THE PERFECT FORECAST FOR CYCLING

By

Yorick Beekman

University of Twente

Bachelor Thesis Industrial Engineering and Management (IEM)

Faculty of Behavioural, Management and Social Sciences (BMS)

Supervisor University of Twente: Matthieu van der Heijden

Second Supervisor University of Twente: Wouter van Heeswijk

Supervisor Company X: Manager Operations

“There are two kinds of forecasters:

those who don’t know,

and those who don’t know they don’t know.”

~ John Kenneth Galbraith

Preface

The report you are about to read is the result of a research conducted at Company X. At Company X I was able to execute my bachelor assignment which is related to the process of forecasting. With fulfilling this assignment, I will meet the requirements needed in order to graduate from the study program Industrial Engineering and Management at the University of Twente, The Netherlands.

A special thanks goes to my supervisor at Company X and all the other people, whether they were involved in my project or not, for letting me be able to do my bachelor assignment at Company X, their contribution to the research and showing support.

Also, I would like to thank my supervisors from the University of Twente, Matthieu van der Heijden, for his feedback and suggestions to make sure the report is of academic level, and Wouter van Heeswijk, for being the second reader of this report. At last, I would like to show gratitude towards my family and friends for always believing in me during my study and research.

I sincerely hope you will enjoy reading my bachelor thesis.

Yorick Beekman

Management Summary

Introduction

Company X is a company based in The Netherlands and specialized in the development of cycling accessories. These accessories are being designed in The Netherlands and produced mostly in Asia at suppliers of Company X. The majority of the customers is based in Europe. However, in the last couple of years, other countries have shown interest as well, also outside of Europe. This led to a growth of the company. In 2020, due to the COVID-19 pandemic a lot of people started cycling. This led to Company X' customers ordering more to keep their stores full. For Company X, this period was very uncertain, but their sales have seen a huge increase. Although 2020 could not have been predicted by anyone, the forecast of Company X might be a problem in the future. Some products like OEM (Original Equipment Manufacturer) or PL (Private Label) do not need forecasting, but lots of other products do and Company X believes this could be done more accurate.

In order to improve the accuracy, a time series forecasting model can be used to predict future demand. This forecast should be made on a product level to determine how many units of a product need to be ordered at the supplier. Furthermore, the forecast should have a horizon of one year since there are some products that do not sell often and have an MOQ (Minimum Order Quantity), thus these products might only be ordered twice a year. Moreover, using a forecast horizon of one year, Company X could inform its suppliers in advance before they will send the actual order to the suppliers. This helps the suppliers to prepare for the coming orders. At last, the time buckets of the forecast should be one month. Using larger time buckets, the information to base decisions on will not be accurate enough. Smaller time buckets would not be possible to forecast accurately.

Forecast Models

Before models can be tested, data has been collected and analysed. Products that are not in the scope of forecasting have been deleted from the dataset and outliers have been changed to a better fitting value. After that, characteristics of the data have been determined. To do so, trend and seasonality analysis have been conducted based on data from the years 2017, 2018, 2019 and 2020. From this, we found that the data shows both a significant trend as well as a significant seasonal pattern. Once the data characteristics are known, possible solutions have been listed. Based on the characteristics of the models and the objectives set by the team as well as the characteristics of the demand data, possible solutions have been determined. Three models have been chosen to be tested: Holt's, Holt-Winters Additive and Holt-Winter's Multiplicative. For the Holt-Winter's models, these have been tested with three kinds of seasonality. This could be seasonality per product individually, per product group (Categorized Seasonality, based on the groups: *bags, baskets, accessories* and *other*) or when all demand has been aggregated (Aggregated Seasonality).

Testing these models, 4 years of data have been used (2017, 2018, 2019 and 2020) which are split up into a training set and a test set. The models have been tested based on the forecast made for 2020 with 3 years of training data. Since 2020 was a remarkable year that does not represent normal demand patterns well, also 2019 has been used as test set with 2017 and 2018 as training set.

Main Findings

Based on the tests, we find that Holt-Winter's Multiplicative with Categorized Seasonalities scores just as good as Holt-Winter's Multiplicative with Aggregated Seasonalities. To determine which model should be chosen as the solution, the ease of computing the forecast has been taken into account. This led to Holt-Winter's Multiplicative with Aggregated Seasonalities being the chosen solution. Using the Weighted Absolute Percentage Error (WAPE) per product, a Weighted WAPE has been calculated with the sales numbers being the weight of a product's WAPE. Using the model, the 12-month Weighted WAPE resulted in an error of 41% for both 2019 and 2020. Looking at the tracking signal for 2019, we see slight underforecasting whereas in 2020 the underforecasting is significant. Underforecasting in 2020 was expected due to the unexpected rise in demand.

Some types of SKUs could be forecasted better than others. This was the case for *Accessories* and *Carriers*. These groups have a Weighted WAPE of around 30% which can be explained by a seasonal pattern that repeats itself. For *Plastic Baskets* and *Mounting Systems* the error was higher, 49% and 57% respectively. This is either due to an erratic demand pattern or a seasonal pattern that does not repeat itself in the forecasted year, making it hard to predict future demand.

Evaluation

After it became clear which model is the solution, it should be compared to the accuracy of the current forecasting method. From this comparison it can be seen that the chosen solution scores better in most cases. Especially looking at the 2019 forecast, which represents a more normal demand pattern, it can be seen that the solution is more accurate. The chosen solution had a Weighted WAPE of 41% compared to 48% for the current method. Also, where the model has a low error, the current method has a low error. Where the model has a high error, the current method scores bad as well. One exception is the group *Metal Baskets,* in this case the model its error was 15% lower than the current method. Which is a result of product XXXX having a much lower error in the model. Since this product has a large sales volume, it has a great influence on the Weighted WAPE. For 2020, the model and current method both had an error of 41%. Looking at the tracking signal for 2020 both methods have been underforecasting, as expected. However, for 2019 it can again be seen that the chosen solution outperforms the current method with a much less biased forecast. Same goes for the root mean square error. For 2020, this error is relatively the same for both the model and the current method. For 2019, the RMSE of the chosen solution is significantly lower than the RMSE of the current method.

Solely using the modelled forecast already shows a better accuracy than the current method. However, to improve the accuracy further, forecasts of the customers of Company X can be used to find large deviations from the forecast. When such a large deviation is found, it can be discussed with that customer what their intentions are or additional information about competitors can be gathered. Adjusting the forecast accordingly could increase the accuracy of the forecast. Furthermore, a new categorization has been introduced based on how low the WAPE for a product has been in the past 12 months. If the WAPE is lower than 30%, it can be said the model predicted the sales for that product accurately. If it is above the boundary of 30%, input from the sales team is definitely needed. However, products with a WAPE below 30% should not be skipped, it is always good to take other indicators into account as these indicators could give valuable information about future demand.

Conclusions and Recommendations

From the finding in this research it became clear that Holt-Winter's Multiplicative model with Aggregated Seasonalities is the most accurate at predicting future demand for Company X. For this model an Excel prototype is being made such that Company X can use this model to help them in the process of making a forecast. Following the forecasted values exactly is never a wise thing to do, so the forecast will always have to be evaluated by the team. In order to improve the accuracy even more, using the customer forecast as well as other indicators present in the Excel prototype will be beneficial. Changing the forecast accordingly will help increase the accuracy.

The main drawback of the Excel model is that after a period, the demand values and outstanding orders have to be updated manually, as well as the modelled and adjusted forecast for the accuracies. Once this data has been added, the model will recalculate the values for level, trend, seasonality and the forecast automatically as well as the indicators and accuracies in the Excel dashboard. Also, after some months the smoothing constants will have to be optimized again. Using an automated program, these manual steps can be done more efficient and less error prone.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

OEM	= Original Equipment Manufacturer
PL	= Private Label
MOQ	= Minimum Order Quantity
MPSM	= Managerial Problem Solving Method
TS	= Tracking Signal
MAPE	= Mean Absolute Percentage Error
WAPE	= Weighted Absolute Percentage Error
RMSE	= Root Mean Square Error
MAD	= Mean Absolute Deviation
SKU	= Stock Keeping Unit
(S)ARIMA	= (Seasonal) Autoregressive Integrated Moving Average
NOS	= Never Out of Stock

Forecast Method Abbreviations

CM	= Current Method
HM	= Holt's Model
HWA	= Holt-Winter's Additive
HWM	= Holt-Winter's Multiplicative
HWAC	= Holt-Winter's Additive with Categorized Seasonalities
HWMC	= Holt-Winter's Multiplicative with Categorized Seasonalities
HWAA	= Holt-Winter's Additive with Aggregated Seasonalities
HWMA	= Holt-Winter's Multiplicative with Aggregated Seasonalities

1 Introduction

The first chapter of this research will provide information about the company and an overview of what this research is about. The introduction to Company X will be handled in section 1.1. After that, problem identification will be conducted in section 1.2. After the problems have been identified, the core problem will be chosen in section 1.3. Next, the objectives of the research are elaborated on in section 1.4 and at last the approach on how to solve the problem is considered in section 1.5

1.1 Company Introduction

The company in question is a family owned company based in The Netherlands which was founded in the second half of the 20th century. It all started with a customer walking into the bike store of the previous owner of Company X and asking for a reed basket. At that point in time they did not sell those baskets, but the bike store made sure one was made for the customer to enjoy cycling even more. This was the starting point of Company X. In the beginning only reed baskets were made, which were imported from former Yugoslavia. Four years later, also steel baskets are part of the product range and products are produced mostly in Asia.

Currently, Company X is a successful business and has earned multiple awards. Like the innovation award for their innovative mounting mechanisms and supplier awards because of being proactive, inspirational, consumer focussed and more. Currently, Company X has around 30 employees in total, of which around 20 work from the office.

Nowadays, the company is specialized in the development of multiple types of cycling accessories. Panniers, baskets, bells and raincoats are some products in the product range. In total, there are around 600 products the company sells. New product designs are made in the product department of Company X. Once a new product has been approved, Company X makes use of their suppliers (mostly) in Asia, in countries like India and China. This is where suppliers produce the products of Company X. After that, the products are shipped to The Netherlands to Company Y, relatively close to the office location of Company X. Here the products are kept in stock. Since Company Y is also specialized in the distribution of products, they also distribute the products to retailers mostly located in The Netherlands, Belgium and Germany for Company X. In the last couple of years other countries like the United States of America, New Zealand and Australia showed interest as well. This leads to a growth in sales orders and Company X might not be fully prepared for this higher amount. One of the problems Company X faces here is the process of forecasting. The method the company currently uses is not accurate enough. Assuming the sales will rise in the coming periods, this accuracy needs to be improved. With a more accurate forecast, costs will be decreased and delivery reliability will be increased.

1.2 Problem Identification

Since Company X is expanding its business to other countries as well, there is a rise in sales orders. Those orders also include orders from customers who have never done any business with Company X before. In this case, it is very hard to estimate the forecast of the following year since there is no data from that company or country yet. Furthermore, there is no use of a forecasting model at all, which makes the process of forecasting time consuming and inconsistent. To prepare for the future, the current forecasting process should be analysed and improved such that Company X can do better forecasting. In this section the action problem will be presented and the core problem will be identified by using a problem cluster.

Since the process of forecasting has become more and more difficult for Company X, an action problem arises. The action problem can be described as the forecast accuracy being too low. This accuracy has been calculated by means of the Mean Absolute Percentage Error (MAPE) and is fluctuating around 70%.

To find out what all the problems are, interviews with employees have been conducted as well as observations of the forecast meetings. Together with conversations in the company, a problem cluster could be made. In the problem cluster, all problems are stated that might be a possible cause for the low accuracy. As depicted in figure 1, there are multiple problems that could influence the forecast accuracy.

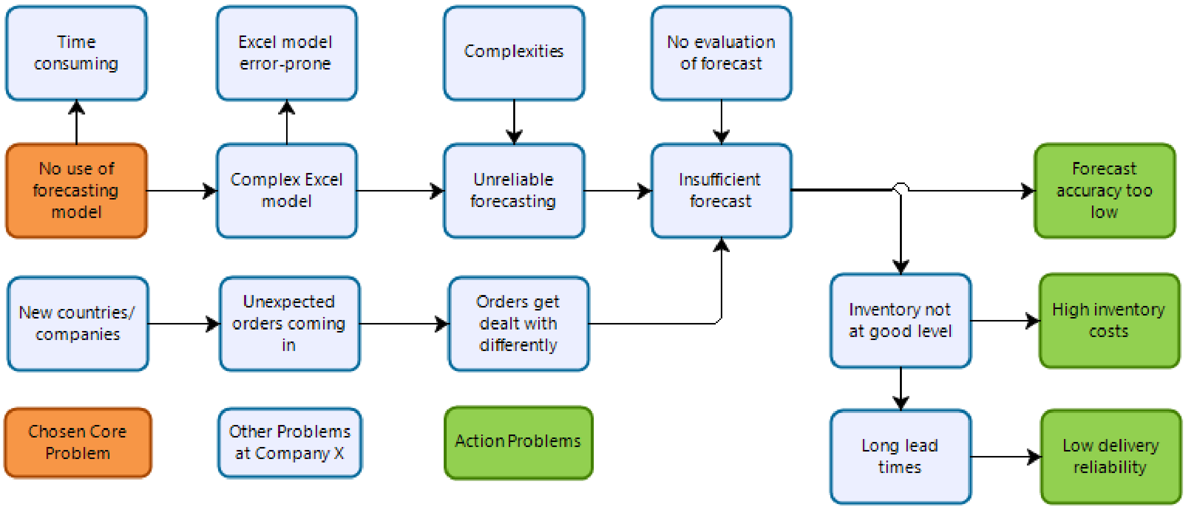


Figure 1 Problem Cluster

A possible core problem is the problem of *complexities*. Already in the beginning the company indicated there are a lot of smaller things going on during the process of forecasting that are not standard. Examples of these complexities could be a sudden shift in demand, the performance of competitors or not knowing what amount of the forecast is meant for pre-orders.

Another problem is *no evaluation of the forecast*. Employees never look at any forecast measures to find out how good or bad the forecast in previous periods was. This means there has not been a change in the method of forecasting since the employees believe it was going well.

Furthermore, the action problem can be influenced by *new countries or new companies*. In this case, there is no data yet about a new country or customer which makes it hard to forecast. One way to overcome this is by looking at the forecast of other customers or by analysing similar products in that particular country. This way it can still be possible to adjust the forecast in such a way that the accuracy is as low as possible.

The last possible core problem is *no use of forecasting model*. The current method is a qualitative way of forecasting where a team decides what the forecast of a product should be. It could be beneficial for Company X to implement a forecasting model to be prepared for the future.

Another problem that will be investigated are the product categories. These are used to divide products into four categories to determine how much attention they need during forecasting. This is not mentioned in the problem cluster since it will definitely be part of the research and thus is not considered when searching for the core problem. These product categories have been implemented in 2015 and it is said that the categories helped during the forecasting. The problem here is that the product categories are not being used anymore and the questions arises whether products are in the right product category as well as whether the criteria that products are being assessed on are the right criteria.

1.3 Core Problem

To find a solution to the action problem, the core problem should be identified. The core problem is the problem that will be most important and has the greatest impact. According to Heerkens and Van Winden (2017), to find out what the core problem could be, four things should be considered.

1. It should be a real problem that occurs in the company
2. The core problem cannot have a direct cause
3. If the problem cannot be influenced, it cannot be a core problem
4. If multiple problems remain, choose the one with highest impact at lowest cost

Looking at the abovementioned aspects, it can be concluded that the problem of having *new countries/companies* cannot be influenced since these new customers arrive when they want. There might be solutions to overcome the problems that arise from this, but it will not solve the problem at its core. The problem of new countries/customers is not chosen as a core problem.

The problem of *complexities* consists of multiple problems of which some can be influenced and some not. Also here, there might be solutions to overcome the problems that arise from this, but it will not solve the problem at its core. Since this consists out of multiple problems, it will take a considerable amount of time to find solutions to all these problems. It is not believed this is possible in the given amount of time and the expected benefit does not match the effort that is needed to solve the problems. For these reasons, this is not chosen as a core problem.

In this case there are two problems left, namely *no use of forecasting model* and *no evaluation of forecast*. Both these problems could be a core problem since they both meet the first three criteria. Thus, it has to be determined which problem will have the greatest impact at lowest cost. The problem that will have the highest impact is assumed to be *no use of forecasting model*, since this is the basis of the whole forecasting process. Solving this problem will cost more effort than finding evaluation metrics, but the benefit from it will be higher.

To conclude, from the abovementioned findings the chosen core problem is: *no use of forecasting model*. This problem meets all criteria to be a core problem and it is expected that solving this problem has the highest impact at the lowest costs.

1.4 Objectives

There are two objectives that can be identified in this research. As mentioned in the previous section, one of these is finding a correct forecasting model. The other one is to improve the product categorization. The following part will explain what the objectives for these solutions are.

1.4.1 Forecast Model

The main focus will be on finding the correct forecast model since this is the core problem. In order to solve this, it should be clear what the objectives are in the research concerning the forecast model.

The goal of this research is to improve the forecast accuracy. This will be checked by means of forecast measures. The way to achieve this goal is by using historical sales data. Based on this data, it will be possible to make a time-series forecast that can be tested on its accuracy. However, a forecast should never be fully trusted. When noticing a low forecast accuracy, adjustments should be made by the forecast team.

A new forecast for the coming calendar year is made at the end of every year. The forecast is made for the whole year to see how well it matches the budget that has been set. Also, they can inform the suppliers about the amount they expect to order so suppliers can prepare for this. The forecast that is made at the end of a year will be adjusted every month. According to Chopra & Meindl (2016), the forecast horizon should be greater than or equal to the lead time of the decision that is driven by the forecast. The decision that has to be made can be described as ‘when to order what amount’, because of this, the forecast horizon should be at least four months. Furthermore, the time buckets of the forecast are originally one month. However, the forecast for one month is divided by the number of weeks in that month to get a weekly forecast. This is done since some suppliers deliver once every three months, but there are also suppliers that deliver once every two weeks. According to the Managing Director it is also not possible to make a forecast per week. It is way too small; it would not be accurate and too time consuming to make. For these reasons time buckets of a month are preferred. Moreover, the forecast should be done on product level. There should be enough historical data for most of the products to make a sufficient forecast. However, there are some more recent products of which the amount of data could be lacking and need to be excluded in the determination of the correct model. Moreover, there are already some products being left out of the scope of forecasting.

Products not included in the scope of forecasting are:

- OEM Products (Original Equipment Manufacturer)
- PL Products (Private Label)
- Rainwear
- Service Products

Main reason for leaving OEM and PL out is because there is no need to forecast them since these orders are identical to the purchasing orders to the suppliers. The rainwear is being left out since these can only be ordered at Company X a couple times a year at certain points in time. For the service products, there is no need to forecast them. These are slow-movers and are rarely used. Leaving these products out, 315 products remain in the scope of forecasting.

Product categories that are in the scope are:

- Bags
- Baskets
- Accessories
- Other

Furthermore, it is not preferred to find a model that needs programming. The employees do not have knowledge about programming and it is believed that using any model that needs programming will make the process of forecasting more difficult for the employees.

At last, Company X indicated they would like to make use of the forecast of their customers. It is believed that those forecasts are of great value. In order to make use of these forecasts, a template can be designed in which trends of customers can be detected. With this template, the forecast team should be able to adjust the forecast if necessary.

All in all, the objective is to find an understandable model for the products in the scope that will make the forecast more accurate than it currently is, without the use of programming. The model should be able to forecast with time buckets of one month and a forecast horizon of one year. This model will be implemented in an Excel prototype to make a forecast. Also an implementation plan will be set up indicating how the forecast should be made. Ultimately, it would be great to see the model being implemented in the systems currently used by Company X such that forecasting becomes a fully automated process. However, this will require programming and knowledge about the systems. This will be too time consuming to execute in the given time frame.

1.4.2 Product Categorization

The second objective is to improve the current product categorization. This should help the forecast team in making the forecast more accurate. Doing so, the forecast improves and costs can be minimized. The current way the classification is done is by using an adjusted Kraljic matrix. At Company X, the axes *predictability of the forecast* and *revenue* have been introduced since it is not a purchasing strategy, but a forecasting strategy. Figure 2 depicts the Kraljic matrix as introduced by Company X. Products that have a bad predictability and generate much revenue, category A, should be reviewed with most care. Products with good predictability and low revenue, category D, are the least important. From interviews with the managers it became clear category C is more important than category B.

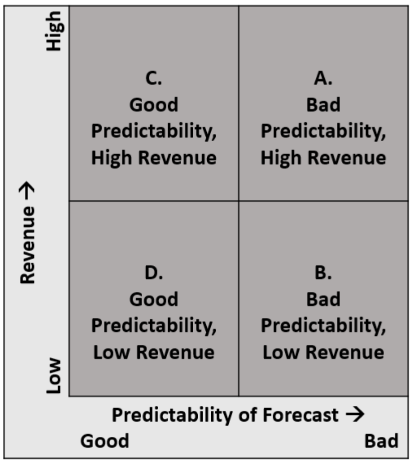


Figure 2 Kraljic Matrix as Introduced by Company X

The new forecast team does not use this product categorization anymore. Main reason according to the Managing Director and Supply Chain Planner is the low number of SKUs. During the forecast meetings they want to address all SKUs individually and not skip any because of good predictability.

To determine in which revenue category a product belongs, there is no clear line between high and low. This is mostly based on how important the product is going to be in terms of profit margin and numbers sold. A product that generates X amount of revenue with a smaller number of products might be in the high revenue category, while a product that also generates X amount of revenue with a lot of products might be in the low revenue category. For the predictability, it becomes even more vague since a bad predictability can be caused by the sales department, customers or performance of competitors. Also here, there is no clear definition of what is good and what is bad. One thing is known, not all products are in the right product category. However, this has never been changed.

Even though the categories have no use anymore, new products are still placed in a category. This is done by the Collection Manager. The categorization is based on multiple aspects. These could include a market research or looking at similar products, but since there are no clear boundaries, it is mostly based on a gut feeling. This counts for both the categorization of revenue as well as the predictability of the forecast. After all, the objective here is to find product categories that are suitable in the process of forecasting such that the accuracy of the forecast can be improved.

1.5 Approach

The research design will be following the guidelines of the *Managerial Problem Solving Method* (MPSM) as indicated by Heerkens and Van Winden (2017). They developed a process consisting out of seven steps that will lead to the solution of an action problem, figure 3 illustrates these seven steps. The first step is to identify the problem, which has been done in sections 1.2 and 1.3, here it also became clear how the forecasting is currently done. Then, a problem solving approach should be formulated. Which takes place in this section. After that, step three of the MPSM will be addressed, the problem is further analysed to find out what data is available, what the current situation is and how the data behaves. Step four and five are about finding possible solutions to the action problem and the best model will be chosen. After the forecast model has been determined, it will be implemented in step six. At last, in step seven the solution should be evaluated to check how it has contributed to solving the action problem. Once the evaluation is done, it could be that a new action problem arises which needs further research. Then the MPSM should be applied again with the new action problem.

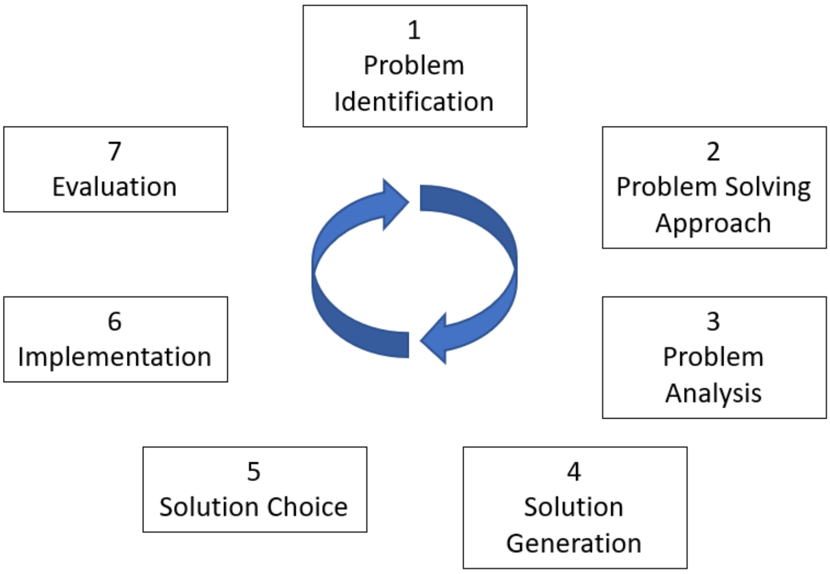


Figure 3 Managerial Problem Solving Method

1.5.1 Problem Analysis

First, it is needed to find out how good the forecasting method is before a solution will be tested. This will be done by conducting a current situation analysis. Forecast measures will be used to evaluate the current forecast. During this step, it will become clear if some products have a higher error in the forecast than others and why this might be the case.

Furthermore in this phase of the approach, it is important to find out what data is available and useful for testing a model. Since forecasting models heavily rely on historical data, it is of great importance this data is valid. Data that is not sufficient to test a model should be left out of the research, this is to prevent inaccurate results

Data that remains should be analysed to find out if there is a trend and/or seasonality in the data. Trend and seasonality are very important criteria when searching for possible solutions since not all forecasting models can handle these components.

Before continuing to the next step of the approach, some research questions need to be answered in this section. Research questions that will be answered are:

- How is forecasting currently done?
- What data is available for testing a model?
- Can a trend be detected in the data of Company X?
- Can seasonality be detected in the data of Company X?

1.5.2 Solution Generation

After the problem analysis step, it is needed to find out which models could possibly fit the data characteristics. Possible solutions will be formulated after a literature research has been conducted. In order to choose the correct model, the models found during the literature research have to be evaluated. This evaluation will be done based on the characteristics of the data and by checking whether the model meets the criteria set by the managers. With this evaluation, it should become clear whether a model might be useful for Company X or that it is irrelevant and could be left out of the research.

After that, the categorization of the products should be determined. The categorization of the products will be done together with the employees. After the categories have been decided on, it should become clear whether products are already in the correct product group, that they need to change groups or that it first has to be determined in which group it belongs in the case of a new categorization method.

At last, Company X would like a solution on how to use the forecast made by their customers. It is believed their forecasts are of great value and can improve the accuracy of the forecast of Company X.

Also during this step some research questions should be answered before continuing to the next step. In this step, the research questions will be:

- Which models are available in the literature?
- What criteria are important to choose a model?
- Which models should be tested in the research?
- Based on which forecast measures should the tested forecast models be evaluated?
- How should products be categorized?
- Are products in the correct category?
- How can forecasts of customers be used?

1.5.3 Solution Choice

To find the most appropriate model, they have to be tested. Historical data will be used to test the model. For this, it is important to have enough historical data. This data will be split up into two groups, a training set and a test set. In general, the training set accounts for 75% of the data and the test set accounts for 25% of the data.

Since the forecast has to be made for a year, one year will be the used as test set. This way the forecast accuracy can be best compared to the current accuracy. So, one year equals 25% of the historical data that will be used. Thus, in total four years of historical data will be used that will be split up into three years of training data and one year of testing data. This also is in line with the amount of data needed to forecast. According to Hyndman and Kostenko (2007), to forecast with Holt-Winter's model, at least 17 observations are needed in the case of monthly demand. For ARIMA models, this is at least 16 observations when using monthly demand. This means it would also be possible to build the model based on two years' worth of training data to test for 2019, since 2020 was such a remarkable year because of the rise in demand due to COVID-19. Also, a combination of 3 years of training data where possible and 2 years of training data where needed could be an option to test the forecast of 2020. This in case some products did not exist in 2017.

In this step, a solution will be chosen from the selected models. This solution will be determined by answering one important research question. After getting an answer to this question, it is possible to continue to the next step in the research. The research question for this section will be:

- Which forecast model performs best based on the forecast measures?

1.5.4 Evaluation and Implementation

The model that will be selected should be compared to the current forecast method. A model scoring better on the forecast measures compared to the other models, does not necessarily mean the chosen solution is better than the current forecast method. To test if the forecast has improved, the forecast measures from the chosen solution will be compared to the measures of the current method with the same products. After this has been done, it can be decided if the chosen solution is more accurate or not.

Main finding in this section is going to be whether the solution should be fully implemented in the company or not. To come to a conclusion here, one research question is very important:

- What is the effect of the new solution compared to the current method?

After it has become clear whether the solution scores better, it could be implemented in an Excel prototype. For this, it should be clear to the employees how the model works. One way of doing so is by making a clear dashboard in Excel that shows what the forecast for a product should be. Also measures to check how well the forecast is performing on a particular product and indicators about future demand could be added. With this, it should be clear whether the forecast for a certain product has to be adjusted or not.

For the implementation, one research question will be answered as well:

- What parameters should be included in the forecast dashboard?

1.5.5 Implementation

If it happens to be the case that the new method works better than the previous method, a full implementation for Company X should be considered. For this, it is useful to present the model in an understandable manner such that employees that are not technical can still understand how forecasting is done. Furthermore, it should be clear how the forecast works. This could be described in a step-by-step approach. This way, employees will be able to forecast in the future or implement it in another system.

In this section it should become clear how the process of forecasting should be done in the future, which arises the question:

- How can the solution be implemented at Company X?

1.5.6 Conclusions and Recommendations

In the last part of this research, the main findings will be summarized in the final conclusion. Next to that, also additional recommendations will be elaborated on. This way it becomes clear what the main findings in the report are and what is recommended to Company X. Also, the recommendations will make it possible to do further research in the future where needed.

2 Data Analysis

As stated by Shrestha and Bhatta (2018): "In time series analysis, it is important to understand the behaviour of variables, their interactions and integrations over time. If major characteristics of time series data are understood and addressed properly, a simple regression analysis using such data can also tell us about the pattern of relationships among variables of interest." Without data analysis it is unwise to choose possible forecasting models since it is unclear what the characteristics of the data are and forecasting models are heavily reliant on certain data characteristics. First, in section 2.1 the current situation will be analysed to find out more about the current forecasting method. This will first be done using all of the products in the scope. Later, only the products that will also be tested in the models will be evaluated. Data analysis will take place in section 2.2. Here, it will become clear what the criteria are that models will need to meet when choosing a possible solution.

2.1 Current Situation

The current forecast is based on sales from previous years and the expected growth of the business the coming year, but the forecast is not being calculated by a model in any way. Furthermore, there are some conversations with big customers to adapt the forecast based on what their intentions are for the coming year. These three aspects form the basis of the current forecast method. At the end of the year, around October, a forecast for the coming year will be made. After each month, this forecast is being adjusted during the monthly forecast meetings with the intention to improve the accuracy of the forecast. During these meetings, the team addresses all SKUs that need to be forecasted. In some occasions, this takes a whole afternoon to adjust the forecast for the products.

To find out how well the current methods performs, some forecast measures will be used. "These measures can be divided into two types: directional and size." (Klimberg et al., 2010). The direction measures tell something about whether a business is overforecasting or underforecasting. The size measures tell about the deviation between forecast and demand. Looking at the size measures, these can be divided into categories again. According to Hyndman (2014), forecast measures can be divided into three categories: scale-dependent errors, percentage errors and scaled errors. Where scale-dependent errors are on the same scale as the data. According to Hyndman (2014): "percentage errors have the advantage of being scale-independent, and so are frequently used to compare forecast performance between different data sets." At last: "scaled errors were proposed by Hyndman and Koehler (2006) as an alternative to using percentage errors when comparing forecast accuracy across series on different scales." (Hyndman, 2014).

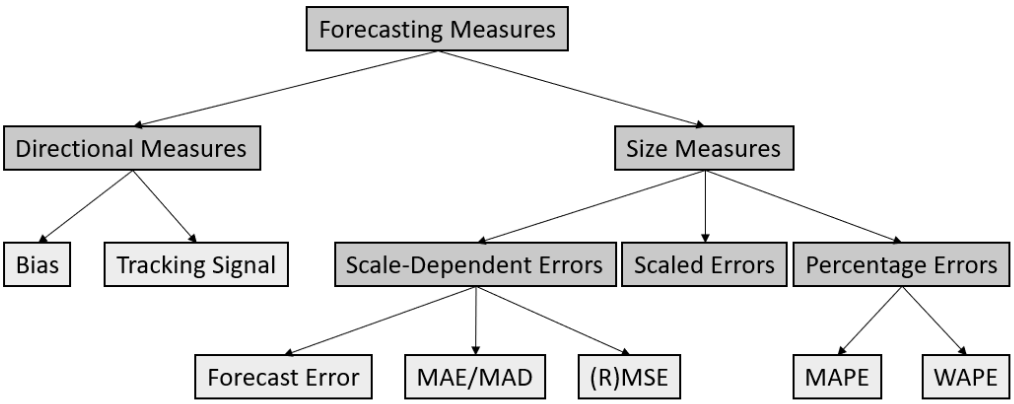


Figure 4 Overview of Forecast Measures

To find out the accuracy of the current forecast, the tracking signal will be determined to find out the direction of the forecast. Furthermore, the Mean Absolute Percentage Error (MAPE) and the Weighted Absolute Percentage Error (WAPE) will be calculated. Also the performance per SKU-type will be evaluated. In appendix A and B the RMSE and MAD are evaluated as well.

The accuracy of the current forecasting method will be based on two different forecasts, the original forecasts and the adjusted forecasts. The original forecast is the forecast made in October. The adjusted forecast will be the forecast as it was four months before delivery date, since this is the lead time of the suppliers. There will also be calculations concerning the forecast one month prior to delivery, these results will be stated in the appendix since this horizon is too short for the company to adapt to large deviations. Furthermore, the forecast for the year 2019 and 2020 will be reviewed. Reason for this is that it is expected that the year 2020 will have a higher error in the accuracy because of the COVID-19 pandemic. Moreover, when calculating any forecast measure, it is needed to check if the data that is used is valid. This will have to be checked for every product before making any calculations.

At first only the products in the scope of the forecast will be evaluated. This means rainwear, OEM, PL, service products, marketing products or any forecast unrelated items should be excluded from the dataset. In the end, the dataset for the 2019 forecasts contains 221 products. For 2020 there are 249 products in the dataset.

2.1.1 Mean Absolute Percentage Error

According to Kim and Kim (2016), the MAPE is one of the most widely used measures in forecast accuracy, due to its advantages of scale-independency and interpretability.

The formula of the Mean Absolute Percentage Error is given by

$$MAPE_n = \frac{1}{n} * \sum_{t=1}^{n} \left| \frac{D_t - F_t}{D_t} \right| * 100 \qquad (2.1)$$

$$Where \ \ D_t = actual\ demand\ in\ period\ t, F_t = forecast\ for\ period\ t, n = number\ of\ periods$$

Since the MAPE is a measurement of error, it should be as low as possible. However, it is hard to determine if the MAPE is sufficient or not. Lewis (1982, cited in Klimberg et al., 2010) came up with a scale for the MAPE depicted in figure 5. However: "depending of the data set, as to whether there is a significant trend or seasonal component, the MAPE may under or overestimate the accuracy." (Klimberg et al, 2010). Moreover, Klimberg et al. (2010) stated: "what does it mean to have a MAD (or MSE or MAPE) of 20, except that the smaller the better." Indicating that there is no clear boundary between good and bad and it depends on the characteristics of the data. It could be that a forecast with a higher MAPE is better than a forecast with a lower MAPE, if demand patterns are hard to capture.

MAPE	**Judgment of Forecast Accuracy**
Less than 10%	Highly accurate
11% to 20%	Good forecast
21% to 50%	Reasonable forecast
51% or more	Inaccurate forecast

Figure 5 Judgement of MAPE

There is however one main disadvantage. As stated by Hyndman and Koehler (2006), the MAPE has the disadvantage of being infinite or undefined if $D_t = 0$ for any period t in the period of interest, and having an extremely skewed distribution when any value of D_t is close to zero. To overcome this so-called *infinite error issue* in the analysis, the following has been done: if the demand of a certain period is 0, the Absolute Percentage Error of that month is set to 100%. However, if it is the case that the demand in a period is 0 and the forecast for that period is 0 as well, the error in that month should be 0% since there is no error.

To find out the MAPE for all demand, a weighted MAPE has been used. In the weighted MAPE the number of sales of a product has been taken into account since products with higher demand are more important to Company X. This has been done by dividing the number of sales of a product by the total number of sales of all products. This fraction accounts for the weight this product gets assigned. To find the overall MAPE, the values of all the weighted MAPEs have been summed. Table 1 depicts the results of the weighted MAPE, based on period 1 to 12 for both 2019 and 2020.

Table 1, Weighted MAPE Current Situation

Weighted MAPE	*2019*	*2020*
Original Forecast	80,50%	58,75%
Adjusted Forecast	81,27%	61,89%

What can be concluded from the MAPE as depicted in table 1 is that the forecast accuracy of 2020 is better than the accuracy of the forecast in 2019, which was not expected because of the COVID-19 pandemic. One side note to this is that underforecasting is penalized less heavily than overforecasting. This is because the error will be divided by the actual demand.

At last, it cannot be concluded that the monthly forecast meetings have a positive effect on the accuracy, this can be seen by the fact that the adjusted forecast has a higher error than the original forecast. However, the adjusted forecast for 1 month prior to delivery date has a much better accuracy. This can be seen in appendix C. Also in appendix C the MAPE with aggregated demand is calculated. To see the errors per month, the Absolute Percentage Error per month of the aggregated demand has been graphed. This shows whether the error in a month increases if the horizon increases. These results are shown in appendix D.

2.1.2 Weighted Absolute Percentage Error

A variant of the MAPE is the Weighted Absolute Percentage Error (WAPE). This measure overcomes the so-called infinite error issue of the MAPE since it divides by the sum of the demand. The WAPE can be calculated by

$$WAPE_n = \frac{\sum_{t=1}^{n} |D_t - F_t|}{\sum_{t=1}^{n} |D_t|} * 100 \tag{2.2}$$

Table 2 depicts the results of the weighted WAPE. The weight for a product is calculated in the same way as for the weighted MAPE.

Table 2, Weighted WAPE Current Situation

Weighted WAPE	*2019*	*2020*
Original Forecast	51,30%	46,92%
Adjusted Forecast	52,63%	45,67%

From table 2, it can again be concluded the forecast from 2020 has been more accurate than the forecast of 2019. However, in this case the adjusted forecast of 2020 is a slight improvement on the original forecast. The results with the forecast one month prior to delivery and when using aggregated demand can be found in appendix E.

2.1.3 Tracking Signal

The tracking signal is the ratio between the bias (sum of the errors) and the Mean Absolute Deviation (MAD). According to Chopra and Meindl (2016), the bias is a useful forecast measure if demand suddenly rises or drops. For Company X, demand suddenly rose in 2020 because of the pandemic. With the tracking signal, it will become clear if Company X underforecasted or overforecasted in 2019 and 2020.

The tracking signal of a period can be calculated by dividing the sum of the errors by the MAD of that period.

$$TS_t = \frac{bias_t}{MAD_t} \tag{2.3}$$

Once the tracking signal has been calculated, is can be determined if Company X is under- or overforecasting. If the tracking signal exceeds the value of 6, it can be said that Company X is overforecasting. The opposite goes for when the tracking signal is lower than -6, then it can be said the company is underforecasting. According to Chopra and Meindl (2016), a tracking signal outside the ±6 range could be explained by a forecasting method that is flawed or by the fact the underlying demand pattern has shifted.

The tracking signal for every product in every period has been determined. After that, the total number of tracking signals is counted by multiplying the number of products by the number of periods. Also, the total number of tracking signals that is higher than 6 or lower than -6 have been counted. This will get an overview of the number of overforecasted and underforecasted periods, table 3 depicts these values. Appendix F also shows these results combined with the adjusted forecast one month prior to delivery.

Table 3, TS Current Situation

Tracking Signal	*2019 (2652 Periods Total)*	*2020 (2988 Periods Total)*
Original Forecast		
Overforecasting	452 (17,0%)	45 (1,5%)
Underforecasting	113 (4,3%)	473 (15,8%)
Adjusted Forecast		
Overforecasting	397 (15,0%)	53 (1,8%)
Underforecasting	89 (3,0%)	365 (12,2%)

During the calculations of the tracking signal, the bias has to be calculated. If the bias is positive, the forecast is higher than the actual sales. When the bias is negative, it could be said the forecast is too low. Advantage of the bias is the ease of the calculation, but it can be interpreted wrongly since it is scale dependent. Table 4 shows the results of the bias.

Table 4, Bias Current Situation

- CONFIDENTIAL -

Furthermore, the tracking signal with aggregated demand has been calculated and can be seen in figure 6. Here it becomes visible that halfway through 2020, when the demand suddenly increased much because of COVID-19, the tracking signal dropped significantly. However in the periods before, the tracking signal was merely positive and, in some cases, even exceeded the value of 6. This can be explained by the fact that at Company X, they want to achieve a certain budget. During the forecast meetings it is often the case that they add a little extra to the forecast in order to reach the budget. This can cause the significant overforecasting.

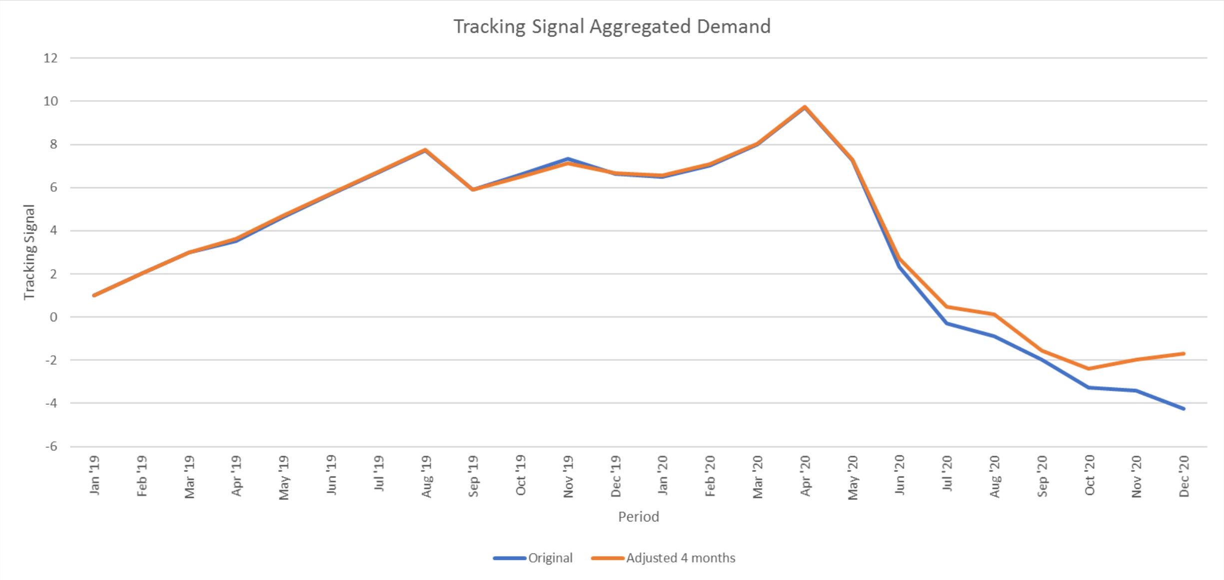


Figure 6 TS Aggregated Demand Current Situation

From table 3 and 4 and figure 6, it can be concluded that Company X has been overforecasting in the year 2019. Looking at 2020, it can be concluded that the company has been underforecasting. Contradictory to the previously mentioned measures, the adjusted forecast scores significantly better in most cases. In appendix G, the graphs clearly show the improvement of the tracking signal between the original and the two adjusted forecasts, moreover the graphs give a better insight in the distribution of the tracking signals.

2.1.4 Performance per SKU-type

It could be that a certain type of SKU scores better or worse in general. To check this, the SKUs have been divided in multiple groups to see if a specific characteristic of the product has a worse or better forecast than other characteristics. The products will be divided in the groups *Colours, Types* and *Introduction Year.* Where the types are: Bags, Baskets, Accessories and Others. For the introduction year, the boundary has been set on 3 years from the forecasted year. To analyse, it will be counted how many times products with certain characteristic are in the top 20 or bottom 20 of the performance based on the WAPE and the tracking signal.

Looking at the performance based on the WAPE, there is no clear distinction between characteristics in the groups, as can be seen in figure appendix H. If a certain characteristic of a group is frequently in the top 20, it does not mean it will be less in the bottom 20. Also, some characteristic score better in 2019, but score worse in 2020. The other way round also happens, a bad scoring characteristic in 2019 can score much better in 2020. However, looking at the introduction year for the 2020 forecast measures, it can be said that newer products are often scoring worse than older products. This could be explained by the fact that there is not much known about those products yet which makes it harder to forecast.

Taking the tracking signal into consideration, it cannot be concluded a certain colour or type has been under- or overforecasted systematically. This can also be seen in appendix I. Also in this case, it could be that in 2019 a certain colour or type has been underforecasted, but in 2020 this was not the case anymore. Moreover, one type can both be underforecasted a lot as well as overforecasted. So there is also no distinction there. This could however indicate that those products are hard to predict. Looking at the introduction year in the 2019 forecasts, it can be seen that there are a lot of under- and overforecasted products that have been introduced after 2016. This again indicates that newer products with less information are harder to forecast accurately.

2.1.6 Performance of Products in Test

To find out how the chosen solution will perform against the current method, it is important to know the performance of the products that will be tested with the model. In the model, the forecast for 2019 as well as for 2020 will be tested with a horizon of one year, based on data from 2017 onward. Also, tests will be done with three years of training data where possible, and two years where needed. The results of the current forecast with those products can be found in appendix J, together with the RMSE and MAD of 2019 and 2020. To find the products that will be tested in the model, products in the scope with at least four years of data are needed together with their forecasts for 2019 and 2020 as they have been made in December of the year before. The measures will be calculated per period, where one period equals one month.

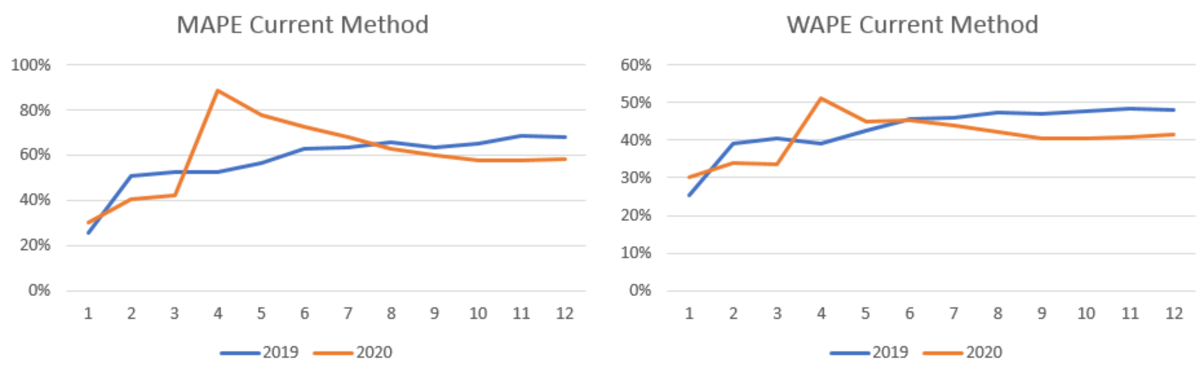


Figure 7 MAPE and WAPE of Products in the Tested Models

Starting off with the WAPE and MAPE (figure 7). It is expected that the further away the horizon is, the less accurate the forecast will be. For 2019, this can be seen clearly. However, for 2020 there is a peak in the error measures in period four. In this period, customers did not want to order since this was the beginning of the COVID-19 pandemic and no-one knew what would happen. In the months after that, the demand skyrocketed which led to a high error in period 5, 6, 7 and 8 as well. After this, demand was more as expected and so the expected increase in the error can be seen.

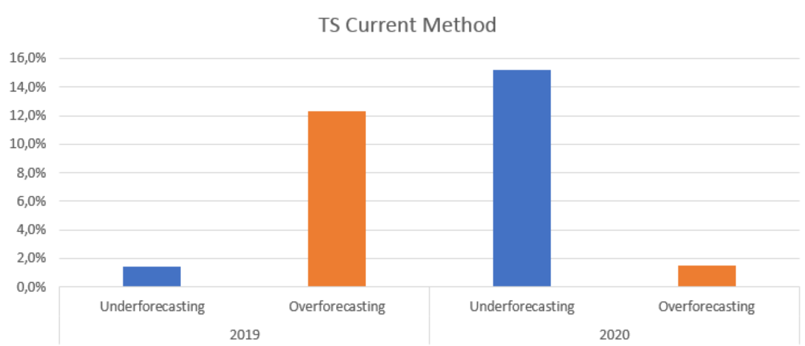


Figure 8 TS of Products in the Tested Models

Figure 8 depicts the tracking signal. Here it can be seen that also the products that will be tested with a model are being overforecasted too much in 2019, this could again be explained by the fact that

Company X sets a budget they want to reach. In 2020 clear underforecasting can be seen, as expected due to the sudden high demand following from the COVID-19 pandemic.

2.1.5 Current Situation Conclusion

From the analysis of the current situation, it can be concluded that the forecast has improved in the year 2020 compared to 2019. This was not the expectation of the team because of the sudden increase in demand due to COVID-19. An explanation for this is that the overforecasting has made a buffer that has been used for the sudden rise in demand, leading to a smaller error. Furthermore, as expected, in 2020 the team has been underforecasting. For 2019, it was assumed that Company X has been slightly overforecasting, but the tracking signal shows that Company X has been overforecasting significantly. It can be assumed this has been happening in the years before 2019 as well, since the same method has been used in the years before 2019.

Looking at the difference between the original and the adjusted forecasts, it can be concluded that the adjusted forecast often does not score better than the original forecast. This was not expected since more information is known, so it should be more accurate.

At last, it cannot be concluded that certain colours or product groups have been forecasted better or worse. However, it can be seen that newer products are more often under- or overforecasted, or that the WAPE is worse. Which is not unusual due to the lack of information for these products.

For an overview of all results, see appendix K.

2.2 Data Analysis

Before solutions can be chosen, it should become clear what the characteristics of the data are. In order to do so, products with at least four years of data will be used. This is done in order to capture at least four cycles. With four years of data, 129 products remain. This data analysis can be seen as the initialization step of a Holt-Winter's model, where trend and seasonality also have to be calculated. In the analysis, it will be checked if trend and/or seasonality occur. First in section 2.2.1, demand will be deseasonalized and it will be checked if there is a significant trend in the data. Once this has been checked, the occurrence of seasonality will be evaluated in section 2.2.2. Based on these characteristics, the possible models will be evaluated in section 3.

2.2.1 Trend Analysis

According to Chopra and Meindl (2016), it is not appropriate to run a linear regression between the original demand data and time. This is because the original demand data are not linear and the resulting linear regression will not be accurate. To overcome this, the demand has to be deseasonalized before running the linear regression. "Deseasonalized demand represents the demand that would have been observed in the absence of seasonal fluctuations" (Chopra & Meindl, 2016). To determine what the deseasonalized demand for a product is, the following formula will be used:

$$\overline{D}_t = \frac{\left[D_{t-\frac{p}{2}} + D_{t+\frac{p}{2}} + \sum_{i=t+1-\frac{p}{2}}^{t-1+\frac{p}{2}} 2D_i\right]}{2p} \tag{2.4}$$

$$Where\ \overline{D}_t = deseasonalized\ demand, D = observed\ demand, p = 12, t = period$$

P stands for the periodicity, which is the number of periods after which a cycle repeats itself. Since we deal with monthly demand, the periodicity is 12.

After demand has been deseasonalized, linear regression can be run. The linear regression will estimate the initial values of the level and the trend. This is done with the deseasonalized demand as dependent variable and the time as independent variable.

Inspecting the data, it becomes clear that some products do not show a significant trend. However, there are multiple products that do show a clear trend, either positive or negative. In total, 47 out of 129 products show a negative trend. The remaining 82 products show a positive trend.

Also a linear regression with aggregated demand has been run. This resulted in an initial level of AAAA and a trend of BBBB. To determine if the trend in the sales numbers is significant, the p-value can be taken into account. The p-value should be lower than 0,05 to accept the null-hypothesis, which would mean there is a trend. Since the p-value from the linear regression is 0,003 it can be concluded the trend is significant. To see the detailed results of the linear regression for aggregated demand, see appendix L. Now it has become clear the data shows a trend. This means that the models that will be tested are preferred to handle trend.

2.2.2 Seasonality Analysis

Now that the deseasonalized demand has been calculated, it should be checked if the data shows seasonality. To check for seasonality the seasonal factor $\overline{S}_t$ should be obtained, which is the ratio of actual demand D_t to deseasonalized demand $\overline{D}_t$. This will be done using equation 2.5.

$$\overline{S}_t = \frac{D_t}{\overline{D}_t} \tag{2.5}$$

After the seasonal factors for every period individually have been determined for the four cycles, the average seasonal factor for a month can be calculated. On average the seasonal factor is 1, if the seasonal factor deviates much from 1 it can be said there is a seasonality. To calculate the average seasonal factor, formula 2.6 is used.

$$S_i = \frac{\sum_{j=0}^{r-1} \overline{S}_{jp+i}}{r} \tag{2.6}$$

In this formula r = 4 since data from 2017, 2018, 2019 and 2020 is used.

Also here, looking at the products individually, seasonality can be spotted for most of the products. However, looking at each product one by one is not efficient and thus the seasonal factors for aggregated demand have been determined. The results are shown in figure 9. From this, it can be seen the seasonal factors fluctuate between 0,53 and 1,42. This indicated strong seasonality.

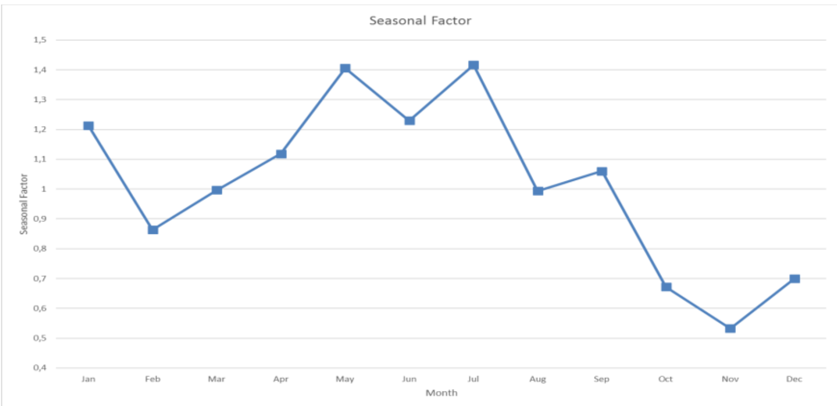


Figure 9 Seasonal Factors Aggregated Demand

2.3 Conclusion

After the data analysis, it has become clear what the characteristics of the data are. It has been concluded there is a significant trend in the data for most of the products. This could be both a positive as well as a negative trend. Moreover, a significant seasonality has been determined in the data. Also here, there might be a difference between individual products, but looking at the aggregated demand, a strong seasonality can be seen. This is useful to know when searching for possible models. Also following from the findings in this section, the demand patterns in 2020 are not very representative due to COVID-19. For this reason, when testing possible solutions, also 2019 will be taken into account since this will represent the demand patterns better. In appendix M also demand patterns have been reviewed, from these results the model should be able to handle smooth demand. Following from the objectives stated in 1.4.1, the models should also be able to forecast with a horizon of one year and time buckets of one month. All in all, it could be said the models are preferred to handle both trend and seasonality, and should be modelled on smooth demand. The models should also be able to forecast for one year with time buckets of one month.

3 Solution Generation

Now it is clear what the characteristics of the data are, possible models will have to be evaluated. The models that will be used for further investigation are preferred to handle the characteristics of the data. Also, the models should meet the requirements set by the forecast team, which are stated in section 1.4.1.

3.1 Forecast Models

"Forecasting methods fall into two major categories: quantitative and qualitative methods." (Hyndman, 2011). This can be seen in figure 10. First, quantitative models, these can again be split up into two sub-categories, namely time-series forecasting and causal forecasting. With time-series forecasting, a model is being used that is based on historical data and it is assumed that components will repeat itself. Costs of implementing time-series models are very low and time to implement the model can range from one day to one month based on the complexity of the data and the model. The second sub-category consists of causal forecasting methods. These methods assume that the dependent variable that has to be forecasted has a cause-effect relationship with one or multiple independent variables. These independent variables could be sales, weather forecast, product features, social chatter and much more. Implementing causal forecasting methods might take up to one month and depending on the method, can be rather expensive (Brillio, 2018). The other category consists of qualitative forecasting methods. These methods apply the knowledge of a forecast team about the business, product, market and customer to come to a conclusion. This is partly what currently happens at Company X during their monthly forecast meetings. Costs of forecasting are high and implementation time is said to be around two to three months. In this research, the time-series forecasting models will be further investigated.

For time-series forecasting, again two different categories can be distinguished, namely the static forecasting methods and the adaptive forecasting methods. As stated by Chopra and Meindl (2016), the static method assumes the level, trend and seasonality do not change as new demand is observed. Parameters are estimated and the same values will be used for all future forecasts. With adaptive forecasting the estimates of the level, trend and seasonality will be updated after a new observation. Advantage of this is that the estimates of the level, trend and seasonality incorporate all data available. Adaptive forecasting models are preferred since the assumption that level, trend and seasonality do not change cannot be made for Company X.

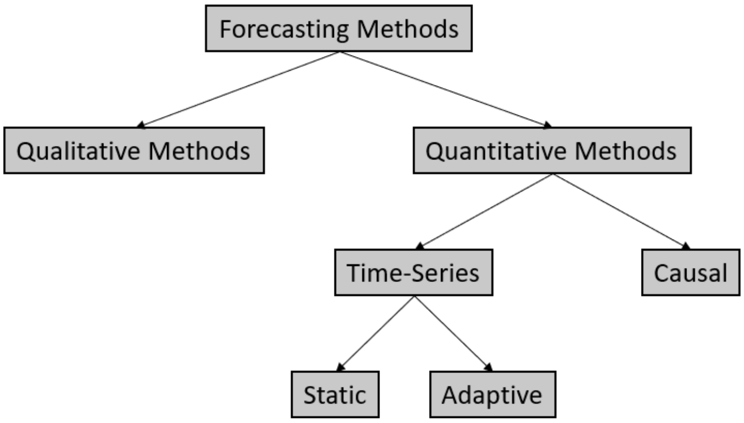


Figure 10 Overview of Forecasting Techniques

3.1.1 Moving Average

The first forecast model that will be evaluated is the moving average. The main advantage is its simplicity, because of this, it is easy to understand for non-technical people and easy to implement which makes it doable in the given time frame. Also, to calculate a moving average, not much data is needed.

However, there are also some important drawbacks. The main disadvantages are that this model is not good at handling trend since it will lag if demand rises or drops, this can be explained by the fact that this method is solely based on past demand values. Moreover, as stated by Holt (2004, cited in Booranawong & Booranawong, 2017), the accuracy of the forecast will significantly decrease when applied to data which shows seasonality patterns. At last, the moving average is not very accurate when forecasting with a horizon of a year.

3.1.2 Simple Exponential Smoothing

The next model is simple exponential smoothing. Using simple exponential smoothing, the forecast for period *t+1* can be calculated via

$$F_{t+1} = \alpha * D_t + (1 - \alpha) * F_t \tag{3.1}$$

The advantages of simple exponential smoothing are almost the same as for the moving average. When comparing the moving average to the simple exponential smoothing another advantage can be added. Ostertagová and Ostertag (2011) state that simple exponential smoothing allows the more recent values of the series to have greater influence on the forecast of future values than the more distant observations.

Of course, this model comes with its disadvantages as well. Also stated by Ostertagová and Ostertag (2011): "Simple exponential smoothing model is only good for non-seasonal patterns with approximately zero-trend and for short-term forecasting because if we extend past the next period, the forecasted value for that period has to be used as a surrogate for the actual demand for any forecast past the next period."

3.1.3 Holt's Model

Holt's model is also known as double exponential smoothing, or as indicated by Chopra and Meindl (2016), *trend-corrected exponential smoothing.* Holt's model is an expansion on the simple exponential smoothing model. The formula for the Holt's model is given by

$$F_{t+1} = L_t + T_t \tag{3.2}$$

The detailed formula can be found in appendix N. This model sums the level and trend. This will be updated once a new value has been observed. To determine the initial level and trend, a linear regression between the demand and the time will be used (Yunishafira, 2018). To optimize the forecast accuracy, the smoothing constants could be determined in such a way the error is minimized.

The most significant advantage here compared to the previously mentioned models is the fact that this model is able to handle a significant trend. Furthermore, the previously stated advantages for moving average and simple exponential smoothing also apply for Holt's model. Moreover, Holt's model will be able to forecast for a longer period in the future compared to moving average or simple exponential smoothing.

The main disadvantage here is still the fact that the model is not able to handle a significant seasonality in the data.

3.1.4 Holt-Winter's Additive/Multiplicative Model

Looking at the model that has been introduced by Holt and Winter, two different models can be identified. The additive and the multiplicative. At first, the additive model will be taken into consideration. This model is given by

$$F_{t+1} = L_t + T_t + S_{t+1} \tag{3.3}$$

In this model, the level, trend and seasonality are all added components. In the multiplicative form, also known as the mixed form, the level and trend are added before they are multiplied by the seasonal factor. The formula for the multiplicative model is

$$F_{t+1} = (L_t + T_t) * S_{t+1} \tag{3.4}$$

Also for these models, the detailed formulas can be found in appendix N.

The advantages and disadvantages between the two models are relatively the same. However, there is one main difference between the models. As stated by Kalekar (2004): “In plots of the series, the distinguishing characteristic between these two types of seasonal components is that in the additive case, the series shows steady seasonal fluctuations, regardless of the overall level of the series; in the multiplicative case, the size of the seasonal fluctuations vary, depending on the overall level of the series.” Based on the other aspects, the models have the same advantages and disadvantages. Comparing Holt-Winter’s to Holt’s model, the advantage is that Holt-Winter’s is able to capture seasonality. Other than that, the (dis)advantages are the same.

At last, the seasonality for these models can be calculated in different ways. The seasonality can be calculated for every product individually. However, it can be hard to estimate the seasonal component for an individual product because of the noise. To overcome this, aggregated demand could be used. “Based on the idea that in general the demand at an aggregate level is relatively less erratic compared with demand at the item level, it is assumed that separating the seasonal pattern from the randomness will be easier, resulting in a better estimate of the seasonal indices.” (Dekker et al., 2004) Aggregating demand can be done in different ways. It can be aggregated based on product groups, which are likely to show the same seasonality pattern, but also aggregating all demand together is an option. When the demand is aggregated, the seasonal indices can be determined. These seasonal indices can be used to forecast on a product level.

3.1.5 (S)ARIMA Model

One model that is also often used is the ARIMA model. ARIMA is short for Autoregressive Integrated Moving Average. This model is considered to be complex and it is preferred to implement the model with the use of programming. ARIMA models cannot handle seasonality very well. For this, the SARIMA model has been developed, a Seasonal ARIMA model. This model has all the advantages the previously mentioned models have as well. It can capture both trend and seasonality and it is relatively good at forecasting with a horizon of one year. However, there are some disadvantages following from this model. The main disadvantage is the fact that the model is very complicated to understand, for the detailed formula, see appendix O. This complexity is not preferred since the employees need to understand it as well. Even while it is such a complex model, it does not mean the accuracy is better. Multiple studies state that the (S)ARIMA model performs worse than exponential smoothing (Makatjane & Moroke, 2016;Köppelová & Jindrová, 2019; Syafei et al., 2018).

3.1.6 Neural Networks

At last, neural networks will be commented on. Neural networks are data-driven machine learning algorithms that mimic the functionalities of the human brain in order to solve the problem. These types of forecasting methods are being used more and more often but are quite complex. Following from Law and Au (1999), neural networks consist of input layers and output layers. Often, there are also one or more hidden layers in between. The input layers are the independent variables, the hidden layer is used to add an internal representation of the handling of non-linear data, the output value is the solution to the problem. Advantages are that the neural network can work with multiple input variables, thus it could capture all different sorts of independent variables. However, with more input variables and hidden layers, the model will become even more complex. Moreover, neural networks are known for the fact they need a lot of data, the implementation time is long and it is also expensive to implement such models. A more elaborate explaining of neural networks can be found in appendix P.

3.1.7 Forecast Model Choice

Now that possible models have been elaborated on, possible solutions should be chosen. In order to choose solutions, the models should meet some criteria. These criteria concern both the data characteristics that have been concluded earlier in section 2.2 as well as criteria set by the team in section 1.4.1 or other practical criteria.

The latter includes:

1) The model should forecast with a horizon of one year, in time buckets of one month
2) The model should be understandable for non-technical employees
3) The model can be implemented in the given timeframe
4) Programming is not expected

Regarding the data characteristics, these arise from the data analysis in section 2.2. The requirements from the data characteristics are:

5) The model should be able to handle seasonality
6) The model should be able to handle trend
7) The model should be able to forecast with two and three years of monthly data

With these criteria, the models can be evaluated to see if it meets the requirements. The assessments of these criteria can be seen in table 5.

Table 5, Criteria Assessment of Possible Solutions

	Moving Average	Simple Exponential Smoothing	Holt's	Holt-Winter's Additive	Holt-Winter's Multi.	SARIMA	Neural Networks
Team requirements							
1 year horizon	No	No	Yes	Yes	Yes	Yes	Yes
Understandable	Yes	Yes	Yes	Yes	Yes	No	No
Doable in given time	Yes	Yes	Yes	Yes	Yes	No	No
Programming is not expected	Yes	Yes	Yes	Yes	Yes	No	No
Data requirements							
Can handle trend	No	No	Yes	Yes	Yes	Yes	Yes
Can handle seasonality	No	No	No	Yes	Yes	Yes	Yes
Can use 2 and 3 years of data	Yes	Yes	Yes	Yes	Yes	Yes	No
Yes: (green) No: (red)							

Since the SARIMA model and the neural network model both are very difficult to understand for non-technical employees, are not likely to be doable in the given time and expect programming, these will not be taken into consideration in the research.

Moreover, both moving average and simple exponential smoothing will not be suitable for multiple reasons. Both models are not able to accurately predict future sales for one year. Also, both models are unable to handle trend and seasonality. For these reasons the moving average and simple exponential smoothing will not be analysed further in the research. Looking at the three models that are left, we can see that both Holt-Winter's Additive model and Holt-Winter's Multiplicative model meet all the requirements. These will be further researched. Holt's model meets almost all requirements, only it cannot capture seasonality. However, this model can still be tested. It could be the case this model will work significantly better for certain products without seasonality and scores better on the overall accuracy.

To conclude, Holt's model, Holt-Winter's Additive model and Holt-Winter's Multiplicative model will be further tested to find the best forecast per product. For Holt-Winter's models the seasonality will be tested in multiple different ways. When testing these models, the smoothing constants will be minimized by optimizing the WAPE. For determining which model will be chosen as the solution, also the RMSE as size-measures will be used and the tracking signal as measure of direction.

3.2 Categorization

Some employees of Company X introduced the Kraljic matrix as depicted in figure 2. The goal of the Kraljic matrix was to improve the process of forecasting. In their eyes, it has helped them. With the new forecast team, the product categorization is not being used anymore. The main reason for this is the fact that the current forecast team wants to address all SKUs individually, since there are not that many SKUs. For this reason, they did not see any added value in the product categories. However, most of the employees have indicated they do not mind using categories again.

With the new categorization only two categories will be used. These categories will be *Modelled Forecast* and *Modelled Forecast With Input From Sales Team*, category 1 and category 2 respectively. The way to distinguish the two categories is by means of the WAPE. This forecast measure has been chosen since it uses percentages rather than units. With percentages, products with high sales numbers and low sales numbers can have the same boundary to decide if it needs more sales input or not because of the scale independency.

The company has decided that once the WAPE over the previous 12 months of a certain product is below 30%, it could be said the forecasting method is performing up to its standard and there is no need to give these products that much attention. If the WAPE of a certain product is above 30% it will be placed in the category where the sales team will have to give their input. Their input is obviously needed in that case since the forecast is not performing up to its standard. Although 30% might seem high, it can be considered good when looking at demand patterns at Company X. Figure 11 depicts the new way of categorization that will be used.

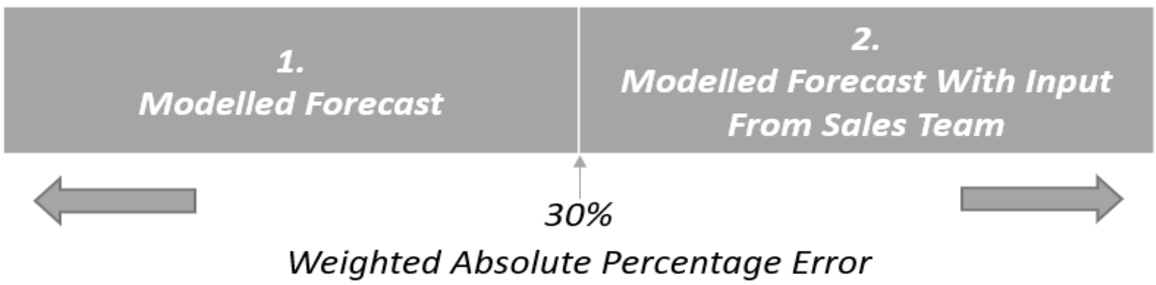


Figure 11 New Product Categorization Method

This new way of product categorization comes with multiple advantages compared to the old method as described in section 1.4.2. These advantages are mainly:

- Clear line between the categories
- No assumptions have to be made
- Easy to reassign products to a new group

All in all, the new way of categorizing products will be by means of the WAPE. If this is above 30%, a product should be reviewed by the sales team. If the WAPE is below 30%, this might not be needed. Nonetheless it is important to keep an eye on other forecast measures and indicators, even if the WAPE is under 30%.

3.3 Customer Forecasts

Currently, Company X receives the forecast from some of its customers. The company believes these are of great value and could be used to improve the forecast accuracy. After benchmarking at Company Z, it has been proven the forecasts help with improving the accuracy. In order to use the forecast of customers from Company X, the customers responsible for 80% of the revenue are selected. With this selection, it has to be checked what the share of a customer is in the sales of a product over the previous 12 months. Then, with the forecast made by Company X and the forecast from its customers, it can be checked what the expected share will be in the next 12 months. If the difference in share is larger than 7,5%, it is assumed to be a large deviation. If this is the case, it could be discussed with that customer what the reason is for this large deviation in their share. Figure 12 depicts the template for using customer forecasts.

According to Company Z, discussing the forecast with customers is very beneficial. Using their forecast, it is possible to detect large deviations in demand which could improve forecast accuracy, but also the fact that there are conversations with the customer is a good thing. These conversations improve the relation with the customer and can give loads of information, also about competitors. It could be that a customer is ordering more at Company X because a competitor cannot deliver. During the conversation, also information about the plans of the customer can be discussed. It could be that a customer orders a larger amount once and then goes back to the competitor again, or the customer could stay at Company X.

- CONFIDENTIAL -

Figure 12 Customer Forecast Template

3.4 Conclusion

Following from the literature research, it is believed there are three possible models that could fit the demand of Company X: Holt's, Holt-Winter's Additive and Holt-Winter's Multiplicative. In the case of Holt-Winter's models, the seasonality will take on different forms. This can either be seasonality per product, seasonality per product group or seasonality with all demand aggregated. For the new categorization, there are two categories a product can be in. The first category is *Modelled Forecast*, the second category is *Modelled Forecast with Input from Sales Team.* Once the WAPE of a product is below 30%, it should be placed in the first category, if it is higher than 30% it will be placed in the second category. However, *Modelled Forecast* does not mean the forecasted values do not need to be checked anymore. It is still wise to adjust the forecast if strange or unwanted values can be seen. Lastly, the use of customer forecast has been discussed. It has been proven that at Company Z the use of these forecasts improved the accuracy and ensures a better relationship with the customer. Also, during meetings with customers, information about competitors might be shared which could help Company X in predicting future demand better.

4 Solution Choice

From chapter three, it became clear which models could be possible solutions. In this chapter, the accuracy of these models will be tested based on forecasting measures. The measures that will be used are the Weighted Absolute Percentage Error (WAPE), Root Mean Square Error (RMSE) and the tracking signal. To find the overall WAPE, it will be weighted in the same way as done in chapter 2.

For testing, four years of data is used. To test the forecast for 2020, the sales from 2017, 2018 and 2019 will be used to train the model. Also for the 2020 forecast, there will be calculations based on three years training data where possible and two years where needed, in case some product did not exist in 2017. This is because the solution should also forecast products with only two years of training data if not possible otherwise. The results of the tests with two and three years of data can be found in appendix Q. Since 2020 was such a remarkable year regarding the demand due to COVID-19, the models will also be tested with 2019 as test set, with two years of training data, namely 2017 and 2018. This is done to see how the models would perform in a more normal year. The forecasts that will be tested are made with a horizon of 12 months. This means that the model would generate a forecast as it should have done at the end of December in the year previous to the test set. Based on how accurate the forecast is, the best solution will be chosen.

Holt-Winter's Additive model and Holt-Winter's Multiplicative model will be tested with multiple forms of seasonalities. These seasonalities are per product individually, per product group (Categorized Seasonality) and when all demand is aggregated (Aggregated Seasonality). In case of the categorized seasonality, the products were first divided into the four main groups that products are placed in: *bags*, *baskets*, *accessories* and *other*. Using a visual check, different seasonal patterns can be seen between groups. The groups can be made more specific by using subgroups like *steel baskets, plastic baskets, bells* etc. However, the differences between the seasonal patterns of subgroups that are in the same main group are minimal. For example, the seasonal pattern of a steel basket is almost the same as the seasonal pattern of a plastic basket. Furthermore, this would lead to a total of 18 groups and in some cases there would be only one product per group. For these reasons, Categorized Seasonality will be based on the groups *bags, baskets, accessories* and *others.*

For Holt-Winter's additive variants, in case of the categorized/aggregated seasonality, the seasonal component that has been found using the categorized/aggregated demand will be divided by the number of products used to calculate the component. After this division, it can be used to generate a forecast on product level.

4.1 Results 2020

Looking at the WAPE with 2020 as test set (figure 13), a clear peak in the error can be seen in period 4. During this period, the demand suddenly dropped due to COVID-19. Later on in period 6, a rise in demand occurred and so the error rose as well. However, period 6 was a period with a lot of underforecasting, so the error did not rise that high since underforecasting is less heavily penalized using the WAPE.

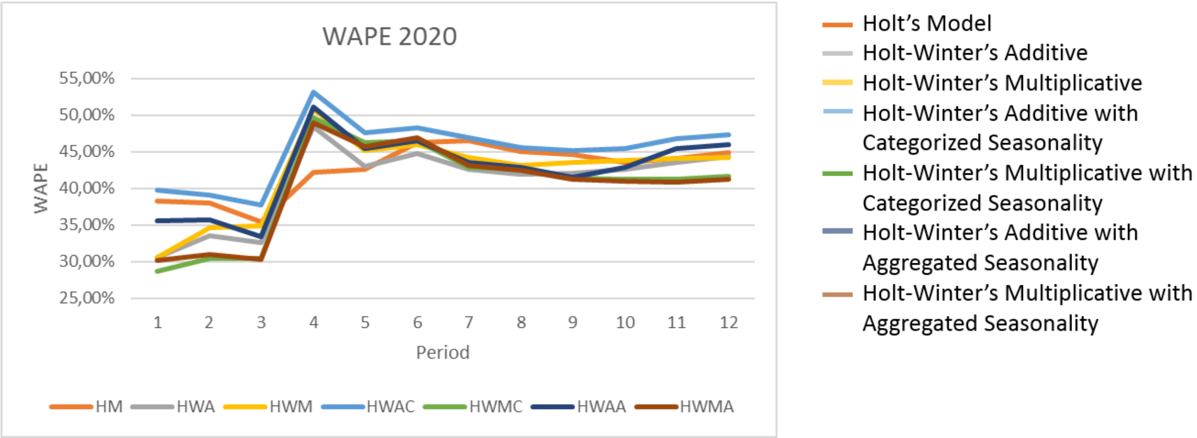


Figure 13 WAPE Tested Models Throughout 2020

In the beginning both Holt-Winter's Multiplicative with Categorized Seasonalities (HWMC) and Holt-Winter's Multiplicative with Aggregated Seasonalities (HWMA) score much better. During periods 4 to 8 the demand was quite different compared to the previous years, leading to a high error in these periods instead of a linear, increasing line what would be expected. Towards the end, demand was more as expected again, leading to an error that is lower than in periods 4 to 8. From period 9 onwards, the expected increase in error can be seen. This was the expectation since the horizon is further away. For HWMC and HWMA the error from period 9 onwards is again smaller compared to the other models.

It can also be seen that the additive variants with categorized/aggregated seasonalities score much worse than the multiplicative variants with categorized/aggregated seasonalities. Reason for this could be because some products have much lower/higher sales numbers. Since the added component is equally divided by the number of product used to calculate the added component, it might not fit all products. For large products, the added component will be too small. For small products, the added component will be too large. This lead to high errors. Using the multiplicative variant this problem does not occur since a factor is used.

Moreover for the additive variant, a difference can be seen between the categorized and the aggregated seasonality, with the categorized seasonality scoring worse. This could be explained by data that is more scattered in a product group. With only a couple of larger products in a group and much more smaller products, the added seasonality for the larger products will be too small. This leads to high errors for those large products. Since these products sell more frequently, its weight for the weighted WAPE is larger. With a high error and a large weight, the weighted WAPE is relatively high. For aggregated seasonalities, the seasonal component is slightly better averaged. If we were to use the additive variant, the aggregated seasonalities would be better to use.

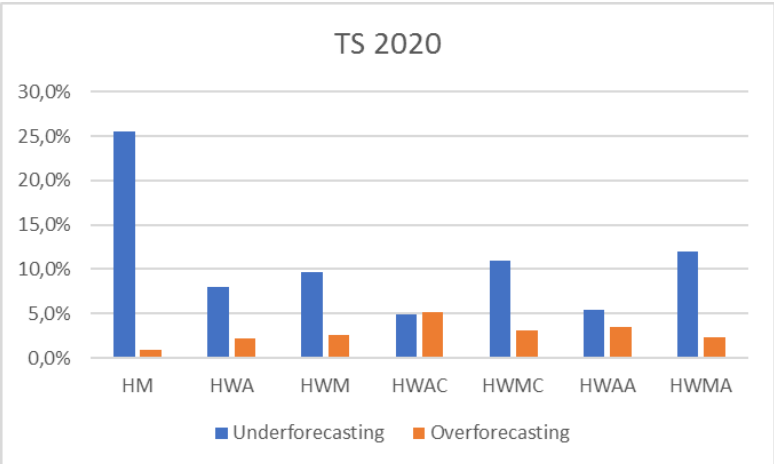
Taking the RMSE into account as well, the results in figure 14 are obtained. Again the sudden rise in demand can be seen in period 4. Contradictory to the WAPE, the error increases even further in periods 5 and 6. This is because absolute values are used instead of percentages. With higher sales numbers comes a higher RMSE. After the high error in period 6, the error again decreases because demand in those periods was lower. Also in the case of RMSE, the HWMC and HWMA score best in the beginning and at the end, the periods with a demand pattern that was more predictable.

- CONFIDENTIAL -

Figure 14 RMSE Tested Models Throughout 2020

Looking at the tracking signal, the expected underforecasting can be seen in figure 15. This underforecasting is the result of the rise in demand due to COVID-19. Only for Holt-Winter's Additive with Categorized Seasonalities, the percentage of over- and underforecasted periods is balanced. This is quite remarkable since the other Additive variants do not show this. It could be that in a normal situation, HWAC is overforecasting which makes it more balanced in case the demand rises.

Holt's model has been underforecasting in more than 25% of the periods. This can be explained by the negative trend in the last month of the training set. In 76% of the cases, the trend of a product was negative in the last period. Since this negative trend is multiplied by the horizon, the forecast will show a negative trend as well, leading to significant underforecasting. Moreover, the WAPE penalizes underforecasting less heavily which could also influence the models to underforecast more.



HM = Holt's Model
HWA = Holt-Winter's Additive
HWM = Holt-Winter's Multiplicative
HWAC = Holt-Winter's Additive with Categorized Seasonality
HWMC = Holt-Winter's Multiplicative with Categorized Seasonality
HWAA = Holt-Winter's Additive with Aggregated Seasonality
HWMA = Holt-Winter's Multiplicative with Aggregated Seasonality

Figure 15 Percentage of Periods Over- and Underforecasted in 2020

Although Holt-Winter's Multiplicative with either categorized or aggregated seasonalities do not score best based on the tracking signal, these models are still preferred because of their performance regarding the WAPE and RMSE. Especially in the periods when demand was more as expected. Since it is easier to compute the seasonalities for aggregated demand compared to categorized demand, the model with aggregated seasonalities is preferred over the model with categorized seasonalities.

4.2 Results 2019

From the previous section, it became clear that Holt-Winter's Multiplicative model with either categorized or aggregated seasonalities performs best based on 2020. However, 2020 was a remarkable year regarding the demand pattern and it is useful to test the forecasting models based on the demand of 2019 as well since this will be more realistic for the future when demand is expected to be more stable.

Again, the WAPE is considered at first in figure 16. For 2019, the WAPE behaves more as expected, since the error increases as the horizon increases. Also in this case, Holt-Winter's Multiplicative model with either categorized or aggregated seasonalities score best. Remarkable enough, the standard Holt-Winter's Multiplicative model scores much worse. This shows that the seasonal pattern cannot be captured well for individual products. This could be the case if the seasonal pattern does not repeat itself for that product in the year that has to be forecasted. This shows it is better to use aggregated seasonalities. This takes erraticness away and so the seasonal component can be better predicted.

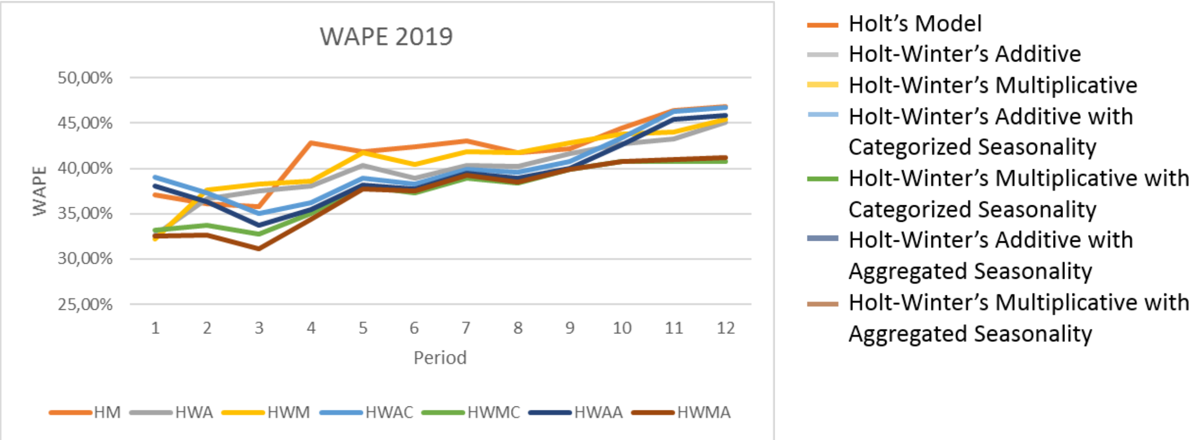
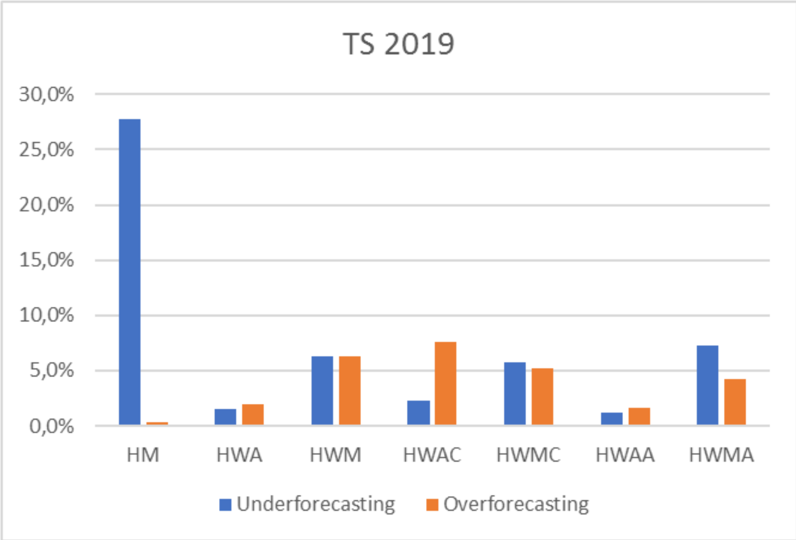


Figure 16 WAPE Tested Models Throughout 2019

Following from figure 17, it can be seen that HWMC and HWMA score best throughout 2019 compared to the other models. Also here, a rise in the error for all models can be seen in period 4. This is explained by the rise in demand because of the seasonal pattern in 2019. With higher sales numbers, the RMSE is expected to be higher as well. Also in period 7 the error increases a little, however this is because in the previous years there was a peak in the demand in period 7 leading to the model expecting the peak in period 7 again. In period 7 of 2019, the peak in demand was much smaller. This led to an increase in the error. In period 9 also a small peak in the error can be observed due to an unpredictable rise in demand. Moreover, towards the end of the year, the RMSE decreases slightly since sales numbers decrease at the end of the year, but the error is not as low as in the beginning of the year. This follows from the forecast being less accurate when the horizon is increased.

- CONFIDENTIAL -

Figure 17 RMSE Tested Models Throughout 2019



HM = Holt's Model
HWA = Holt-Winter's Additive
HWM = Holt-Winter's Multiplicative
HWAC = Holt-Winter's Additive with Categorized Seasonality
HWMC = Holt-Winter's Multiplicative with Categorized Seasonality
HWAA = Holt-Winter's Additive with Aggregated Seasonality
HWMA = Holt-Winter's Multiplicative with Aggregated Seasonality

Figure 18 Percentage of Periods Over- and Underforecasted in 2019

Considering the tracking signal in figure 18, it can again be concluded that Holt's model performs much worse than the other models. This high percentage of underforecasting can once more be explained by the negative trend in the last period of the training data which occurs in 64% of the cases. Together with the bias from the lower penalty on underforecasting from using the WAPE, Holt's model is also in 2019 extremely biased. Only Holt-Winter's Additive with Categorized Seasonalities shows significant overforecasting. A reason for this is that there are a couple of products in a group with higher sales numbers. Since the seasonality of the categorized demand is divided by the number of products, it is likely that for most products the seasonal component is too large. 21 out of the 36 seasonal components show a positive added seasonality. In the other cases, the added component is negative. This could explain the higher percentage of overforecasting compared to underforecasting. Holt-Winter's Multiplicative with Aggregated Seasonalities show slight underforecasting, which was expected because of using the WAPE. The other four models are relatively unbiased. With Holt-Winter's Additive and Holt-Winter's Additive with Aggregated Seasonalities scoring best.

Based on the forecast measures in this section, it can again be concluded that both Holt-Winter's Multiplicative with Aggregated Seasonalities and Holt-Winter's Multiplicative with Categorized Seasonalities score best. Only in the first period, Holt-Winter's Multiplicative scores slightly better. In all other periods Holt-Winter's Multiplicative with Categorized Seasonalities and Holt-Winter's Multiplicative with Aggregated Seasonalities score better than the other models. Since aggregated seasonalities are less time consuming to compute, these are chosen over categorized seasonalities.

4.3 Conclusion

It could be said that Holt-Winter's Multiplicative with Categorized Seasonalities and Holt-Winter's Multiplicative with Aggregated Seasonalities perform equally good. However, when it comes to the ease of forecasting, it can be said that Holt-Winter's Multiplicative with Aggregated Seasonalities is easier to compute than Holt-Winter's Multiplicative with Categorized Seasonalities. This is because there is no need to determine if a product belongs to the group *bags, baskets, accessories* or *other.* Also, when aggregating all demand, only one forecast with all demand needs to be made to find the seasonal indices for the products, whereas four forecasts would be needed with categorized demand. For this reason, Holt-Winter's Multiplicative with Aggregated Seasonalities is chosen as the solution that could help the team improve the accuracy of the forecast of Company X.

5 Evaluation and Implementation

Before the solution will be implemented, it is needed to know how it performs compared to the current method. This will happen in section 5.1. After that, in 5.2, a strategy will be addressed concerning the basic steps of forecasting and the implementation of the model will be elaborated on.

5.1 Evaluation

To see how Holt-Winter's Multiplicative with Aggregated Seasonalities (HWMA) compares to the current method (CM), the forecasts for 2019 and 2020 have been taken into account. To compare the forecasts, the results from section 2.1.6 can be used for the current method since these only include the products that are also in the model. For the comparison between the current method and the chosen solution with 2 and 3 years of training data, see appendix R.

At first, the weighted WAPE will be addressed. From figure 19, it can be seen that in 2019 the WAPE for Holt-Winter's Multiplicative with Aggregated Seasonalities is performing significantly better. On average, the chosen solution has a WAPE that is around 6% lower every month than the WAPE of the current method. For 2020, the WAPE of both methods are relatively the same. However, when demand was not influenced by COVID-19 yet, which is the case for the first three months, it can be seen that the model outperformed the current method. Over the course of 2020, the model had a WAPE that was only 0,5% lower on average than the current method.

Concerning the categorization as decided in section 3.2, the difference between the chosen solution and the current method regarding the number of products with a WAPE lower than 30% is very small. For 2019 the current method had 11 out of 129 products under the boundary, the model had 14 products under the boundary. For 2020, almost no products have been forecasted accurately. For the current method only 3 out of 129 products had a WAPE lower than 30%, the chosen solution had only one product more under the boundary, namely 4. Since there are many products that are not under the 30% boundary, it could be said that the demand of products is very dynamic and thus finding the right seasonal pattern can be difficult.

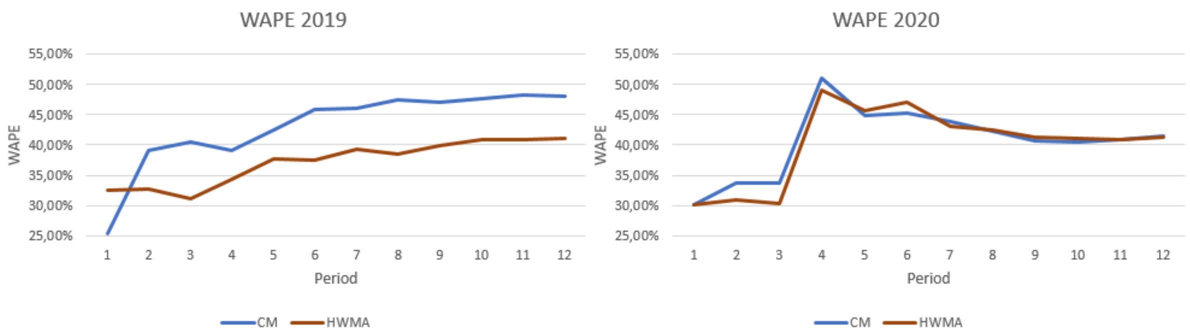


Figure 19 WAPE 2019-2020 of CM and HWMA

Considering the RMSE in figure 20, again in 2019 a significant difference can be seen where the model outperforms the current method in almost all periods, except for period one. Over the course of 2019, the model scores on average CCCC units better per month than the current method. Once more, in 2020 the errors for both methods behave relatively the same, but in the end the current method slightly outperforms the chosen solution (DDDD units less on average per month). Same as with the WAPE in 2020, in the first three months the chosen solution shows it performs better than the current method, which is promising when demand in the coming years is more stable.

- CONFIDENTIAL -

Figure 20 RMSE 2019-2020 of CM and HWMA

Also the tracking signal is reviewed. First of all, in 2019 the chosen solution is less biased than the current method. The current method is overforecasting this much because the forecast team wants to meet the budget that has been set for that particular year. A model does not take this into account and will be less biased. Normally, a biased forecast is not a good thing, but since Company X has been overforecasting in the previous years, the WAPE and RMSE for 2020 are lower than expected. This is because the amount that has originally been overforecasted for 2020 is now used for the unexpected extra demand, and thus decreasing the expected error. For 2020, both methods extremely underforecast, which was expected due to the sudden rise in demand. However, also in 2020 the chosen solution is slightly less biased than the current method.

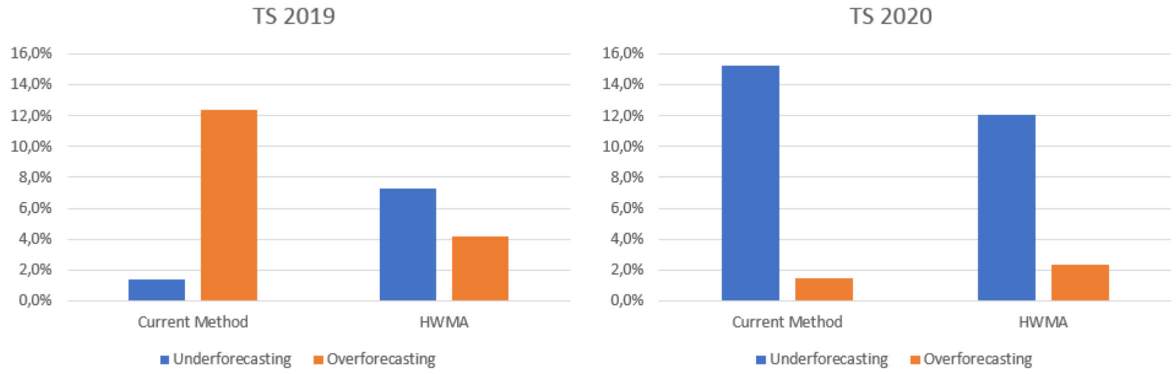


Figure 21 TS 2019-2020 of CM and HWMA

Regarding the different types of SKUs, in most cases it is that when the model scores bad, the current method scored bad as well. When the model scores good, the current method scored good as well. However *Plastic Baskets* and *Mounting Systems* have a higher error from the model (49% and 57% respectively) than from the current method (44% and 42%). This is because the seasonal patterns of these groups are slightly different than the seasonal pattern following from the aggregated demand.

From the model, the subgroups *Carriers* and *Accessories* have a relatively low error of around 30% which is mostly due to a seasonal pattern that repeats itself and corresponds to the seasonal pattern of the aggregated demand. At last, for the group *Metal Baskets* the error following from the model is much lower than from the current method, namely 15%. This is due to the error of product XXXX being 30% lower in the model which has a great influence on the weighted WAPE because of its large sales volume.

From the abovementioned forecast measures, it can be said that Holt-Winter's Multiplicative with Aggregated Seasonalities performs better than the current method. This is mainly the case for 2019, when there were no large unexpected occasions that influenced the demand heavily. This is promising for the coming years, assuming demand will not show any unexpected rises or drops like in 2020. For

2020, the chosen solution slightly outperforms the current method based on the WAPE and tracking signal, but for the RMSE the current method scores marginally better.

Furthermore, some types of SKUs could be forecasted better than others. This was the case for *Accessories* and *Carriers*. These groups have a weighted WAPE of around 30% which can be explained by a seasonal pattern that repeats itself and is corresponding to the seasonal pattern of the aggregated demand. For *Plastic Baskets* and *Mounting Systems* the error was higher, 49% and 57% respectively. This is either due to an erratic demand pattern or a seasonality that does not repeat itself in the forecasted year.

All in all, it can be said that Holt-Winter's Multiplicative with Aggregated Seasonalities is better at predicting the forecast than the current method based on the abovementioned measures.

5.2 Implementation

Now it is clear the chosen solution is outperforming the current method, a full implementation should be considered such that Company X can make use of Holt-Winter's Multiplicative with Aggregated Seasonalities in their process of making a forecast. To do so, it is useful to know the steps on how to make a forecast. For this a strategy will be described in section 5.2.1. After that, in section 5.2.2 the prototype will be elaborated on, also the adjustments that should be made by the team if the forecast is expected to be inaccurate are described.

5.2.1 Strategy

According to Hyndman and Athanasopoulos (2018) there are five basics steps in the process of making a forecast. Following these steps, it should be possible for the team to make a forecast that is as accurate as possible. Figure 22 visualizes the steps as described by Hyndman and Athanasopoulos (2018).

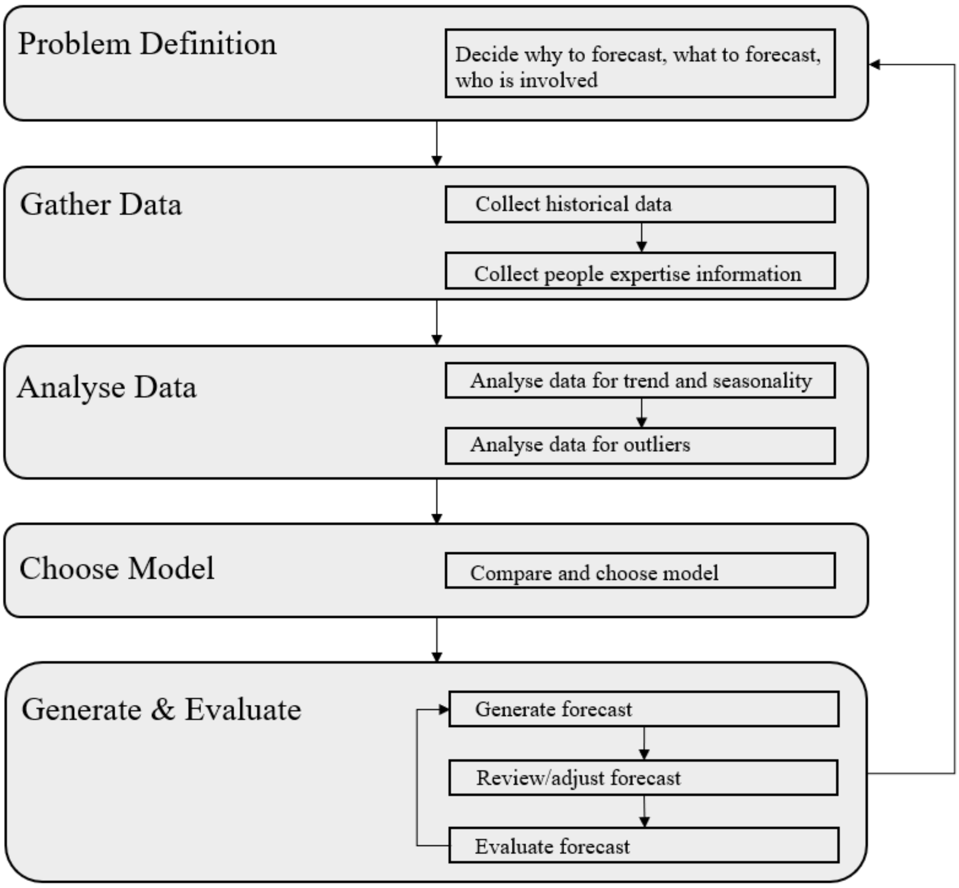


Figure 22 Process of Making a Forecast

The first four steps are conducted during the research, these steps are about preliminary research and finding the best possible model. Step five is about using and evaluating the model, this can be done after the solution has been implemented and used by Company X. The first four steps contain the following activities:

1. Problem Definition: during this step it is important to understand what the reason of forecasting is. Furthermore, what will be forecasted and who will be involved in the process of forecasting are also important questions that should be answered here.
2. Gather Data: in this step the available data should be gathered. This can consist out of two different types, statistical and expertise information.
3. Analyse Data: this step is about finding the characteristics of the data. Does the demand show trend and/or seasonality? Are there any outliers in the dataset that should be changed/deleted? These activities are important to find the best possible models.
4. Choose Model: once the characteristics of the data are known, possible models can be formulated. Once these are known, it can be tested which model performs best.

These steps already have been conducted during the research. Only step five remains to be done. This is up to Company X. Once the end of step five has been reached, a new updated forecast can be generated. If the forecast has been performing too bad, it can be advised to revise all steps and possibly find a new model.

5. Generate & Evaluate: in the last step the actual forecast will be generated. Once the forecast has been generated by the model it should be reviewed and, if needed, it should be adjusted such that the team is satisfied with the forecasted demand. After every period the accuracy of the forecast should be evaluated. If the team is not satisfied with the model anymore, step 1 to 4 should be repeated, otherwise a new forecast for the next periods can be made.

5.2.2 How to Forecast

A forecasting tool has been implemented in Excel. With this tool the employees will be able to make a statistical forecast for products with at least two years of sales data, excluding the products that are not in the scope of forecasting. The statistical forecast should always be reviewed since it is unwise to exactly follow the statistical forecast. In order to improve the accuracy of the forecast, there are some indicators and forecasting measures included in the model.

- CONFIDENTIAL -

Figure 23 Left Side of the Dashboard

What can be seen in figure 23 is the left part of the dashboard. The dashboard is the place where the most important information will be shown. For every product individually the forecast per period for the coming 12 months can be seen, together with the total number of forecasted sales in the coming 12 months which also shows the revenue that product is expected to generate. As an indicator, the number of pre-orders to the number of forecasted sales is showed. On top of the dashboard is the possibility to add an extra growth or damping to the forecast for a category. To the right of the forecast, other indicators and forecast measures are placed as shown in figure 24.

After the forecast has been made, it should first of all be checked for remarkable values. This could be a negative forecast or a forecast consisting of merely 0's. Often, the reason for this is overfitting of the training data by high smoothing constants. For products with remarkable forecasts it should be checked if the smoothing constants are high and if so, they can be changed to a lower value. According to Chopra and Meindl (2016), it is best to choose smoothing constants no larger than 0,2. This will most likely assure the model is not overfitting the training data, making the forecast more realistic.

Once remarkable values have been adjusted, an extra growth or damping could be added by filling in the percentage for the categories at "Extra Groei". This can be useful if the team expects a growth/shrinkage in sales for a category that is unlikely to be captured by the model. This way, it can be added and the forecasted values will be updated. However, this should be handled with care since the model already has some growth or shrinkage taken into account. To make it clearer, the dashboard depicts the growth or shrinkage coming from the modelled forecast as well. When changing the added growth, the total growth will show the growth from the model together with the added growth by the team. If the team expects a rise or shrinkage in demand for individual products, this could be adjusted manually as well by overwriting the forecasted value in the dashboard.

- CONFIDENTIAL -

Figure 24 Right Side of the Dashboard

To the right of the forecast, there are indicators and forecast measures. First of all, the total percentage of pre-orders compared to the total forecasted sales. If this percentage is much higher/lower than other products it might be useful to check what this reason could be. The forecast could be too low or it might be that certain products get more pre-orders leading to a higher percentage.

Next to that, the percentage of forecast that consists of pre-orders is shown per period. This indicator can show when it is likely that the forecast for a period is too low. If it can already be seen that in period 10 the outstanding orders make up 70% of the forecasted demand, this probably indicates that the forecast is too low for that period and should be adjusted accordingly.

Next to the percentage of pre-orders, the WAPE of every product over the last 12 periods can be found. If this is below the boundary of 30% it can be said the model has accurately captured the demand pattern and not too much attention should be given to these products. However, it will always be needed to check the other forecast measures and indicators to accurately predict future demand. Next to the WAPE of the model, also the WAPE of the adjusted forecast is calculated. With this, it can be seen if the input from the forecast team has had a positive effect on the accuracy.

Next to the WAPE the tracking signal is a useful forecast measure. This indicates whether the forecast that is made has been over- or underforecasting the demand. If it can be seen that the tracking signal exceeds the limits of ±6 it should be considered to adjust the forecast to be either higher or lower. Again, the tracking signal after input from the forecast team is shown as well.

At last the root mean square error is included. This will indicate how far off the forecast has been on the same scale as the product. Since it is scale dependent it cannot be compared to the other products which makes it more difficult to use. However, it can be said that the lower the RMSE is, the better. Also here, calculations based on the adjusted forecasts will be done to see how beneficial the input from the team is.

For a better accuracy, *customer forecasts* can be used as described in section 3.3. The main purpose of these forecasts is to detect rises or drops in demand from a customer that are significant enough to conclude that more/less products need to be ordered. Moreover, if this happens it is beneficial to discuss with the customer what the reason for the rise/drop in demand is. According to Company Z, by talking to their customers, the relation between Company Z and their customers improved. Moreover, valuable information can be exchanged about competitors. For example, a competitor that cannot deliver and the reason for that. During the meetings with customers, it might also become clear what the intention of the customer is. It could be that the rise in demand will happen only once, or that the customer will keep ordering more in other periods as well. Something that is currently unknown to Company X.

At last, it could be said that when demand during lead time is higher than expected, it should be delivered from stock if possible. If the deviation in demand is too large and it cannot be delivered from stock, it will be placed in backorder to be delivered in the coming periods. Since the lead time is around four months, it is beneficial to have a forecast that is not only accurate for the coming month, but also for the next 4+ months.

6 Conclusion and Recommendations

The last chapter will cover the conclusions, discussion, recommendations and recommended further research. First, in section 6.1 the main findings will be evaluated. In section 6.2, limitations will be discussed. Section 6.3 will cover recommendations following this research. At last, section 6.4 will suggest on what could be further researched.

6.1 Conclusion

This research aimed to find a possible forecast model that would help Company X in the process of making a forecast, such that the forecast is more accurate. The modelled forecast has to be made with a horizon of one year with time buckets of one month. Based on the tested models and evaluation, it can be concluded that Holt-Winter's Multiplicative model with Aggregated Seasonalities performed best on the data of Company X. Based on solely modelled demand values, the error for the 2019 test set was around 6% lower on average each month than the current method. With the right adjustments of the team, the error will be even lower. Also, the model has a much less biased forecast. Together with the customer forecasts and indicators in the forecast model, Company X will be able to adjust the forecast in the correct way such that the forecast accuracy improves as well as its relation with the customers.

In order to come to this result, data has been collected and validated to check for possible outliers. Based on a visual check, outliers have been found and changed accordingly. With the remaining data, the current accuracy could be calculated. The current method showed a weighted WAPE of 48% in 2019 and 41% in 2020. Also, clear overforecasting can be seen in the year 2019, which can be explained by the forecast team wanting to meet a certain budget. It is likely that this happened in the previous years as well. For 2020, as expected, there has been a lot of underforecasting since the demand increased a lot.

Following from data analysis, it has been concluded a significant trend and seasonality can be detected in the data. These are two very important characteristics when choosing possible models. Moreover, the possible solutions should meet the criteria that have been set by the team; forecast with a one year horizon, forecast with time buckets of one month, model should not be too technical and programming is not preferred.

Based on the data characteristics and criteria from the team, three models could be a possible solution. Namely: *Holt's Model, Holt-Winter's Additive Model* and *Holt-Winter's Multiplicative Model.* For Holt-Winter's Models the seasonality component takes on different forms, namely seasonality per product, seasonality per product group (Categorised Seasonality) and seasonality with all demand aggregated (Aggregated Seasonality). From the tests, Holt-Winter's Multiplicative model with Aggregated Seasonalities performed best with a weighted WAPE of 41% in both 2019 and 2020.

Based on the evaluation of the chosen solution and the current method, it has been concluded that the chosen solution outperforms the current method. The accuracy of the chosen solution is solely based on a modelled forecast, so adjusting the forecast where necessary will help increase the accuracy even further. Since the model outperforms the current method, it has been implemented in an Excel model. This model not only shows the forecasted values, but also some useful indicators and forecast measures that can be used to improve the accuracy.

To increase the accuracy of the forecast, it is advised to also make use of the customer forecasts to detect sudden rises or drops in demand. Once this is seen, it can be discussed with the customer why this happens and what will happen after these periods. When this is known, the team could act accordingly.

From the abovementioned findings and the forecasting strategy as proposed by Hyndman and Athanasopoulos (2018), Company X should be able to increase the accuracy of their forecast in the future.

6.2 Discussion

Also this research has its limitations that could affect the validity. These limitations will be discussed in the following section.

- The COVID-19 pandemic has played a huge role in the demand patterns of Company X. The accuracy of the forecast made by the model might be lower in the following periods which is due to the unusual demand pattern in 2020. Assuming that demand patterns will behave as it did before COVID-19, it will take around three years before the remarkable fluctuations due to COVID-19 are fully excluded from the data.
- An improvement can be made during the process of optimizing smoothing constants. In this research, the WAPE has been used since the managers are the most comfortable with this. However, the WAPE has the disadvantage to penalize underforecasting less heavily than overforecasting. This leads to a slightly biased forecast with more underforecasting than overforecasting. In order to eliminate this bias, it would be better to use another measure like the Mean Square Error or the Mean Absolute Deviation. These errors penalized under- and overforecasting the same.
- During the selection of usable data to ensure a valid research, a visual check has been conducted in order to find any outliers in the monthly sales of Company X. Using this visual check it could be that some outliers are being skipped, so this might not be the best option for outlier detection. To conduct a better outlier detection, a more sophisticated statistical method could be used.

6.3 Recommendations

Following from this research, there are multiple recommendations that can be proposed. These recommendations could be useful in order to generate a forecast that would be more accurate.

- Make sure that products that generate most revenue have an excellent forecast accuracy. Following from the model, products with high revenue already have a better accuracy than the products with lower revenue since their contribution to the seasonal patterns of the aggregated demand is relatively large. The products with high revenue and good accuracy should be NOS products since it is relatively certain they will be sold, either now or next year. However, there are some products that generate much revenue and are sold often, but its demand pattern is uncertain and so is its accuracy. For these products it would be good to review the forecast better.
- Track the accuracy of the input of the forecast team. Since it became clear from the current situation analysis that the accuracy did not necessarily improve once the forecast has been adjusted, it will be a good thing to evaluate what the influence of the input from the team is.
- Use more years of data. If more years of data are used, it will be easier for the model to capture the seasonal patterns as well as the trend. With three or only two years, it might be that one year with a remarkable seasonal pattern might be of big influence on the forecast.
- Decrease the lead time of suppliers. This might not be an easy task, but doing so the forecast horizon could be decreased. With a smaller forecast horizon, the accuracy increases since it is not needed to forecast that far in the future. With a better forecast accuracy, the delivery reliability and inventory costs can be optimized as well.

6.4 Further Research

In order to improve the process of forecasting more, some suggestions can be made for further research. These suggestions are listed below.

- Implement the forecast in an automated program. With the chosen solution in an Excel model it is needed to manually insert the new data and update some values to do all calculations. If the model would be implemented in an automated system, it would be more efficient and less error prone. There are already systems that can make a forecast automatically.

- Method for forecasting new products. Since the model does need at least two years of data, it is not possible to make a statistical forecast for products that exist less than two years. However, some models exist for predicting demand for new products, but these are often very complex since they take multiple independent variables into account. Also, they might not predict future demand accurately. This is due to the lack of data and the fact that it is very complicated to forecast how well a new product will be accepted by the consumers. There are also qualitative methods for this like the Delphi method.
- Machine learning models. Although these models are difficult to understand and implement, it could be a good idea to further investigate possible models that use machine learning models like Neural Networks, which were earlier addressed in section 3.1.6. It could be that these models outperform the classical models. However in multiple researches the classical model outperformed machine learning models. To see which model would perform best based on Company X' data, this could be further researched.
- Analyse the optimal amount of inventory. An often-heard point of discussion at Company X is the amount of inventory. Some employees say the current level is fine, others say it is way too low. Since the sales for Company X are hard to predict, it might be better to have more inventory to be prepared for unexpected increases in demand. Since the delivery reliability has not been great in the previous years, it would be interesting to analyse what the correct amount of inventory is for Company X, keeping the difficulty of predicting future demand in mind.

References

Baldigara, T., & Mamula, M. (2015). Modelling international tourism demand using seasonal ARIMA models. *Tourism and Hospitality Management*, *21*(1), 19–31.

Booranawong, T., & Booranawong, A. (2017). SIMPLE AND DOUBLE EXPONENTIAL SMOOTHING METHODS WITH DESIGNED INPUT DATA FOR FORECASTING A SEASONAL TIME SERIES: IN AN APPLICATION FOR LIME PRICES IN THAILAND. *Suranaree Journal of Science & Technology*, *24*(3).

Boylan, J. E., Syntetos, A. A., & Karakostas, G. C. (2008). Classification for Forecasting and Stock Control: A Case Study. *The Journal of the Operational Research Society*, *59*(4), 473–481. http://www.jstor.org.ezproxy2.utwente.nl/stable/30133025

Brillio. (2018). *Choosing The Right Forecasting Technique*. https://www.brillio.com/insights/choosing-the-right-forecasting-technique/#:~:text=Causal forecasting is the technique,can impact the dependent variable.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?– Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250.

Chopra, S., & Meindl, P. (2016). *Supply Chain Management* (6th ed.). Pearson.

Dekker, M., van Donselaar, K., & Ouwehand, P. (2004). How to use aggregation and combined forecasting to improve seasonal demand forecasts. *International Journal of Production Economics*, *90*(2), 151–167. https://doi.org/https://doi.org/10.1016/j.ijpe.2004.02.004

Heerkens, H., & Winden, A. Van. (2017). Solving Managerial Problems Systematically. In *Solving Managerial Problems Systematically*. Noordhoff Uitgevers.

Hyndman, R. J. (2011). *Forecasting: An Overview.*

Hyndman, R. J. (2014). Measuring forecast accuracy. *Business Forecasting: Practical Problems and Solutions*, 177–183.

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

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

Hyndman, R. J., & Kostenko, A. V. (2007). Minimum sample size requirements for seasonal forecasting models. *Foresight*, *6*(Spring), 12–15.

Kalekar, P. S. (2004). Time series forecasting using holt-winters exponential smoothing. *Kanwal Rekhi School of Information Technology*, *4329008*(13), 1–13.

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, *32*(3), 669–679.

Klimberg, R., Sillup, G., Boyle, K., & Tavva, V. (2010). Forecasting performance measures - What are their practical meaning? *Advances in Business and Management Forecasting*, *7*, 137–147. https://doi.org/10.1108/S1477-4070(2010)0000007012

Köppelová, J., & Jindrová, A. (2019). Application of exponential smoothing models and arima models in time series analysis from telco area. *AGRIS On-Line Papers in Economics and Informatics*, *11*(665-2019–4145), 73–84.

Law, R., & Au, N. (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. *Tourism Management*, *20*(1), 89–97.

Makatjane, K., & Moroke, N. (2016). Comparative study of holt-winters triple exponential smoothing and seasonal Arima: forecasting short term seasonal car sales in South Africa. *Makatjane KD, Moroke ND*.

Nenes, G., Panagiotidou, S., & Tagaras, G. (2010). Inventory management of multiple items with irregular demand: A case study. *European Journal of Operational Research*, *205*(2), 313–324.

Ostertagová, E., & Ostertag, O. (2011). The simple exponential smoothing model. *The 4th International Conference on Modelling of Mechanical and Mechatronic Systems, Technical University of Košice, Slovak Republic, Proceedings of Conference*, 380–384.

Shrestha, M. B., & Bhatta, G. R. (2018). Selecting appropriate methodological framework for time series data analysis. *The Journal of Finance and Data Science*, *4*(2), 71–89. https://doi.org/https://doi.org/10.1016/j.jfds.2017.11.001

Syafei, A. D., Ramadhan, N., Hermana, J., Slamet, A., Boedisantoso, R., & Assomadi, A. F. (2018). Application of Exponential Smoothing Holt Winter and ARIMA Models for Predicting Air Pollutant Concentrations. *EnvironmentAsia*, *11*(3).

Wang, W., & Lu, Y. (2018). Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model. *IOP Conference Series: Materials Science and Engineering*, *324*(1), 12049.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82.

Yunishafira, A. (2018). Determining the Appropriate Demand Forecasting Using Time Series Method: Study Case at Garment Industry in Indonesia. *KnE Social Sciences*, 553–564.

Appendices

Appendix A: Root Mean Square Error

The Root Mean Square Error (RMSE) has been taken into account as well. The RMSE is an often-used forecast measure, also because it can deal with zero-demand periods. As stated by Wang and Lu (2018): "For evaluation, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are widely adopted in many recommendation systems to measure the difference between the predicted scores and users' actual ratings." However, there has been a discussion going on about which error is better, the RMSE or the MAE. In a research where the MAE and RMSE are compared, Willmott and Matsuura (2005) state the following: "it seems to us that there is no clear interpretation of RMSE or related errors, and we recommend that such measures no longer be reported in the literature." Willmott and Matsuura concluded that the MAE is superior to the RMSE. However, according to Chai and Draxler (2014), under the circumstances of calculating model error sensitivities, MAEs are definitely not preferred over RMSEs.

The RMSE should be interpreted as the standard deviation of the residuals, also known as errors. To calculate the RMSE the following formula will be used, where the number of periods is 12:

$$RMSE_n = \sqrt{\frac{1}{n} * \sum_{t=1}^{n} (D_t - F_t)^2} \tag{A.1}$$

Also here, to get an overall value of the RMSE, the RMSEs of the products are multiplied by a weight since products with a higher number of sales should be more important. This is done the same way as for the MAPE. Table 6 shows the results of the weighted RMSEs of the forecasts. Table 7 shows the results when demand is aggregated. In case of aggregated demand, the adjusted forecast for 2020 scores better than the original.

Table 6, Root Mean Square Error Current Situation

- CONFIDENTIAL -

Table 7, Aggregated Root Mean Square Error Current Situation

- CONFIDENTIAL -

The values in the table show that the adjusted forecast scores worse on the RMSE compared to the original forecast since its value is higher. The higher the value is, the higher the standard deviation of the residuals is and thus the worse the forecast fits the real demand.

Appendix B: Mean Absolute Deviation

The MAD might be the most widely used forecast measures. This can be explained by the fact that this measure is easy to interpret and compute. The MAD also works on all values, so zero values are no problem. To calculate the MAD for a product, all the absolute errors in a period will be summed, after that they are divided by the number of periods, which is 12 in this case. After the MAD is calculated for the products individually, it will again be weighted to get an overall value. The formula of the Mean Absolute Error is given by

$$MAD_n = \frac{1}{n} * \sum_{t=1}^{n} |D_t - F_t| \tag{B.1}$$

Table 8 depicts the results from the calculations of the weighted MAD. Table 9 depicts the results when demand is aggregated.

Table 8, Mean Absolute Deviation Current Situation

- CONFIDENTIAL -

Table 9, Aggregated Mean Absolute Deviation Current Situation

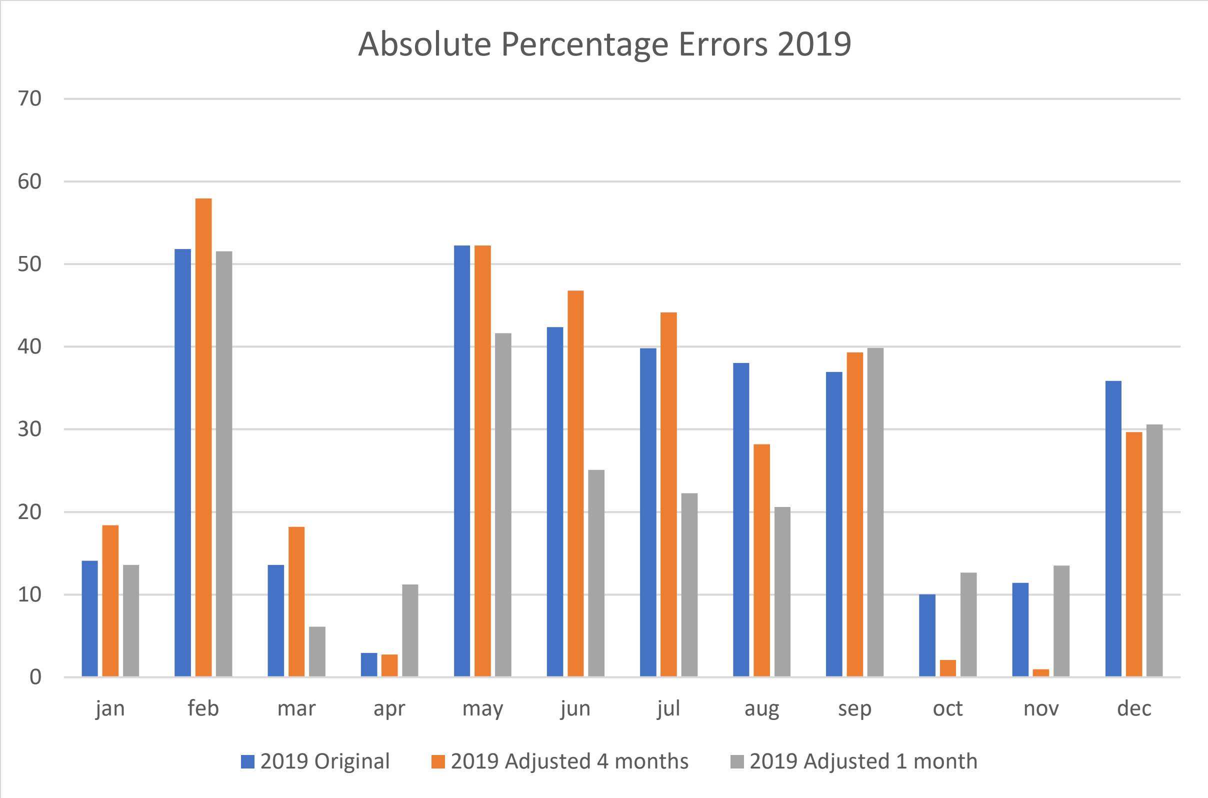
- CONFIDENTIAL -

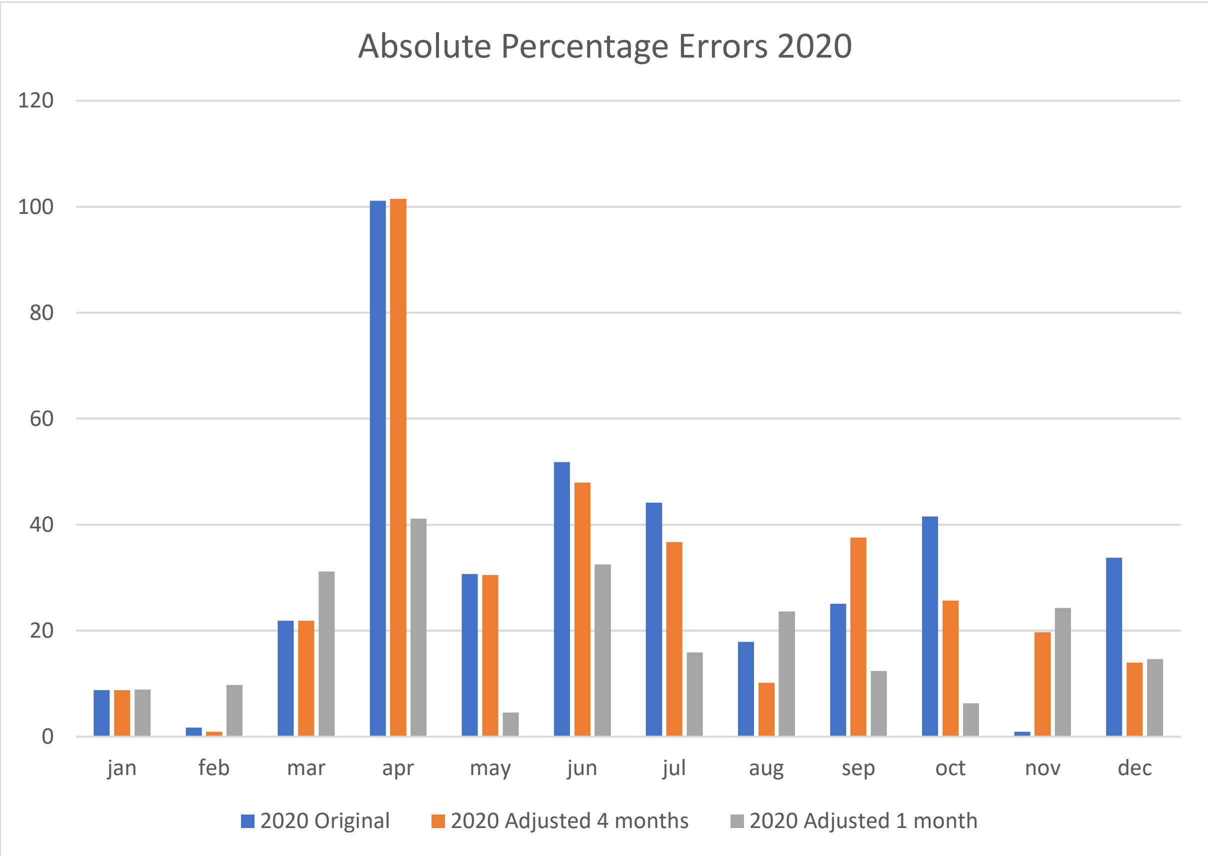
Appendix C: Mean Absolute Percentage Errors

Weighted MAPE	*2019*	*2020*
Original Forecast	80,50%	58,75%
Adjusted FC 4 months prior	81,27%	61,89%
Adjusted FC 1 month prior	74,09%	55,02%

Aggregated MAPE	*2019*	*2020*
Original Forecast	29,10%	31,62%
Adjusted FC 4 months prior	28,40%	29,62%
Adjusted FC 1 month prior	24,06%	18,78%

Appendix D: Absolute Percentage Errors Per Month





Appendix E: Weighted Absolute Percentage Error

Weighted WAPE	*2019*	*2020*
Original Forecast	51,30%	46,92%
Adjusted FC 4 months prior	52,63%	45,67%
Adjusted FC 1 month prior	42,07%	40,32%

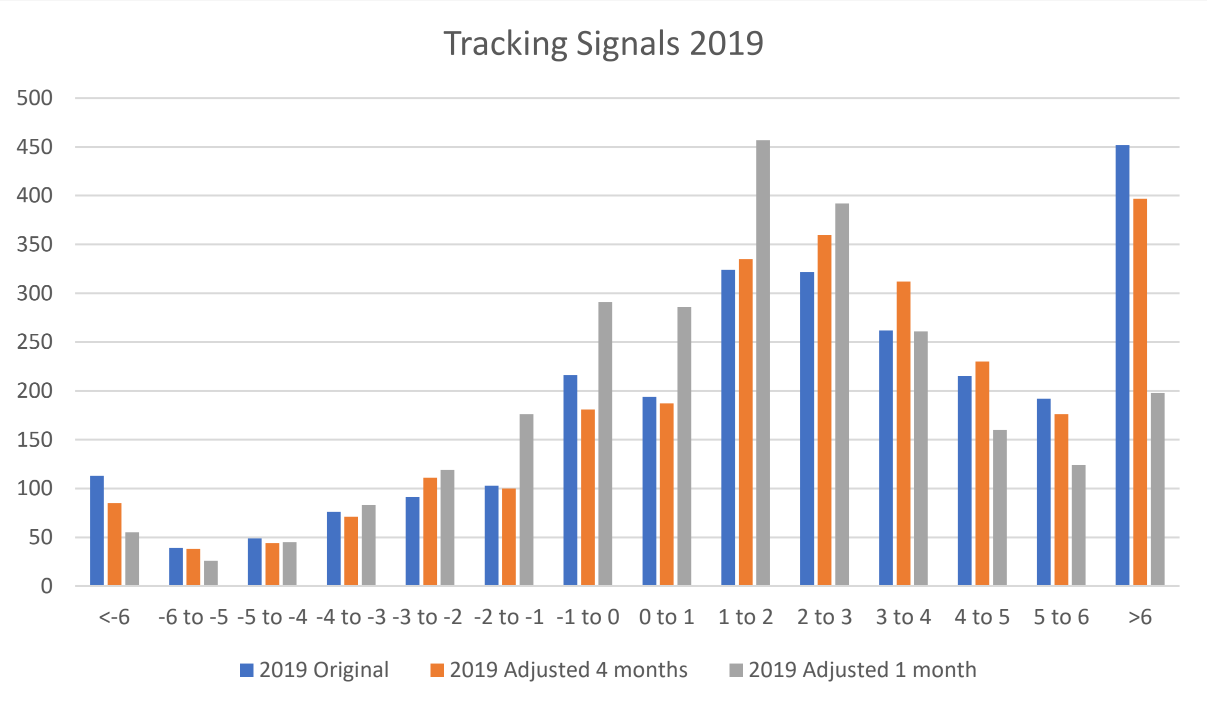
Aggregated MAPE	*2019*	*2020*
Original Forecast	28,78%	32,46%
Adjusted FC 4 months prior	29,01%	30,31%
Adjusted FC 1 month prior	23,81%	18,24%

Appendix F: Tracking Signals Table

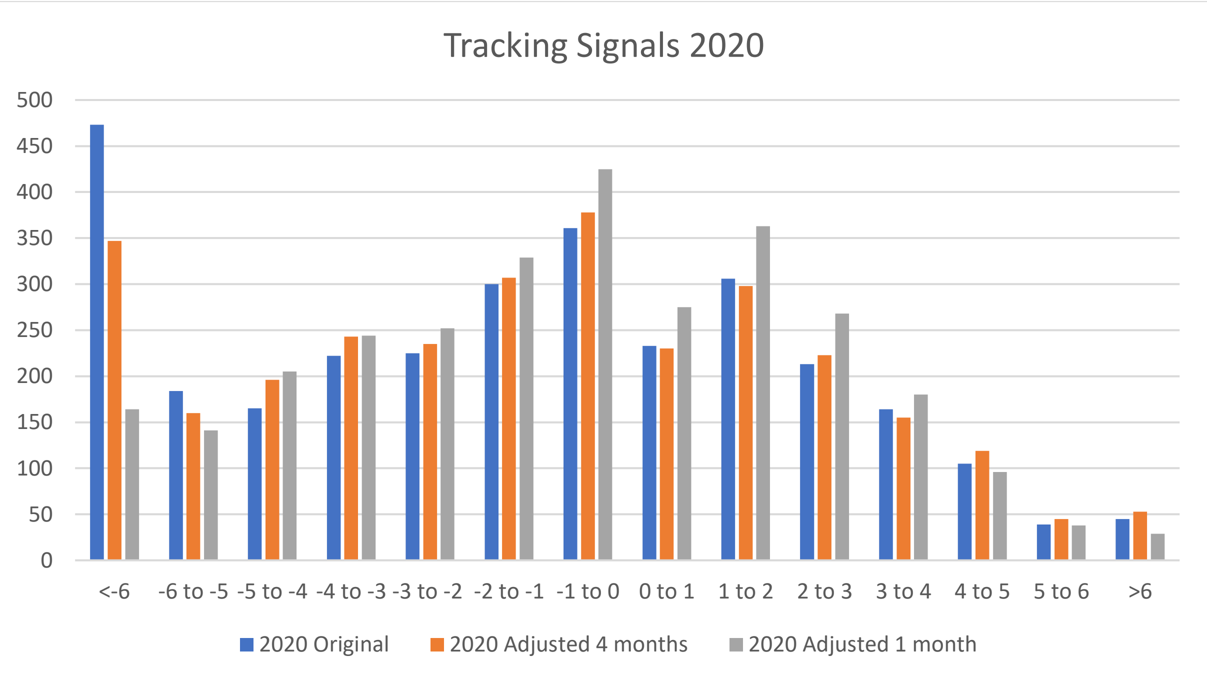
Tracking Signal	*2019 (2652 Periods Total)*	*2020 (2988 Periods Total)*
Original Forecast		
Overforecasting	452 (17,0%)	45 (1,5%)
Underforecasting	113 (4,3%)	473 (15,8%)
Adjusted FC 4 months		
Overforecasting	397 (15,0%)	53 (1,8%)
Underforecasting	89 (3%)	365 (12,2%)
Adjusted FC 1 month		
Overforecasting	198 (7,5%)	29 (1%)
Underforecasting	55 (2,1%)	169 (5,7%)

Appendix G: Tracking Signals Graph

Tracking Signals in 2019, divided in buckets of size 1.



Tracking Signals in 2020, divided in buckets of size 1.



Appendix H: Performance per SKU Based on WAPE

2019, Original Forecast

- CONFIDENTIAL -

2019, 4 Months Adjusted Forecast

- CONFIDENTIAL -

2019, 1 Month Adjusted Forecast

- CONFIDENTIAL -

2020, Original Forecast

- CONFIDENTIAL -

2020, 4 Months Adjusted Forecast

- CONFIDENTIAL -

2020, 1 Month Adjusted Forecast

- CONFIDENTIAL -

Appendix I: Performance per SKU Based on TS

2019, Original Forecast

- CONFIDENTIAL -

2019, 4 Months Adjusted Forecast

- CONFIDENTIAL -

2019, 1 Month Adjusted Forecast

- CONFIDENTIAL -

2020, Original Forecast

- CONFIDENTIAL -

2020, 4 Months Adjusted Forecast

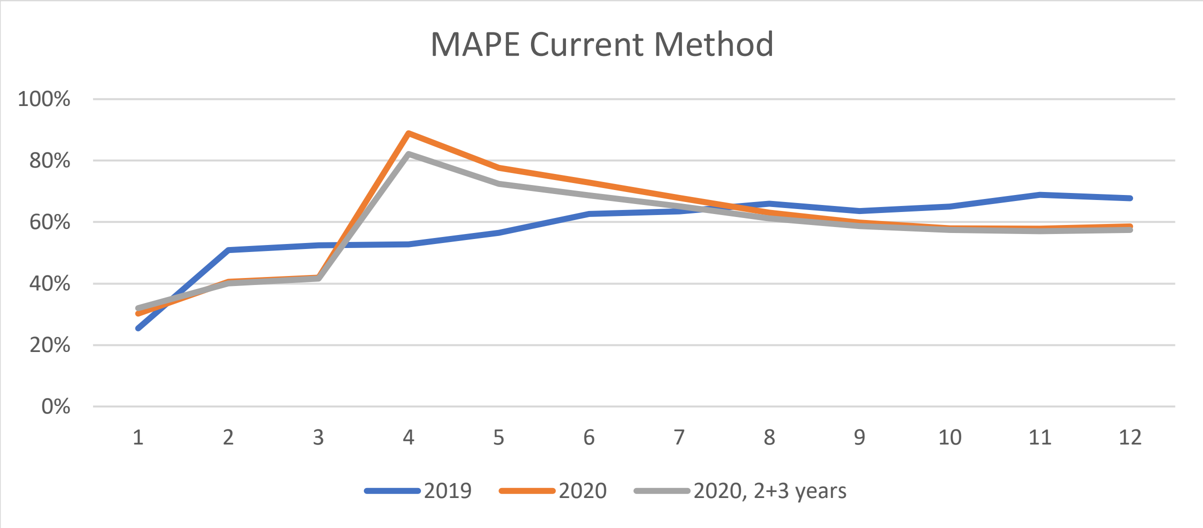
- CONFIDENTIAL -

2020, 1 Month Adjusted Forecast

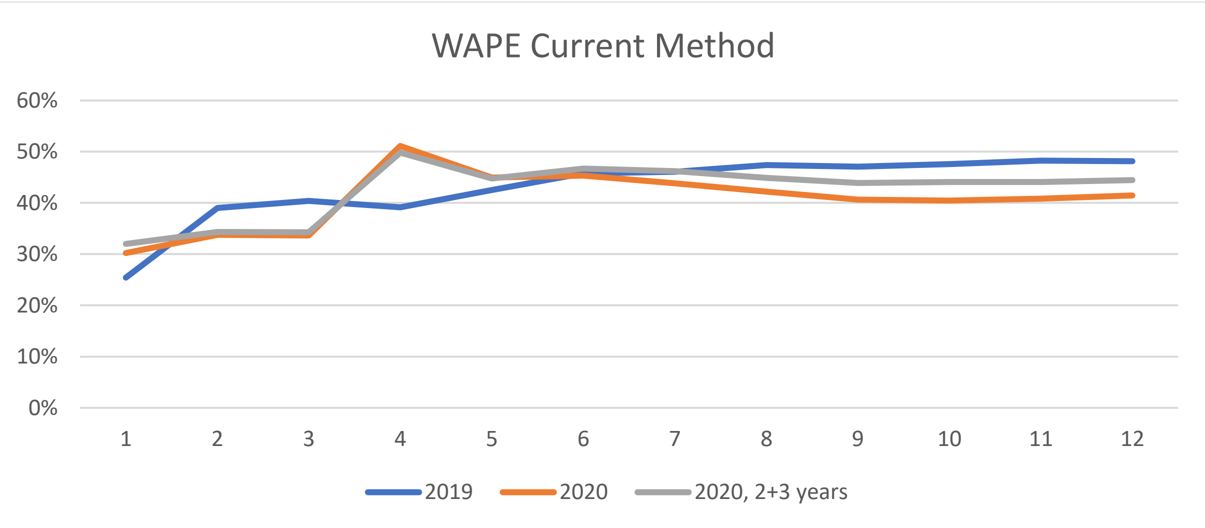
- CONFIDENTIAL -

Appendix J: Current Forecast Accuracy

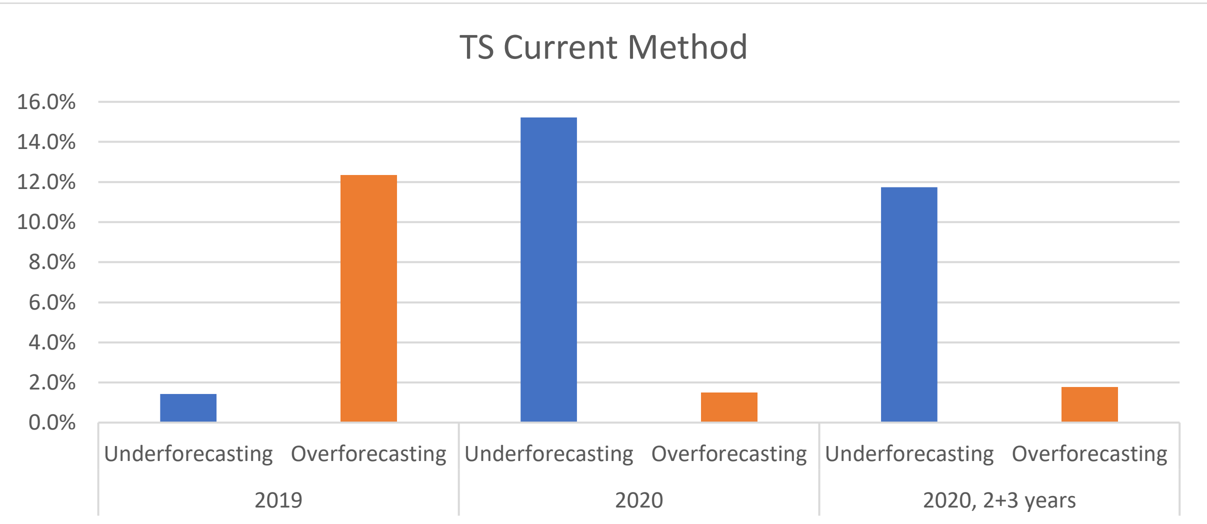
MAPE



WAPE



Tracking Signal



MAD

- CONFIDENTIAL -

RMSE

- CONFIDENTIAL -

Appendix K: Forecast Measure Overview

2019, original forecast.

- CONFIDENTIAL -

2019, adjusted forecast 4 months prior to delivery.

- CONFIDENTIAL -

2019, adjusted forecast 1 month prior to delivery.

- CONFIDENTIAL -

2020, original forecast.

- CONFIDENTIAL -

2020, adjusted forecast 4 months prior to delivery.

- CONFIDENTIAL -

2020, adjusted forecast 1 month prior to delivery.

- CONFIDENTIAL -

Appendix L: Linear Regression Aggregated Demand

- CONFIDENTIAL -

Appendix M: Demand Patterns

According to Boylan, Syntetos and Karakostas (2008), there are four demand patterns, namely: *erratic, lumpy, smooth* and *intermittent.* To find the best model, it should be known which demand pattern can be determined from the dataset of Company X. This is done by checking the Average inter-Demand Interval (ADI) over a year with periods of one month and the squared coefficient of variation of demand sizes (CV^2). Following from the scheme in figure 25, it can be determined what demand pattern the data has.

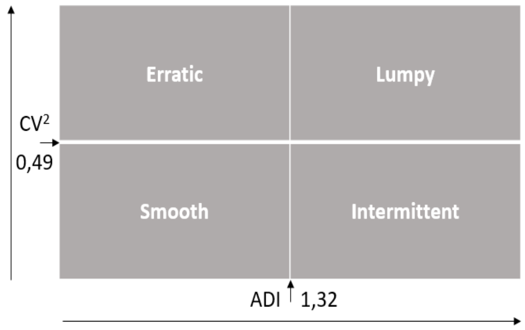


Figure 25 Demand Patterns

For the demand in 2019, it can be seen that 77% of the products have a CV^2 lower than 0,49, also 99% show an ADI lower than 1,32, meaning that most of the products show a smooth demand pattern. From the calculations with demand from 2020, it becomes clear that 73% of the data has a CV^2 that is lower than 0,49. Moreover, 100% of the demand has an ADI that is lower than 1,32. These numbers also show that demand often falls in the *smooth* category.

As stated by Nenes, Panagiotidou and Tagaras (2010): "slow-moving items have intermittent demand with each demand size equal to one item or very few items." Since almost no products show intermittent demand, it can also be concluded that the items are not slow-movers.

All in all, this shows that most of the products are in the *smooth* category, with some products being in the *erratic* category. Moreover, it cannot be said these products are slow-movers.

Since around 75% is in the *smooth* category, this will be used to find a solution.

Appendix N: Model Details Holt(-Winter's)

Holt's model
$F_{t+h} = L_t + (h)T_t$
$L_t = \alpha * D_t + (1-\alpha) * (L_{t-1} + T_{t-1})$
$T_t = \beta * (L_t - L_{t-1}) + (1-\beta) * T_{t-1}$
$a, b = smoothing\ constants, t = period, h = horizon$

Holt-Winter's additive model
$F_{t+h} = L_t + (h)T_t + S_{t+h}$
$L_t = \alpha * (D_t - S_t) + (1-\alpha) * (L_{t-1} + T_{t-1})$
$T_t = \beta * (L_t - L_{t-1}) + (1-\beta) * T_{t-1}$
$S_t = \gamma * \left(D_{t-p} - L_{t-p-1} - T_{t-p-1}\right) + (1-\gamma) * S_{t-p}$
$\alpha, \beta, \gamma = smoothing\ constants, t = period, h = horizon, p = periodicity$

Holt-Winter's multiplicative model
$F_{t+h} = (L_t + (h)T_t) * S_{t+h}$
$L_t = \alpha * \left(\frac{D_t}{S_t}\right) + (1-\alpha) * (L_{t-1} + T_{t-1})$
$T_t = \beta * (L_t - L_{t-1}) + (1-\beta) * T_{t-1}$
$S_t = \gamma * \left(\frac{D_{t-p}}{L_{t-p}}\right) + (1-\gamma) * S_{t-p}$
$\alpha, \beta, \gamma = smoothing\ constants, t = period, h = horizon, p = periodicity$

Appendix O: SARIMA Model

A stated by Baldigara and Mamula (2015), the SARIMA model is given by

$$\Phi(B^S) * \varphi(B)(1 - B^S)^D(1 - B)^d * y * t = \Theta(B^S) * \theta(B)\varepsilon_t \tag{O.1}$$

$Where,$

- $\phi(B) = p - order\ non - seasonal\ Auto\ Regressive\ model$
- $\Phi(B) = P - order\ seasonal\ Auto\ Rgressive\ model$
- $\theta(B) = q - order\ non - seasonal\ Moving\ Average\ model$
- $\Theta(B) = Q - order\ seasonal\ Moving\ Average\ model$
- $(1 - B)^d = d^{th}\ non - seasonal\ difference$
- $(1 - B^s)^D = D^{th}\ seasonal\ difference\ of\ seasons\ s$
- $\varepsilon_t = error\ term\ {\sim}IID(0, \sigma^2)$
- $B = backshift\ operator$
- $S = seasonal\ order$

Appendix P: Neural Network Model

A simple neural network is depicted in figure 26. This neural network has 4 input nodes.

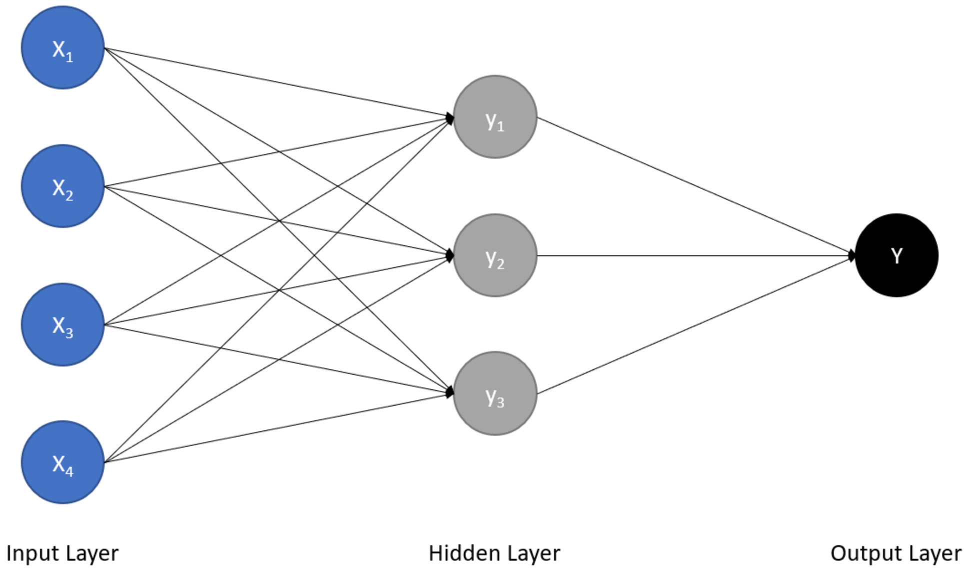


Figure 26 Neural Network

Each of these layers contains nodes and these nodes are connected to nodes in adjacent layers. With 4 input nodes, the nodes in the hidden layer compute y_j in the following way:

$$y_j = \sum_{i=1}^{4} x_i \, w_{ji} \tag{P.1}$$

$$where\ x_i = input\ variable\ and\ w_{ji} = weight\ from\ node\ j\ to\ i$$

To transform the output such that the values fall in an acceptable range, the sigmoid function is used. This transformation is done before the value reaches the next layer. This function makes sure the output value of the hidden layer falls between 0 and 1. The sigmoid function is given by

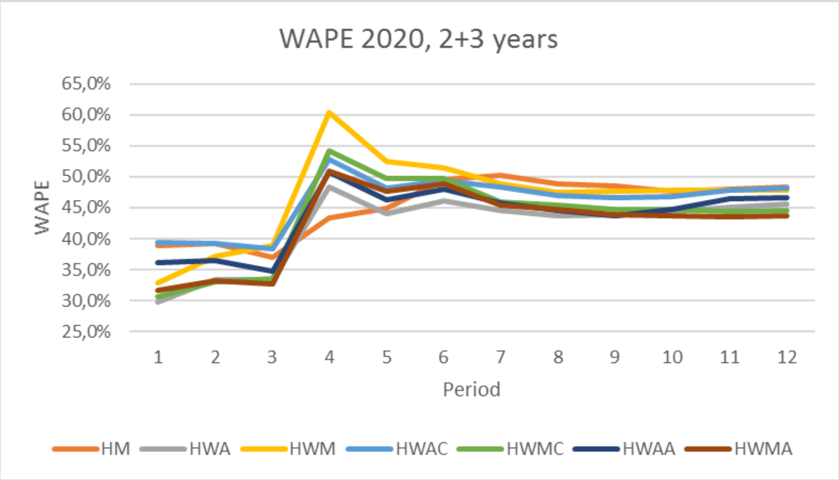
$$y_T = \frac{1}{1 + e^{-y}} \tag{P.2}$$

In the end, the value of Y in the output node is given by

$$Y = \sum_{i=1}^{3} y_{Ti} w_i \tag{P.3}$$

Appendix Q: Results with 2+3 Years Training Data

WAPE

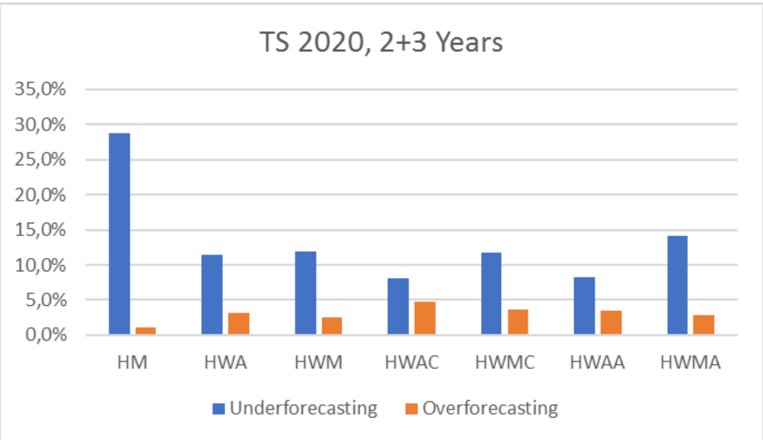


Holt's Model
Holt-Winter's Additive
Holt-Winter's Multiplicative
Holt-Winter's Additive with Categorized Seasonality
Holt-Winter's Multiplicative with Categorized Seasonality
Holt-Winter's Additive with Aggregated Seasonality
Holt-Winter's Multiplicative with Aggregated Seasonality

RMSE

- CONFIDENTIAL -

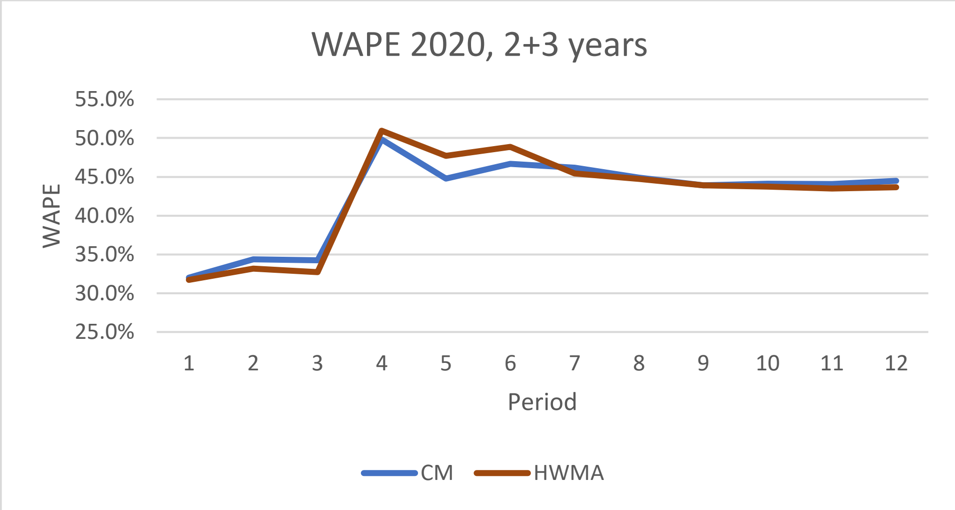
Tracking Signal



HM	= Holt's Model
HWA	= Holt-Winter's Additive
HWM	= Holt-Winter's Multiplicative
HWAC	= Holt-Winter's Additive with Categorized Seasonality
HWMC	= Holt-Winter's Multiplicative with Categorized Seasonality
HWAA	= Holt-Winter's Additive with Aggregated Seasonality
HWMA	= Holt-Winter's Multiplicative with Aggregated Seasonality

Appendix R: Comparison Current Method and Solution 2+3 Years

WAPE



RMSE

- CONFIDENTIAL -

Tracking Signal