	Average	16,2%	17,0%
	Std dev	0,15%	1,72%
Supermarket	Average	25,2%	25,4%
	Std dev	0,45%	0,70%

5.5 Forecasting model's results

5.5.1 Results obtained for the representative stores

The forecasting results for the representative stores of each cluster, a hypermarket store, a wholesaler store and a supermarket store, are shown in Table 7. It contains the results for the statistical models and Random Forest model, evaluated by the performance indicators NRMSE and wMAPE. These results were obtained for the test period, the months of January and February of 2020.

For the hypermarket store, the statistical model with the best performance was TBATS, although the statistical models have a similar performance between then. Random Forest obtained the best result, being able to improve the performance of the statistical models.

Moreover, for the wholesaler store, seasonal ARIMA was the statistical model with the best performance. In this case, the differences between the performance of the statistical models are more pronounced. Random Forest obtained an identical performance to the seasonal ARIMA.

Lastly, for the supermarket store, the best statistical model was Holt-Winters. Nevertheless, Random Forest model was able to significantly improve the performance of the statistical methods. The difference between the results obtained by Random Forest model and the statistical ones is more pronounced for the supermarket store.

Table 7: Forecasting results for the representative stores of each cluster

		Holt-Winters	**SARIMA**	**TBATS**	**Random Forest**
Hypermarket	NRMSE	25,1%	24,9%	24,3%	21,1%
	wMAPE	31,2%	29,8%	31,7%	26,9%
Wholesaler	NRMSE	16,9%	15,4%	17,8%	15,4%
	wMAPE	42,4%	38,0%	42,3%	38,0%
Supermarket	NRMSE	24,8%	26,3%	25,1%	18,1%
	wMAPE	23,7%	23,8%	24,1%	17,8%

A representation of the forecasted values against the observed ones for the hypermarket store is displayed on Figure 38 and Figure 39. Figure 38 presents the results obtained for the TBATS model since it was the statistical model with the best performance for this store. Figure 39 presents the prediction obtained with Random Forest model.

Both models capture the minimum values present in the first half of the test period. Yet, they still over predict the number of pallets shipped in those days. One of the reasons behind this is that these minimum values are more pronounced during January than in the previous months.

Furthermore, it is visible that the time series behavior changes in the second half of the test period. Both models struggle in modeling these changes, although Random Forest seems to be able to adapt better to them.

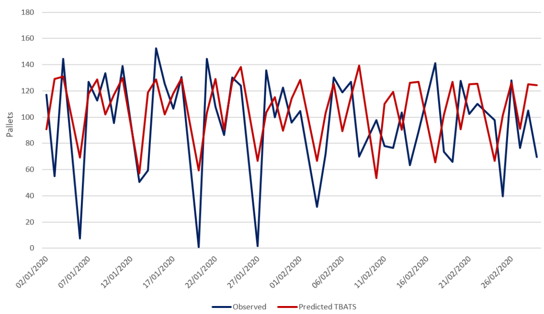


Figure 38: TBATS predictions for store 657

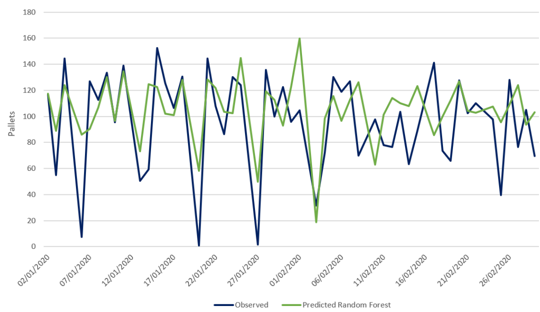


Figure 39: RF predictions for store 657

Likewise, a representation of the predicted values against the observed ones for the wholesaler store is displayed in Figure 40 and Figure 41. The best statistical model was seasonal ARIMA.

The wholesaler store usually does not have deliveries on Saturdays, and both models are able to capture this pattern. In fact, both models show a similar behavior between them. They fail to predict the break on the 4th of February, they overestimate the number of pallets on the 10th of February and miss the peak on the 14th of February. A change in the seasonal pattern during February is visible.

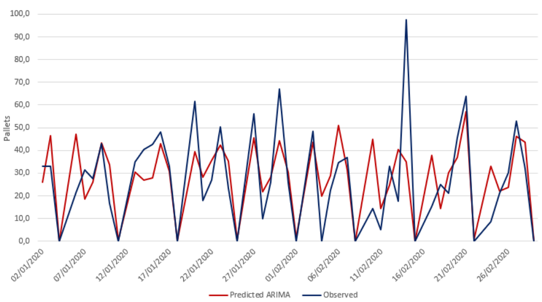


Figure 40: SARIMA predictions for store 2022

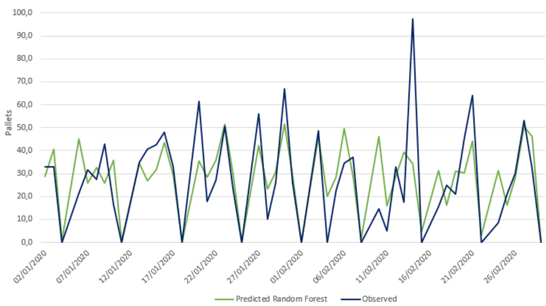


Figure 41: RF predictions for store 2022

At last, a similar representation is displayed on Figure 42 and Figure 43 for the supermarket store. For this store, the statistical model with the best performance was Holt-Winters, thus the results of this model are presented in Figure 42.

For the supermarket store, the minimum values are not as sharp as the ones for the hypermarket store. Nevertheless, both models tend to overestimate these days. Additionally, both models fail to capture the break on the 4th of February, and underestimate days 8 and 22 of February. In this store, it is also visible a change in the seasonal pattern of the number of pallets shipped per day in February. Random Forest seems to provide a better prediction than Holt-Winters for this month.

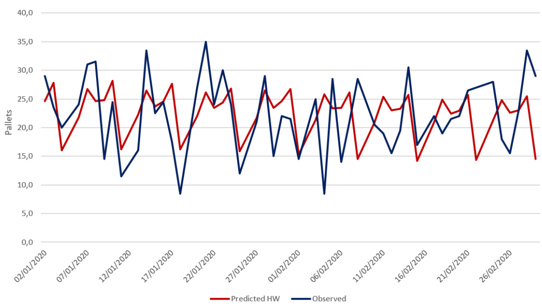


Figure 42: HW predictions for store 347

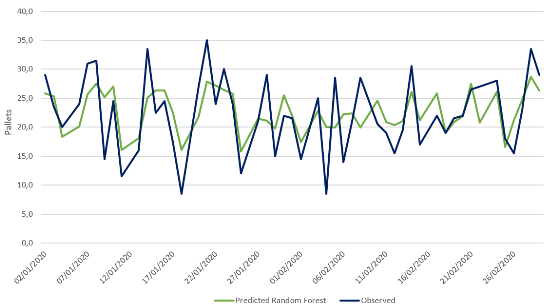


Figure 43: RF predictions for store 347

5.5.2 Overall forecasting results

The statistical forecasting methods under analysis were applied to each one of the 190 stores, as described above.

Regarding the machine learning methods, the number of lagged weeks and external variables to include in the model as well as the model's parameters were optimized for each of the three representative stores. These stores represent the cluster they belong to. Therefore, a Random Forest model was run individually for every store taking into account the cluster they are a part of and the features selected for its representative store.

The NRMSE and the wMAPE were computed for each store and method applied. The results are presented by calculating the average and the standard deviation of these performance indicators by store.

Furthermore, the Ensemble method was computed, choosing the statistical method with the lowest NRMSE per store. The SARIMA model had the best performance for most of the stores, 48%, followed by the Holt-Winters model, which had the best performance for 39% of the stores.

Table 8 represents the overall results obtained by the implementation of each of the statistical forecasting methods, the Ensemble method and the Random Forest model. Figure 44 represents a box plot of the NRMSE obtained by the implementation of each model for every store.

Table 8: Aggregated forecasting results

		Holt-Winters	**SARIMA**	**TBATS**	**Ensemble**	**Random Forest**
NRMSE	Average	21,0%	21,1%	21,7%	20,5%	19,2%
	Std dev	2,90%	2,97%	2,98%	2,73%	2,99%
wMAPE	Average	27,4%	27,6%	28,5%	26,7%	25,4%
	Std dev	8,11%	8,87%	8,96%	8,05%	9,10%

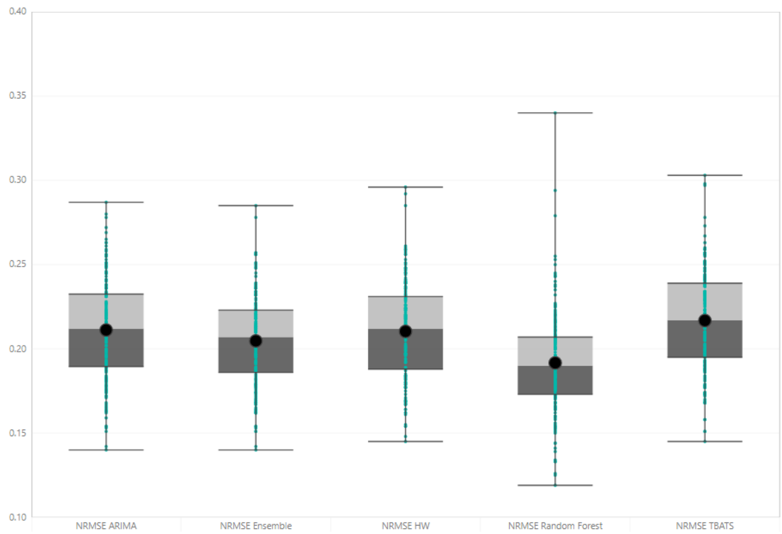


Figure 44: NRMSE obtained for each store by forecasting method applied

It is shown that the statistical methods have a similar performance between them, with TBATS model having, on average, a slightly worse performance than the others.

Random Forest, on average, was able to improve the performance of the statistical forecasting methods. However, it led to a slightly higher standard deviation of the performance indicators, meaning it might not provide consistently better results for every store. From Figure 44 it is visible that there are at least a couple of stores for which Random Forest model had a much inferior performance than the statistical ones.

Moreover, the overall percentage of improvement in the performance of the models obtained by Random Forest was inferior to the one obtained for the hypermarket and supermarket stores.

A comparison between the performance of the forecasting models by cluster is represented in Figure 45. Cluster 1 represents the wholesaler stores, Cluster 2 the hypermarket stores and Cluster 3 the supermarket stores.

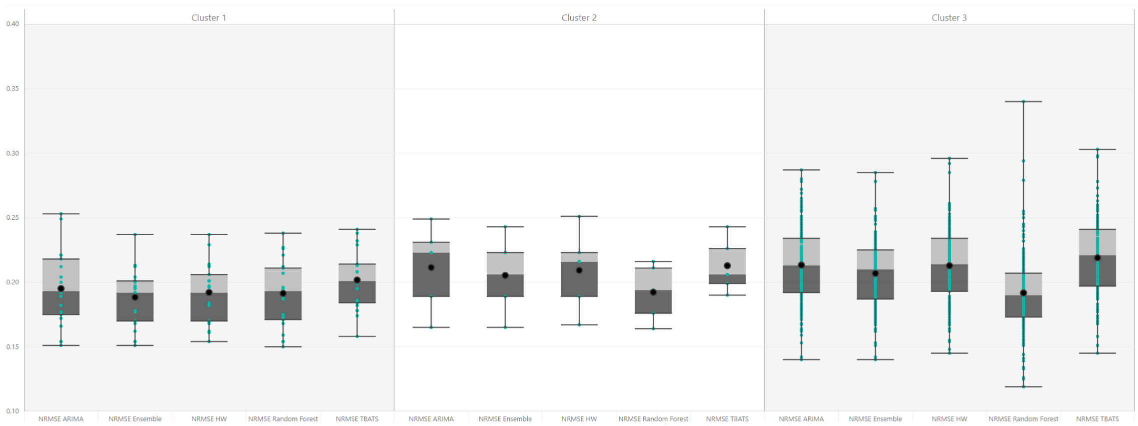


Figure 45: NRMSE obtained for each store by the method applied grouped by cluster

For the wholesaler stores, there is no clear difference between the performance of the models. The Ensemble method and Random Forest are the models with the best performance.

On the other hand, for the hypermarket stores, the superior performance of Random Forest is visible in Figure 45. The statistical methods have a similar performance between them.

Finally, Random Forest is also the model with the best performance, on average, for the supermarket stores. In this cluster, there is a wider range in the NRMSE obtained for each store, coherent with the fact this is the cluster with the highest number of stores. There are at least two stores for which Random Forest model had a much inferior performance than the statistical models.

5.5.3 Comparison between the results obtained for January and February

At the beginning of February, as shown in the time series displayed above, there was a shift in the expedition pattern. This change is justified by the implementation of a new working schedule in the warehouse, that also changed the time window for expeditions.

Figure 46 shows the differences in the performance of the models for the months of January and February.

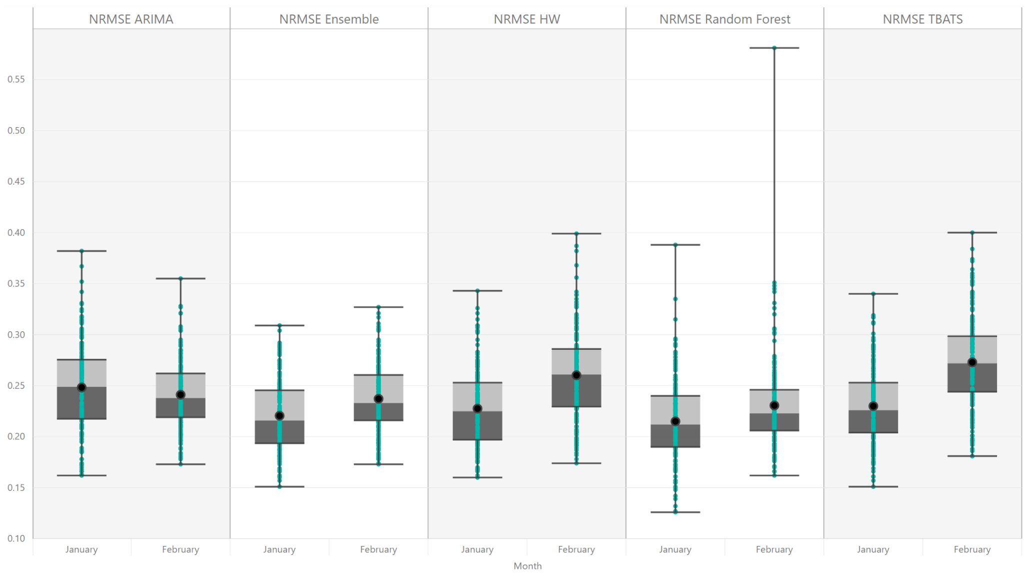


Figure 46: NRMSE of each store for the months of January and February

Holt-Winters and TBATS models show a clear fluctuation in the results obtained, with the results for January being more accurate, as expected. Random Forest also performed better for January, but the difference between months is not as marked as the one present in Holt-Winters and TBATS models. Despite this fact, there is a store for which the performance of Random Forest during February was highly unsatisfactory.

On the other side, the SARIMA model had a slightly better performance during February, which is unexpected. This can be explained if SARIMA is using a reduced number of past observations in the auto-regressive part as well as in the moving average part.

Despite these differences, Random Forest was the model with the best performance for both months, on average.

5.5.4 Comparison between the results obtained by the forecasting methods and the volumetry based method

Table 9 presents a comparison between all the methodologies applied. The average NRMSE and wMAPE per store as well as the standard deviation were computed.

It is possible to conclude that the methodology based on volumetry proposed had, on average, a better performance than the forecasting models proposed.

Table 9: Aggregated results for the methods applied

		HW	SARIMA	TBATS	Ensemble	RF	Volumetry
NRMSE	Average	21,0%	21,1%	21,7%	20,5%	19,2%	15,9%
	Std dev	2,90%	2,97%	2,98%	2,73%	2,99%	4,45%
wMAPE	Average	27,4%	27,6%	28,5%	26,7%	25,4%	20,4%
	Std dev	8,11%	8,87%	8,96%	8,05%	9,10%	6,84%

A box plot comparing the NRMSE obtained for each store by each method is represented in Figure 47. Although it is visible that the volumetry based method has a better performance, on average, there are some stores for which the model has a considerably worse performance. This might indicate that the model might not be well parameterized for these stores.

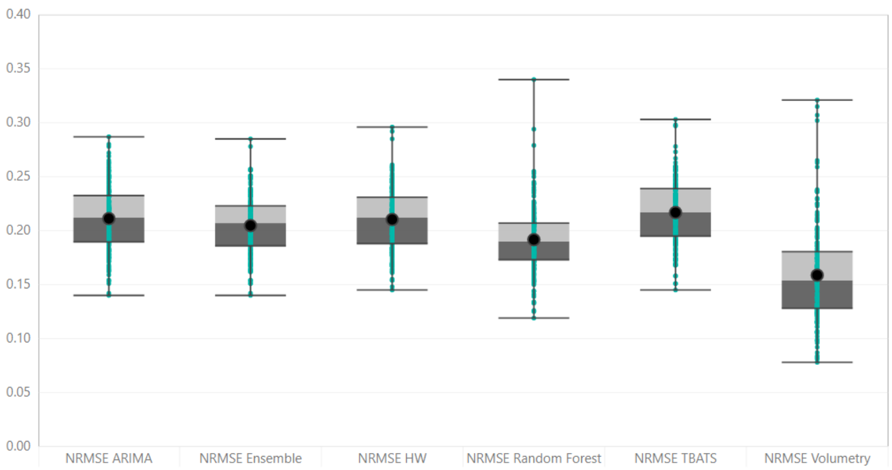


Figure 47: NRMSE obtained for each store by the method applied

Chapter 6

Conclusion

In this dissertation, two approaches to improve the forecast of the number of pallets shipped per day to each store were taken. The first was a practical approach based on volumetry and the request made by each store. The second was a more theoretical approach applying time series forecasting to this problem. The methods tested included both statistical models, Holt-Winters, Seasonal ARIMA and TBATS and machine learning models, Random Forest and LSTM Neural Networks. The present chapter includes a discussion of the results obtained by these models along with an overview of the pros and cons of each model when applied to this particular problem. It finishes with a summary regarding the limitations of this dissertation and further work suggestions.

6.1 Results discussion

From the forecasting methods for time series applied, Random Forest was the one with the best performance. The statistical methods had quite similar performances between them.

The statistical models have the advantage of being simple and relatively easy to implement while producing viable results. In this case, the implementation was performed in RStudio, and anyone with some experience in programming should be able to implement them. These methods can be implemented through other platforms as well.

On the other hand, machine learning methods require some in-house expertise on the subject in order to be implemented and sustained correctly. Two machine learning algorithms were tested, Random Forest and LSTM Neural Networks. Contrarily to what was expected, Neural Networks did not provide a better nor more consistent performance than Random Forest. Besides, they have the disadvantage of being computationally inefficient, taking several hours to train them. In addition, Neural Networks have more parameters to tune than Random Forest, with results being rather sensitive to them. To train a Neural Network properly a deeper knowledge on the topic is necessary. Random forest, on the other side, is computationally efficient, has few parameters to tune and its simple to use, for a machine learning model.

When compared to the statistical models implemented in this dissertation, Random Forest has the advantage of being more accurate and allowing the use of explanatory variables in the model.

On the other side, it is necessary to train the model from time to time while that is not needed for the statistical models. Despite being relatively simple to implement, statistical models are even more simple and require less in-house expertise to be maintained over time. However, statistical methods provide better results when conditions are stable, and struggle to adapt when external conditions change.

Both statistical and machine learning models require a constant update of the information in order to be able to perform. In the case of multivariate models, not only the variable being predicted but also the explanatory variables need to be updated. So the choice of multivariate variables has to take into consideration the trade-off between the increase in the accuracy of the model and the increased difficulty to sustain it.

In the particular case of this dissertation, the number of pallets shipped per day per store sometimes showed strong variations. Forecasting models were not able to always capture this irregular behavior.

The first method proposed in this dissertation is also a viable option to pursuit and, in this case, it produced the best result. It has the advantage of being the most simple method and not requiring any additional knowledge. Additionally, since the main input to this tool is the order made by each store, fluctuations in the number of pallets shipped are probably detected more easily by this tool. Like the forecasting methods under analysis, this tool has the disadvantage of being highly dependent on the input given, thus being subject to human error.

To conclude, the decision regarding the best method to implement depends on a variety of factors, such as the prediction accuracy needed, the in-house expertise available, the information available, the urgency of the implementation and so on. On the one side, statistical methods are widely used, are fast and easy to implement and provide a viable forecast. On the other side, machine learning methods such as Random Forest have a great potential and will most likely provide a more accurate forecast. Finally, the alternative option based on volumetry and the request made by each store also constitutes a possible solution especially if well parameterized and provided with the measurements and the right inputs needed.

6.2 Limitations and further work directions

This work has some limitations.

The first one is regarding the length of the data set. The historical data used in this dissertation is constituted by the year 2019 and beginning of 2020. This can affect the performance of the statistical models, highly reliable on historical data to capture trend and seasonality. For instance, since TBATS model is able to model multiple seasonal patterns, having more years of data could be interesting for it to be able to model more seasonal patterns.

Still concerning the data set, the number of pallets shipped per day and the number of boxes utilized in Random Forest model, as an explanatory variable, have their origin in different information systems of the company. Thus, there might be some inconsistencies in the data that could be affecting the performance of the random forest model. Furthermore, this could be the reason why this variable was not able to improve the performance of the Random Forest model for the Wholesaler and supermarket stores.

Another limitation is the capacity of the Random Forest model optimal features obtained for the representative stores to be applied to the other stores of their respective clusters. In order to improve this, increasing the number of variables considered to perform the clusters could be an option. By doing this, the similarity between the stores in the same cluster could be enhanced, as more characteristics would be taken into consideration. Therefore a model built for a store from each one of these newly calculated clusters would probably have a better performance.

Additionally, the multivariate model could be improved by including further explanatory variables. Moreover, more methodologies could be tested, both statistical and machine learning methods. Respecting Neural Networks in particular, more network architectures could be analyzed.

Besides improving the accuracy of the estimate, the process of performing the forecast could be changed so that it wasn't so dependent on an employee to perform it. All methodologies proposed, if implemented at the company, would currently depend on an employee to perform them and update the data to be fed to the methods. An integration of these forecasting methods in the information system or the implementation of a simple robotic process automation (RPA) system, for example, would make the estimate process less influenced by human errors.

Furthermore, the variability in the number of pallets shipped per day to each store is an issue that should be tackled. Although there is a visible weekly seasonality in the shipment's volume to each store, there are also relevant differences in this volume between homologous days, causing an irregular pattern. The estimated accuracy and operation would benefit from reducing this variability in the shipment size. If the shipment size to each store was more stable it would facilitate operations, for instance, workforce planning and scheduling, distribution planning, and so on.

To conclude, although this dissertation has limitations and room for improvement the forecasting framework presented along it can be applied to numerous situations.

Bibliography

Alon, I., Qi, M., and Sadowski, R. J. (2001). Forecasting aggregate retail sales:. *Journal of Retailing and Consumer Services*, 8(3):147–156.

Arunraj, N. S., Ahrens, D., and Fernandes, M. (2016). Application of SARIMAX Model to Forecast Daily Sales in Food Retail Industry. *International Journal of Operations Research and Information Systems*, 7(2):1–21.

Aye, G. C. ., Balcilar, M., Gupta, R., and Majumdar, A. (2013). Department of Economics Forecasting Aggregate Retail Sales : The Case of South. (February):1.

Bratina, D. and Faganel, A. (2009). Forecasting the Primary Demand for a Beer Brand Using Time Series Analysis. *Organizacija*, 41(3):116–124.

Chakure, A. (2020). Random Forest Regression - Towards Data Science. https://towardsdatascience.com/random-forest-and-its-implementation-71824ced454f.

Chu, C. W. and Zhang, G. P. (2003). A comparative study of linear and nonlinear models for aggregate retail sales forecasting. *International Journal of Production Economics*, 86(3):217–231.

Coimbra, E. A. (2013). *Kaizen in logistics and supply chains*. McGraw-Hill Professional.

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

Di Pillo, G., Latorre, V., Lucidi, S., and Procacci, E. (2016). An application of support vector machines to sales forecasting under promotions. *4or*, 14(3):309–325.

Diário de Notícias (2019). Em dois anos abriram em Portugal quase 300 supermercados - DN. https://www.dn.pt/edicao-do-dia/09-dez-2019/em-dois-anos-abriram-em-portugal-quase-300-supermercados-11596245.html.

Günther, H. O. and Seiler, T. (2009). Operative transportation planning in consumer goods supply chains. *Flexible Services and Manufacturing Journal*, 21(1-2):51–74.

Hbner, A. H., Kuhn, H., and Sternbeck, M. G. (2013). Demand and supply chain planning in grocery retail: An operations planning framework. *International Journal of Retail & Distribution Management*, 41(7):512–530.

Hosseini, S. D., Akbarpour Shirazi, M., and Karimi, B. (2014). Cross-docking and milk run logistics in a consolidation network: A hybrid of harmony search and simulated annealing approach. *Journal of Manufacturing Systems*, 33(4):567–577.

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

Imai, M. (2005). *Gemba Kaizen*. McGraw-Hill Professional.

Jain, A., Menon, M. N., and Chandra, S. (2015). Sales Forecasting for Retail Chains. pages 1–6.

Ma, S., Kolassa, S., and Fildes, R. (2018). Retail forecasting: research and practice Evaluation and development of forecasting approaches applicable to Internet-based telecommunications View project Sensory-cognitive and Physical Fitness Training in Mild Cognitive Impairment View project. (October).

Makridakis, S., Spiliotis, E., and Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLoS ONE*, 13(3):1–26.

Martins, S., Amorim, P., and Almada-Lobo, B. (2018). Delivery mode planning for distribution to brick-and-mortar retail stores: discussion and literature review. *Flexible Services and Manufacturing Journal*, 30(4):785–812.

Pavlyshenko, B. M. (2019). Machine-learning models for sales time series forecasting. *Data*, 4(1):1–11.

Pedro Curvelo (2019). Vendas do retalho crescem 3,4% em 2018 para 20.945 milhões - comercio - Jornal de Negócios. https://www.jornaldenegocios.pt/empresas/comercio/detalhe/vendas-do-retalho-crescem-34-em-2018-para-20945-milhoes.

Srivastava, P. (2017). Long Short Term Memory - LSTM. https://www.analyticsvidhya.com/blog/2017/12/fundamentals-of-deep-learning-introduction-to-lstm/.

Sternbeck, M. G. and Kuhn, H. (2014). An integrative approach to determine store delivery patterns in grocery retailing. *Transportation Research Part E: Logistics and Transportation Review*, 70(1):205–224.

Toyota Motor Corporation (2020). Toyota Production System | Vision & Philosophy | Company | Toyota Motor Corporation Official Global Website.

Tyralis, H. and Papacharalampous, G. (2017). Variable selection in time series forecasting using random forests. *Algorithms*, 10(4).

Van Belle, J., Valckenaers, P., and Cattrysse, D. (2012). Cross-docking: State of the art. *Omega*, 40(6):827–846.

Zhang, G. P. and Qi, M. (2005). Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research*, 160(2):501–514.

Appendix A

Holt-Winters additive and multiplicative model components

The components for the additive model are as follows:

$$
\begin{aligned}
\text{Forecast equation} \qquad & \hat{y}_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)} \\
\text{Level equation} \qquad & \ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\
\text{Trend equation} \qquad & b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \\
\text{Seasonality equation} \qquad & s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}
\end{aligned}
$$

The components for the multiplicative model are as follows:

$$
\begin{aligned}
\text{Forecast equation} \qquad & \hat{y}_{t+h|t} = (\ell_t + hb_t)s_{t+h-m(k+1)} \\
\text{Level equation} \qquad & \ell_t = \alpha\frac{y_t}{s_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1}) \\
\text{Trend equation} \qquad & b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1} \\
\text{Seasonality equation} \qquad & s_t = \gamma\frac{y_t}{(\ell_{t-1} + b_{t-1})} + (1-\gamma)s_{t-m}
\end{aligned}
$$

Appendix B

ARIMA and Seasonal ARIMA model equations

The ARIMA model can be described according to Equation 1, written in backshift notation.

$$\begin{array}{ccccc} (1-\phi_1 B-\cdots-\phi_p B^p) & (1-B)^d y_t & = & c+(1+\theta_1 B+\cdots+\theta_q B^q)\varepsilon_t \\ \uparrow & \uparrow & & \uparrow \\ \mathrm{AR}(p) & d \text{ differences} & & \mathrm{MA}(q) \end{array} \tag{1}$$

In backshift notation,

$$B^n y_t = y_{t-n} \tag{2}$$

The SARIMA model can be described according to Equation 3.

$$\phi_{\mathrm{p}}(\mathrm{B})\Phi_{\mathrm{p}}\left(\mathrm{B}^{\mathrm{S}}\right)(1-\mathrm{B})^{\mathrm{d}}\left(1-\mathrm{B}^{\mathrm{S}}\right)^{\mathrm{D}} y_{\mathrm{t}} = \theta_{\mathrm{q}}(\mathrm{B})\Theta_{\mathrm{Q}}\left(\mathrm{B}^{\mathrm{S}}\right)\varepsilon_{\mathrm{t}} \tag{3}$$

Where $\phi_{\mathrm{p}}(\mathrm{B})$ is the auto-regressive operator of order p, $\theta_{\mathrm{q}}(\mathrm{B})$ is the moving average operator of order q, $(1-\mathrm{B})^{\mathrm{d}}$ is the differencing operator of order d, y_{t} is the dependent variable and ε_{t} the residual error in SARIMA model. Regarding the seasonal components, $\Phi_{\mathrm{p}}(\mathrm{B})$ is the Seasonal auto-regressive operator with p-order, $\Theta_{\mathrm{q}}(\mathrm{B})$ is the seasonal moving average operator with q-order, $(1-B)^D$ is the seasonal differencing operator of order D, and S the seasonal length.

Appendix C

TBATS model equations

Model:

$$y_t^{(\lambda)} = l_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + d_t$$

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

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

$$d_t = \sum_{i=1}^{p} \varphi_i d_{t-i} + \sum_{i=1}^{q} \theta_i e_{t-i} + e_t$$

Seasonal part:

$$s_t^{(i)} = \sum_{j=1}^{(k_i)} s_{j,t}^{(i)}$$

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos(\omega_i) + s_{j,t-1}^{*(i)} \sin(\omega_i) + \gamma_1^{(i)} d_t$$

$$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} \sin(\omega_i) + s_{j,t-1}^{*(i)} \cos(\omega_i) + \gamma_2^{(i)} d_t$$

$$\omega_i = 2\pi j / m_i$$

Where:

$y_t^{(\lambda)}$ - time series at moment t (Box-Cox transformed)

$s_t^{(i)}$ - ith seasonal component

l_t - local level

b_t - trend with damping

d_t - ARMA(p,q) process for residuals

e_t - Gaussian white noise

Model parameters:

T - Amount of seasonalities

m_i - Length of ith seasonal period

k_i - Amount of harmonics for ith seasonal period

λ - Box-Cox transformation

α, β - Smoothing

ϕ - Trend damping

φ_i, θ_i - ARMA(p, q) coefficients

$\gamma_1^{(i)}, \gamma_2^{(i)}$ - Seasonal smoothing (two for each period)

M. Source: https://medium.com/intive-developers/forecasting-time-series-with-multiple-seasonalities-using-tbats-in-python-398a00ac0e8a, 11-04-2020, 19:30

Figure 48: Tbats model equations

Appendix D

Performance indicators

$$wMAPE = \frac{\sum_{i=1}^{n} ABS\left(\mathrm{y}_i - \hat{y}\right)}{\sum_{i=1}^{n} \mathrm{y}_i} * 100\% \tag{4}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(y_i - \hat{y}\right)^2}{n}} \tag{5}$$

$$NRMSE = \frac{RMSE}{y_{max} - y_{min}} \tag{6}$$

Where y_i represents the observations and $\hat{y}$ the predicted values, with n being the number of observations.

Appendix E

Top 20 stores by the number of pallets shipped per day

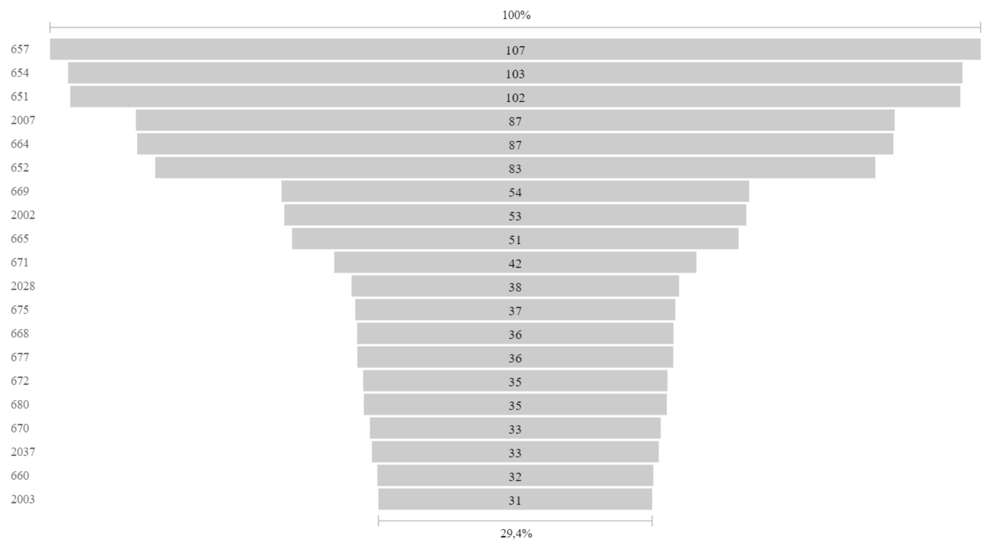


Figure 49: Top 20 stores with the highest average number of pallets shipped per day

Appendix F

ABC-XYZ analysis scatter plot

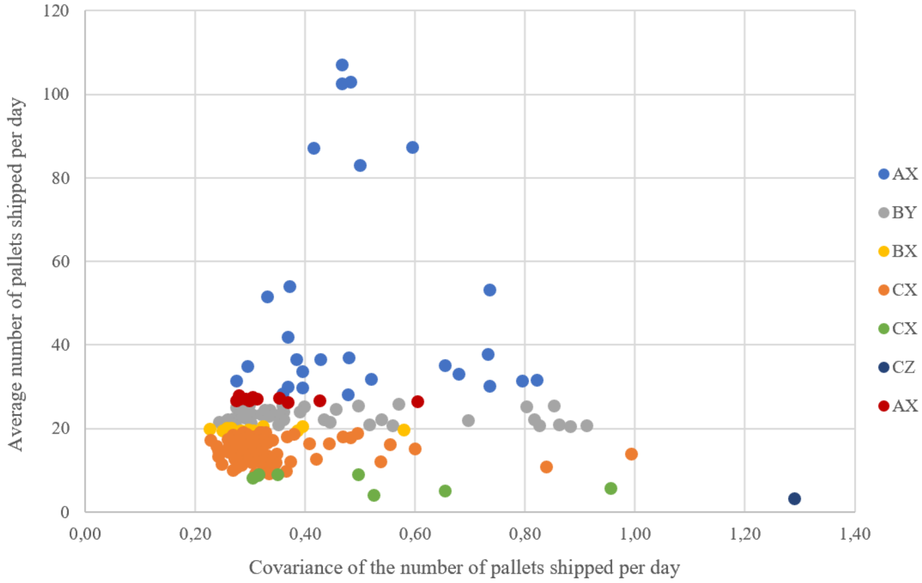


Figure 50: Scatter plot of stores organized by groups

Appendix G

Autocorrelation plots for the representative stores

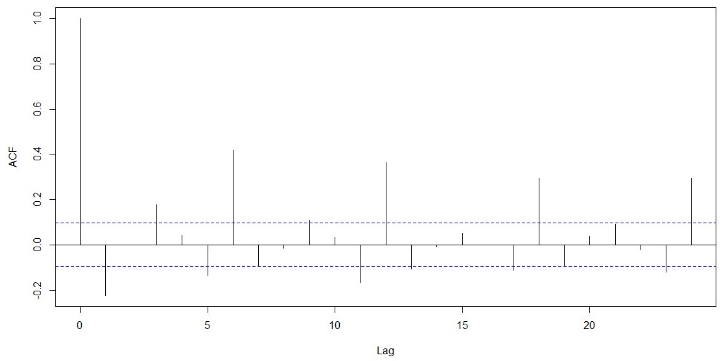


Figure 51: Autocorrelation plot of the number of pallets shipped per day to store 657

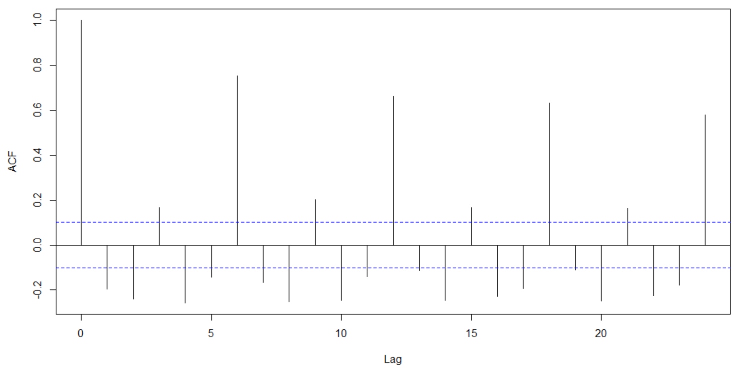


Figure 52: Autocorrelation plot of the number of pallets shipped per day to store 2022

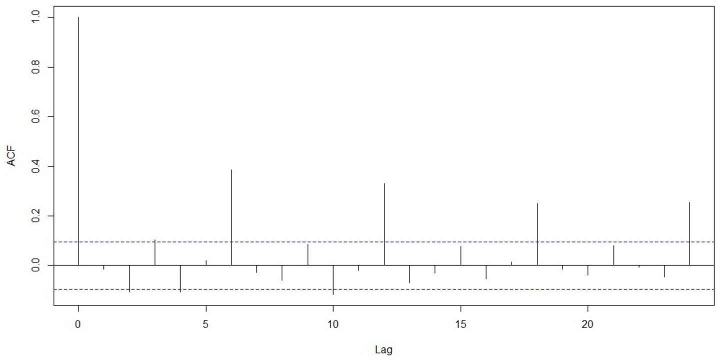


Figure 53: Autocorrelation plot of the number of pallets shipped per day to store 347

Appendix H

Standard for the performance of the estimate with the forecasting tool

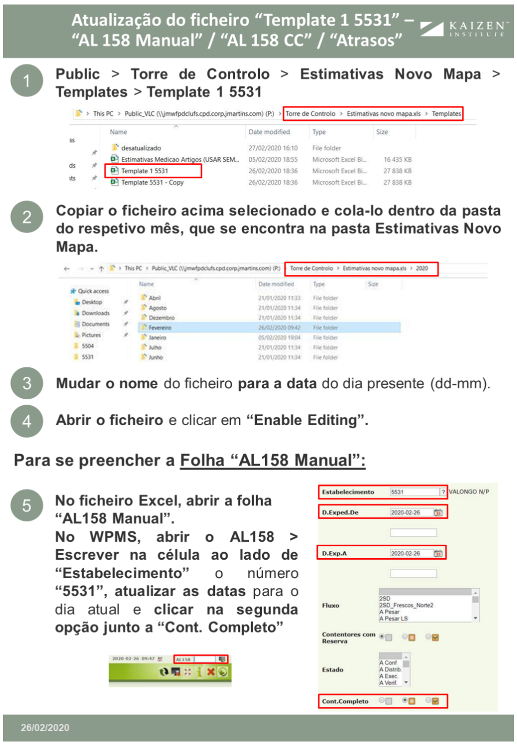


Figure 54

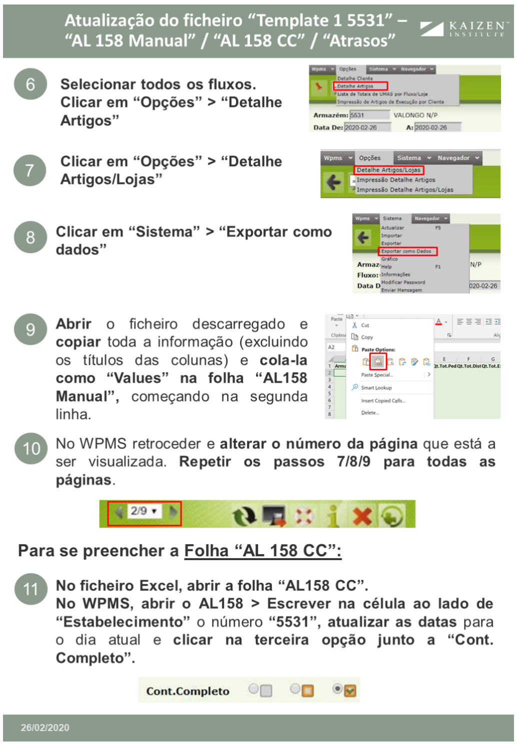


Figure 55

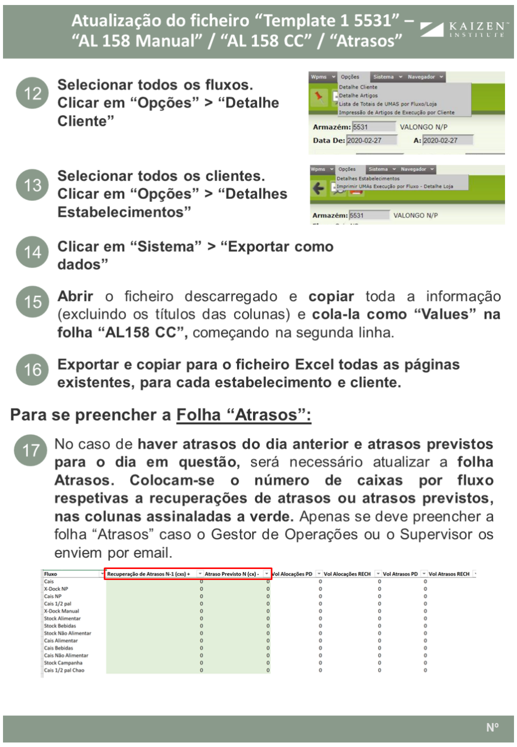


Figure 56

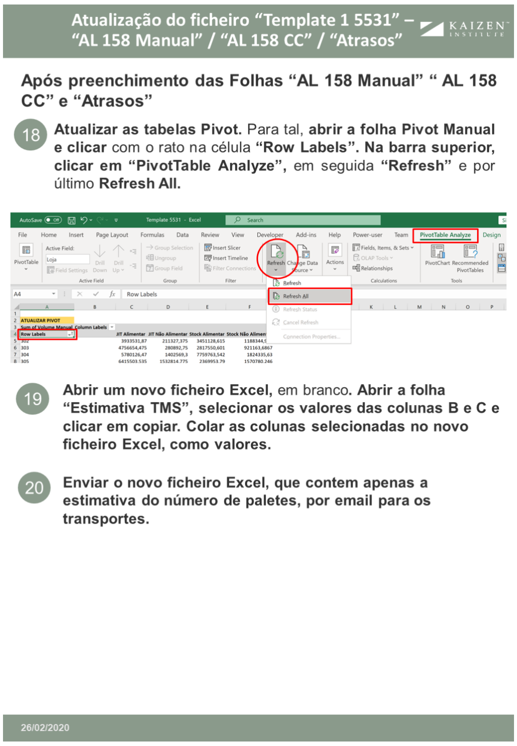

Atualização do ficheiro "Template 1 5531" – "AL 158 Manual" / "AL 158 CC" / "Atrasos"
KAIZEN INSTITUTE
Após preenchimento das Folhas "AL 158 Manual" " AL 158 CC" e "Atrasos"
18
Atualizar as tabelas Pivot. Para tal, abrir a folha Pivot Manual e clicar com o rato na célula "Row Labels". Na barra superior, clicar em "PivotTable Analyze", em seguida "Refresh" e por último Refresh All.
19
Abrir um novo ficheiro Excel, em branco. Abrir a folha "Estimativa TMS", selecionar os valores das colunas B e C e clicar em copiar. Colar as colunas selecionadas no novo ficheiro Excel, como valores.
20
Enviar o novo ficheiro Excel, que contem apenas a estimativa do número de paletes, por email para os transportes.
26/02/2020

Figure 57

Appendix I

Augmented Dickey-Fuller test results

```
        Augmented Dickey-Fuller Test

data:  ts(dados657$Paletes)
Dickey-Fuller = -5.8191, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

Figure 58: Output of the augmented Dickey-Fuller test for store 657

```
        Augmented Dickey-Fuller Test

data:  ts(dados2022$Paletes)
Dickey-Fuller = -5.2271, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

Figure 59: Output of the augmented Dickey-Fuller test for store 2022

```
        Augmented Dickey-Fuller Test

data:  ts(dados347$Paletes)
Dickey-Fuller = -6.0324, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

Figure 60: Output of the augmented Dickey-Fuller test for store 347