# Forecasting food waste: The case of a Swedish grocery company

**A quantitative & exploratory study of how to forecast food waste for a Swedish retailer**

**Authors:**

Joel Klefbom Ridderstråle

William Lundström

**Examinator:**

Wiley Wakeman

**Supervisor:**

Mariya Ivanova

**Submission:**

May 19th 2020

## Acknowledgments

We would like to thank our supervisor Mariya Ivanova for guiding and coaching us in writing our first thesis. Especially valuable were all the pieces of advice that provided us with a framework for how quantitative research should be handled. We would also like to thank Professor Nikolaos Kourentzes for introducing us to the forecasting discipline and for taking the time to help us with structuring the methodology. We are most grateful that he consistently gave us recommendations for how to proceed when crossroads appeared. Finally, we would like to thank Tobias Petersson, Mathias Jönsson, Magnus Palm, and Stellan Svensson along with all the people at Axfood. They provided a lot of guidance and knowledge surrounding food retailing and waste management. They also gave us access to their data banks and supplied company PCs on which all forecasting took place. For all this, we are very grateful.

Abstract

In the retail stage of a food supply chain, the occurrence of food waste is often caused by inaccurate forecasting of sales, which leads to incorrect ordering of products. Previous literature has suggested using more accurate demand forecasts on the sales data in order to combat the problem. However, a few authors have suggested applying forecasting methods to the waste data in order to reduce the food waste. This thesis explores the potential predictive power that forecasting methods possess in explaining food waste by comparing more advanced methods to the simplest form of forecasting, namely the *Naïve* forecast. Comparisons are made between a *Stepwise Regression* method, with a set of explanatory variables, lags and manually constructed variables, *Exponential Smoothing* (*ES*) methods and a combination model of the *Stepwise Regression* method and the *ES* methods, were used in order to assess which of these are the most accurate in terms of predicting food waste data at different aggregated levels. By using a set of different error metrics, more nuanced conclusions can be drawn regarding which aspects of the food waste data is explained by the models. The variables from the *Regression* model can further describe if and what factors actually explain food waste and if these differ for aggregated levels of the data. At the highest aggregated level, the *Combination* model has the most accurate forecast. At lower aggregated levels the results show that the *ES* has the most accurate predictive power in terms of explaining seasonal structure, whereas the *Stepwise Regression* model is, on all aggregated levels, most proficient in terms of explaining outliers. Some of the *Stepwise Regression* model's explanatory variables display consistent, significant correlation to food waste, indicating that these could be leveraged by practitioners. Our study suggests that it is feasible to forecast food waste data and that distinct waste-affecting variables exist. These variables can be used to explain the causes of waste, and subsequently minimize it. Grocery retailers can, by adopting the conclusions drawn in this study, and integrating them into their strategy, effectively reduce their overall food waste. This helps retailers to save money, mitigates environmental effects and aids them in reaching critical sustainability goals.

# Table of Contents

# 1 Introduction

This thesis aims to evaluate the performance of different forecasting methods on a grocery retailer's food waste data for perishable products. We aspire to investigate if food waste can be explained by statistically significant explanatory variables. The data utilized comes from three retail stores located in Stockholm, Sweden. The insights gained from this study could help retail businesses to improve their forecasts and more accurately determine order sizes, minimizing stockouts and overordering, among other benefits. At the time of writing this thesis, there is little research or knowledge surrounding the topic of forecasting food waste data, both in academia and in practice. Therefore, filling this knowledge-gap and demonstrating how forecasting can help reduce food waste will have important implications for academia as well as for businesses.

In 2015, the United Nations set an objective of lowering food waste by 50% in retail and consumer levels by 2030, as part of their Sustainable Development Goals (SDGs). Additionally, objectives were set to lower waste earlier on in the supply chain (Rosa, 2018). While not obligatory to conform to, the UN goals are helpful in setting guidelines for governments and companies on a global scale. Therefore, reducing food waste through measures such as improved planning and better forecasting (Thyberg & Tonjes, 2016) becomes an important priority for any company that aspires to become more sustainable and help achieve global goals. For example, Axfood aims to halve their food waste by the year 2025, using 2015 as the base year (Axfood, 2019) and so do Coop and ICA (Coop, 2019; ICA, 2018). The public attention over the ethical dimension of food waste puts increasing pressure on companies to show their effort in preventing and reducing food waste (Teller et al., 2018). Companies are facing external pressure from activist groups, journalists, customers (*This Campaigner Wants to End the Global Food Waste Scandal - VICE*, n.d.) and politicians to conform to stricter sustainability standards. Internal demands from stakeholders and shareholders for a focus on sustainability, such as those made by investors and employees, can also be a driving force for the change experienced in the retail industry. These factors make efforts in the reduction of food waste even more imperative and relevant to the retailers' image, longevity and economic and environmental sustainability.

Nordic food retailers operate in a saturated and consolidated market where the realization of economies of scale and scope are both critical in securing current and future competitiveness and profitability (Halloran et al., 2014). For example, Axfood, Coop and ICA together have a 86% market share of the Swedish market (HUI, 2018). This market environment creates incentives for retailers to streamline the supply chain and aim toward even leaner operations. A single actor could potentially reduce a large part of the total food waste in the market, while simultaneously eliminating a high absolute value of its costs, making that actor more competitive and sustainable.

In line with this way of thinking, evidence suggests that a smaller but still significant portion of food in grocery retailing is still wasted. This implies unnecessary costs for all parts of the supply chain. In 2018 the estimated value of food loss at the retail levels in the U.S. was $20 billion (Teller et al., 2018). Approximately 10% of the available, edible food supply in the U.S. ended up as food waste, which equals to 19.5 million metric tons (Buzby & Hyman, 2012). The food waste in the retail sector in the EU-28 countries is approximately one-third of the U.S. levels, equivalent to €7 billion (Stenmarck et al., 2016). 88 million tons of food is wasted in EU in total and approximately 5% of that figure stems from retailers per annum, according to Stenmarck et al. (2016). Moreover, in 2018, Swedish grocery retailers accounted for 100.000 tons of food waste, making up 7.7% of the country's total food waste (Naturvårdsverket, 2020).

## 1.1 Problem areas

According to Thyberg & Tonjes (2016) and Buzby & Hyman (2012), poor forecasting is a leading cause behind food waste in the developed world. Improving the quality of how we forecast therefore seems like an obvious way to reduce the negative impacts that stem from grocery retailers. Nevertheless, Halloran et al. (2014) state that retailers have few incentives to address the issues. This stems from the fact that grocery retailers lack incentives to order the correct amount of food because they can often send it back to producers or wholesalers (Eriksson et al., 2017). For instance, 10% of over-production and high levels of wastage in the UK food supply chain can be derived to a combination of product take-back agreements (TBA:s), contractual penalties, and poor demand forecasting (Parfitt et al., 2010). Furthermore, a clear understanding of the actual scale of food waste at a global, national and store level is lacking (Halloran et al., 2014; Fildes et al., 2019). Naturally, this situation makes it a lot harder for actors to understand the potential financial incentives to reduce it. These are all factors that cause the lack of incentives which ultimately has led to inaction in reducing food waste, according to Halloran et al. (2014).

In addition, many leading retailers have recently adopted sustainability goals in their strategies, including reducing food waste. Retailers also see great advantages in communicating the efforts made to become more sustainable (Halloran et al., 2014). Tools for forecasting food waste could help retailers reach their sustainability goals as well as help them be more proactive regarding reducing food waste. These efforts could also, simultaneously, improve and streamline their supply chains and operations. While some retailers employ tactics such as deep discounts close to expiration dates (Lebersorger & Schneider, 2014a), such actions have not been sufficient enough to curtail food waste adherent to expiration, even if it helps to lessen the impact. Many retailers also use demand forecasting, but with high fluctuations in demand there is a challenge in modelling accurate forecasts (Teller et al., 2018). In conclusion, is a plethora of reasons as to why a model which forecasts food waste could create value for retailers.

Lastly, based on the literature review made for this thesis, it is clear that reliable results with regard to how to forecast food waste are quite scant. Although we found a wide range of research where forecasting models were applied to different use cases and data, none of these studies applied forecasting models on the actual food waste data from retail stores. In effect, this thesis will contribute to the research in this relatively unexplored area. Teller et al. (2018) explored the root causes of food waste in retail stores, something which has been valuable for this thesis in order to further delimitate the research question. To the best of our knowledge, at the time of writing, this thesis is among the first in the world attempting to forecast of food waste and explain the causes behind it.

## 1.2 Purpose and research question:

The purpose of this thesis is to develop a model for forecasting food waste data. Thus, the question this thesis will try to answer is: *is it possible to forecast food waste caused by the retail organization and test if explanatory variables can enhance the predictive power of the forecast*. The results of this thesis will extend the prior literature and can help retail organizations to reduce food waste by insights gained from exploring how and if food waste is possible to forecast. Based on the product's level of perishability and financial importance for the retailer, the following categories were chosen: Meat, Fish, Cheese, Dairy, and Refrigerated Vegetarian products.

### 1.2.1 Expected Research Contributions

This thesis aims to expand the knowledge regarding how to forecast food waste. By comparing between four forecasting models; *Naïve* forecast, Exponential Smoothing (*ES*) models, *Stepwise Regression* models with a number of explanatory variables and lastly a combination model of the *Regression* and the *ES*, conclusions can be drawn concerning which of these models best explain food waste data. The technicalities of these methods will be explained in the methodology section. Furthermore, by using a set of residuals as evaluation criteria, the models will be applied to a set of different aggregated levels in order to asses if these factors affect which forecasting model predicts the data most accurately. The expected research contributions are to develop an understanding or whether or not there is a possibility to forecast food waste data. Furthermore, insights as to how explanatory variables can be used to enhance the predictive power will also contribute to the development of the practical knowledge regarding how to combat food waste using forecasting methods.

### 1.2.2 Delimitations

Research by Teller et al. (2018) shows that the occurrence of food waste is complex to explain. This is mainly due to the huge variety of root causes. Our study aims to predict food waste related to the retail organization's activities by looking primarily at; on-shelf availability and promotion campaigns. We focus on looking at goods that are perishable. The products included in the categories chosen all have defined sizes and weights, meaning that sales through delis or bulk goods are not part of the study. This is mainly due to the different nature of these products, which subsequently leads to worse data, as indicated by Axfood, where the data stems from. It is more difficult to accurately log waste for food that needs to be weighted and lacks a defined package size. Other less perishable, but still spoilable goods such as rice, chocolate, cooking oil etc. could also have been included as they are wasted. However, with a much lower rate of occurrence for waste, these types of goods were not prioritized since they carry less financial importance for the stores and were therefore excluded in the exported datasets. The effort for including these are seen as out of scope and potential rewards in analytic improvement are believed to be minimal.

### 1.2.3 Limitations

The thesis is based on a single Swedish Grocery Retailer, Axfood AB. The data is collected from three separate stores situated within the Stockholm region. Axfood has both franchise and company owned stores. Those included for the purposes of this study are part of the latter category. Because the thesis is limited to a Swedish food retailer the results may not be generalizable across other markets/settings, a limitation that will be touched upon in further detail in the discussion section of this paper. The data spans only three years, from the 1st of March 2017 to the 29th of February 2020. As will become more evident, the hot summer of 2018, in the region may have impacted the results to some minor extent. Furthermore, this thesis only applies forecasting on the waste data and waste is always the dependent variable. No forecasts were produced on any other commonly forecasted variables, such as sales data. Additionally, no qualitative data, in terms of interviews with either store managers or personnel at headquarters were conducted.

### 1.3 Disposition

The thesis is structured as follows. We will first present the prior research literature related to forecasting food waste and food waste in general. After that, we will turn to the methodology. In this part we will elaborate on data description and environment, selection method of variables, products and categories, and model evaluation. In the next part, we describe the modelling of the forecasts. This is followed by the results, discussion of the implications of the results, and conclusions. Lastly, we present potential future research areas.

## 2 Literature review

In this section, we discuss the prior literature on retail forecasting and relevant studies on food waste, that either adds context to our study or provides theoretical background.

### 2.1 Retail forecasting

The general literature on retail forecasting is plentiful (Fildes et al., 2019). However, little to no research specifically focuses on retail forecasting of food waste. There is only a small portion of the literature that touches upon how to forecast food waste in retail stores. Therefore, relevant articles about general principles for forecasting used in the making of this study are included, along with studies about food waste.

Research papers with some relevance focus on forecasting retail data. Fildes et al. (2019) made provide a valuable review of the research literature on retail demand forecasting. However, food waste and forecasting of perishable products are barely mentioned in the study, but other general principles of forecasting were helpful when developing the methodology for this thesis. Van Donselaar et al. (2016) touch upon the topic of forecasting food waste by modelling a regression model with five dummy variables for different price discount classes to capture potential threshold and saturation effects for perishable products. This research aided us in the structuring of our thesis. In addition, Arunraj & Ahrens (2015) developed a *"seasonal autoregressive integrated moving average with external variables"* (*SARIMAX*) model to forecast daily sales of perishable products. Their approach to hybrid models inspired our decision to use a combination model of the *Stepwise Regression* model and the *Exponential Smoothing* model, in order to capture a wider range of effects.

Most of the research papers mentioned in this section are based on observations made in the US, Germany, and the UK. As far as we are aware, no study of this kind has ever been conducted for the Swedish grocery market. While Sweden is a relatively small country, the findings emanating from this particular region could still be relevant for at least other Nordic countries, sharing similar weather conditions, holidays, and market saturation patterns. Furthermore, diversifying the geographical sample used in forecasting research certainly helps to further develop the field by adding more variance. Our research extends the available literature by being one of the first studies that forecasts retail food waste data and does so using longitudinal data spanning a three-year period from five perishable product categories in three retail stores.

To address the research question, we made use of papers that address the general principles of forecasting. Athanasopoulosa & Kourentzes (2020) provides plenty of practical guidelines for how to model hierarchal forecasts. Their paper discusses how to approach data with a lot of sparsity, which is highly relevant for this study since the observations in the food waste data often contain zero values, making it necessary to aggregate the data. Athanasopoulosa & Kourentzes (2020) recommend which error metrics to be used when making hierarchical forecasts and how to make the metrics relative. Fildes et al. (2019) offers a coherent framework for the different dimensions of forecasting- a framework that we made use of when choosing between different crossroads in terms of which dimensions and hierarchical levels the forecasts should be conducted on.

Ord et al. (2017) provides valuable advice and guidelines regarding the principles of business forecasting. Much of the methodology in this thesis is based on the contents of this book, as it provides guidance and insight into basic principles of forecasting, collected in one piece of literature. It is a textbook, and as such gives straightforward and scientific explanations for topics such as error values, model construction and method selection. In its second edition it also contains guidelines and exercises on how to use the programming language R and RStudio, which formed the basis of the forecasting in this thesis. Our study serves as an example of how such relatively basic principles can be put into use in order to advance forecasting research and how this can assist businesses.

It is difficult to determine beforehand which forecasting model will be the most accurate when forecasting food waste. This difficulty is mainly due to the differences between food waste data and other types of data, such as sales. The available literature covers many aspects of retail business forecasting, such as forecasting  demand for a variety of categories, measure saturation levels and price elasticity, along with the intricacies of data aggregation. These areas of research have comprehensive conclusions regarding utilization of error measures, explanatory variables and general method tradeoffs. However, little attention has been paid to the other side of the equation, namely waste. We have found no literature that specifically addresses the question of forecasting food waste data. Therefore, it is difficult for us to use findings from earlier research efforts explicitly addressing how to model suitable forecasts for forecasting food waste.

There are articles that address solid waste, but not specifically retail food waste. Due to the differing nature of the two types of waste, this literature was deemed not relevant for our purposes. Solid waste includes not only household waste from all parts of the supply chain, but also materials such as metal, plastic and paper along with food. Although no studies about forecasting food waste were found, we will use existing literature on retail forecasting and build upon established assumptions in order to expand the research area into the field of food waste.

## 2.2 Food waste

A lot of literature regarding the causes of food waste, methods for managing and reducing it as well as estimates of past and future food waste for different industries and regions have been published in the academic arena as well as in official governments reports and from NGO’s. Although much of this literature is not directly relevant for the forecasting methodology in this paper, it is an indicator of the fact that statistical models for the reduction of food waste is a very topical research area. It also helps us in identifying waste drivers that can be added to as variables to the forecast models.

As mentioned earlier, Thyberg & Tonjes (2016) and Buzby & Hyman (2012) state that poor forecasting is a leading cause in food waste for the developed world. Both Lebersorger & Schneider (2014) and Eriksson et al. (2017) state that retailer's food waste data in the Swedish market is all too often impaired by lacking routines as well as TBA:s. The insights from these papers gave us guidelines regarding the limitations which food waste data usually entail. They also gave credence to the assumption that more forecasting is needed in order to curtail food waste issues.

Teller et al. (2018), by combining qualitative and quantitative research methodologies, explores the root causes of food waste in retail stores using a sequenced, multi-method approach. First, exploratory research, which involved 28 case studies from various retail formats, were conducted. By adding secondary data from previous research, the authors found several causes linked to food waste. By linking dependencies between the findings from the case studies and the secondary data, they could identify root causes. These root causes were then presented to 12 food waste experts through semi-structured interviews during which various fields in the area of food waste were also discussed. Teller et al. (2018) found that food waste in retail stores is the result of a combination of internal and external factors. The internal factors are resources and operations processes of a retailer and store, and the external factors are demand patterns and in-store consumer behavior.

Moreover, Teller et al. (2018) identified a number of root causes of food waste, whilst also clarifying that these are not the only causes of food waste. Root causes of food waste for the retail (parent) organization are product quality standards, width and depth of product range, on-shelf availability, promotions and marketing campaigns. The root causes at store levels are type of store format, number of product category, store operations and store personal & management. The last root causes of food waste can be derived from the customers' demand patterns, their instore behavior and high expectations regarding product quality. The information presented by the authors gives us insight into what factors form the basis for what leads to food waste. Even if all effects are not fully explored within the modeling of this thesis, the knowledge of their existence and how they affect waste is still valuable for analysis and discussion.

# 3 Methodology

The methodology section is divided into three parts: 3.1 Data description, 3.2 Selection methodology and 3.3 Forecasting method evaluation methodology. Data description describes the prerequisites of the data used for the forecasts. The selection methodology presents what decisions were taken and why, regarding the modelling, what hierarchies and categories to forecast, and which variables to use in order to explain food waste. The method evaluation methodology describes the process used for evaluating the result.

## 3.1 Data description

This part will explain the underlying data upon which the forecasts will be based. In section 3.1.1, the research environment, data description, and also additional parameters which affect the underlying data, such as lead times and routines regarding food waste, is presented. Limitations with the data is presented in section 3.1.2. Section 3.1.3 explains how the data was cleaned from errors and other factors that can affect the result. Lastly, the data is explored in section 3.1.4 in order to give some insight into how the underlying data, which the forecasts are based on, looks like.

### 3.1.1 Research environment and data description

The empirical data that forms the basis for analysis and the forecasting modeling was provided by Axfood AB. The retailer has an approximate 20% market share in the Swedish grocery trade. The group owns some 300 stores, and are through e-commerce and collaborations connected to 1200 additional locations. As already mentioned, Axfood has a strategic focus on sustainability and aims to be climate neutral by 2020, while also cutting food waste by 50% by 2025.

The data stems from three stores located in the Stockholm region, the stores are all part of the *Hemköp* chain. The data set contains information from the period 2017-03-01 to 2020-02-29 regarding daily sales and waste in SEK per stock keeping unit (SKU) per store. Each observation has additional descriptive information about the SKU: s belonging subcategories and categories. Furthermore, the data set contains descriptive information about whether the sale was campaign related or not. The campaigns are divided into local or central campaigns. Central campaigns are nationwide and local campaigns affect one specific store or a smaller sales region. Table 1 displays the different variables of the data for each observation from the original file received from Axfood AB.

*Table 1 Data exported from Axfood*

| Variable | Example data | Description of variable |
|---|---|---|
| **Description of Category** | DAIRY | Descriptive text of the category. This is the most aggregated category the data set contains. |
| **Description of subcategory 1** | BUTTER | Descriptive text of the subcategory 2. This subcategory is less aggregated than the Category variable. |
| **Description of subcategory 2** | BUTTER ECO | Descriptive text of the subcategory 2. This subcategory is less aggregated than the subcategory 1 variable. |
| **Description of SKU** | BUTTER ECO 500G | Descriptive text of the SKU. This variable explains the data on stock keeping unit level and is the least aggregated variable. |
| **Description of store** | STORE 1 | Descriptive text of the store. |
| **Date** | 2018-08-26 | Date of when the observation was recorded. |
| **Sales amount excluding VAT** | 1 058 | The amount of sales for an SKU, per store per day. |
| **Waste amount** | 5 110 | The amount of waste for an SKU, per store per day. |
| **Campaign related to the sale** | LOCAL PROMOTION | This variable contains three unique values: central promotion, local promotion and normal sale. The variable explains if the SKU had a promotion or not related to the sale. |

According to Axfood, lead times[1] from central warehouses to the store differ from 24 to 72 hours, depending product category. The waste routines are standardized across the four original stores chosen (as will become evident later on, one store was omitted from the final analysis). Products are flagged by the store's system when the expiration date approaches. When products are flagged differ depending on the consumer's consumption time of the product. This enables staff to take proactive measures in order to boost sales of the products and still allow the customer to consume the product before its expiry date occurs. Proactive measures include

[1] Lead time refers to the delay between a location placing an order for an item and said item arriving at the destination.

putting red stickers[2] on the products, or similar promotion activities. Directly after the expiration date has passed, the staff logs the product as waste in the system. This makes the logged waste data overlap exactly with the actual expiration date of the product. There are different types of waste used when logging products as waste. For example, staff can log waste for a product that will be consumed in the personnel breakroom. The waste data exported for this thesis only include observations with the cause "identifiable waste". This is done to ensure that external factors will not affect the data and results since this identifiable waste category only includes wasted products that have passed expiry dates or products that have been donated to charity.

In order to make an informed and balanced choice of the product categories to include in the forecast, we developed a set of criteria. Studies have shown that perishable products contribute most to avoidable food waste (Lebersorger & Schneider, 2014;Brancoli et al., 2017). Therefore, perishability became one of the selective factors. Furthermore, in order to combat the problem of retailers having few incentives to reduce waste, which Halloran et al. (2014) discuss, the categories should have high turnover in general and thus be important financial categories for the retailer. Having these criteria, the products selected from the database stem from the following categories: Cheese, Dairy (milk, yogurt, butter), Fish, Meat and lastly Refrigerated Vegetarian products.

Fruits and vegetables also fit these criteria and are large enough to warrant forecasting, and as they have a relatively short shelf life, they can often be drivers of food waste (Lebersorger & Schneider, 2014;Brancoli et al., 2017). However, since the partner company lacked sufficient data for the particular product category the results obtained through forecasting would not be as reliable as from the other categories.

### 3.1.2 Limitations with the data

As with most data sets, ours comes with limitations. The data set for this study has a lot of infrequency in terms of observations that have actual logged food waste in them. Only 1.4 % of the approximately 3,400,000 million original observations contain food waste data. Furthermore, the data in the waste variable can be caused by human errors. Since the logging of food waste is mostly done manually, there is a risk that some observations of waste amounts

[2] Red stickers are put on products that are about to expire, to indicate a deep discount of the product.

are incorrect. Although extreme values can be identified and adjusted, smaller errors are harder to find. The data also contains negative sales and waste data. The reason for this is that negative waste or sales occur if waste is logged incorrectly or if items sold are returned by customers. This will counteract the inaccurately logged waste, but it is still hard to verify that all the observations in the food waste variable represents reality.

Although the stores have routines for managing food waste, there is always the risk that the personnel do not follow them. For example, personnel might not log food that is thrown away, log the food long after the actual expiration date or accidently use the wrong cause of the waste when logging a product. Moreover, a lot of food waste is sent back to the suppliers through TBA:s (Cicatiello & Franco, 2020). It is problematic to adjust the data for these factors as doing so would require additional information which was not compiled for this study.

The stores chosen are in direct ownership of Axfood AB and the routines in these stores are determined by Axfood AB. The other types of stores are privately owned under the *Hemköp* franchise, leading to the store owners with more freedom regarding routines. We choose to not include these stores because Axfood indicated that the risk is higher that routines for logging food waste is not as uniform and well managed compared to the group of wholly owned stores.

#### 3.1.3 Data cleaning

Originally, data from four stores were exported from the Axfood's database. However, one of the stores only had 930 daily observations. This is 166 fewer days of observations compared to the other three stores, which all have 1096 daily observations each. Unfortunately, it is the last 166 days of the time series that are missing from this store, which could therefore have some significance on the result. The reason for this is that the *ES* model weights recent data as more important, which is why losing the last observations is problematic. All observations from this store were therefore removed from the data.

Initially, the data-file contained 3 482 789 observations. 293 observations which contained summarizing values of sales were removed since they only contained compiled sales per subcategory data. This made them unusable for time series forecasting as they only contain total sales and cannot be tied to a specific date. Our focus is on the three-year period between March 1, 2017 and February 29, 2020. Hence, 78 623 observations which contained dates before 2017-

03-01 and after 2020-02-29 were also removed. Lastly, all 539 638 observations from the store that only had 930 observations were removed. After the cleaning, 2 737 038 observations were left. Table 2 presents the sample selection process. Some observations in the waste data contained extreme values well over 10 000 SEK, despite being tied to a single SKU each. These values were replaced by zero in order to reduce the extreme effects which outliers have on the *ES* model and the *Regression* model, when fitting these models.

*Table 2 Cleaning of sample selection*

| Process | Number of observations |
|---|---|
| Initial observations: | 3 482 789 |
| Observations containing summarizing values: | 293 |
| Observations recorded before 2017-03-01 and after 2020-02.-29: | 78 623 |
| Observations removed from store with 930 daily observations: | 539 638 |
| Observations after cleaning process: | 2 737 038 |

### 3.1.4 Data exploration

In order to increase our understanding of the waste, the data was explored. This was necessary in order to ascertain whether enough observations containing values above zero were present, which are needed in order to produce forecasts. Without sufficient data, the models cannot be constructed properly, and the forecast fails.

*Table 3 Summary statistics by category*

| Category | Vegetarian | Dairy | Cheese | Meat | Fish | All categories |
|---|---|---|---|---|---|---|
| Total waste amount (SEK) | 25 708 | 662 349 | 195 770 | 1 001 466 | 405 560 | 2 290 853 |
| Total sales amount (SEK) | 2 069 020 | 141 410 546 | 79 601 468 | 97 621 700 | 54 152 344 | 374 855 078 |
| Number of observations | 45 941 | 1 410 703 | 711 551 | 323 582 | 245 261 | 2 737 038 |
| Distinct SKU: s | 127 | 1 277 | 875 | 660 | 473 | 3 412 |
| Distinct Subcategories 2 | 6 | 75 | 41 | 38 | 17 | 177 |
| Distinct Subcategories 1 | 6 | 154 | 69 | 38 | 23 | 290 |
| Number of waste observations > 0 | 612 | 21 624 | 3 415 | 10 038 | 3 022 | 38 711 |
| Average amount of waste per observation >0 (SEK) | 42 | 31 | 57 | 100 | 134 | 59 |
| Share of observations containing waste data | 1,33% | 1,53% | 0,48% | 3,10% | 1,23% | 1,41% |
| Standard deviation of sales | 34,73 | 165,12 | 171,87 | 464,47 | 626,27 | 295,36 |
| Standard deviation of waste | 6,88 | 8,37 | 8,33 | 29,78 | 26,68 | 14,98 |
| Coefficient of variation of sales | 0,77 | 1,65 | 1,54 | 1,54 | 2,84 | 2,15 |
| Coefficient of variation of waste | 12,29 | 17,84 | 30,27 | 9,62 | 16,14 | 17,89 |

The data from Table 3 was exported after the data cleaning was conducted in order to present the relevant data for this study. As can be observed, Meat is the category which has most waste in terms of SEK and on average per observation. Furthermore, at 3.1%, Meat contains the most waste in relation to the total number of observations. Dairy has the largest number of observations with values greater than zero.

The number of waste observations greater than zero summarizes to 38 711 and thus 35.5 observations on average per day, given that the number of daily observations is 1096. Furthermore, all of the categories, except Refrigerated Vegetarian products, contain approximately slightly more than 3 observations per day. This indicates that a forecast would be feasible, at least on an aggregated level. The waste could also be condensed into a few subcategories or SKU:s, within each category, making it possible to conduct forecasts on less aggregated levels. As can be seen in the Table 4, the number of observations per store is relatively equally distributed. This lessens the impact of any variation between the stores' waste routines. That is, if one store would have represented the majority of the observations, that store's routine would have been affecting the overall result to a much greater extent.

*Table 4 Summary data per store*

| Store name | Waste (SEK) | Sales (SEK) | Observations |
|---|---|---|---|
| STORE 1 | 783 389 | 161 229 191 | 973 862 |
| STORE 2 | 612 310 | 72 686 900 | 717 317 |
| STORE 3 | 895 154 | 140 938 987 | 1 045 859 |

Figure 1 and Figure 2 display all logged waste and sales data respectively by day across the three years for all categories and stores. Apart from a slight increase every July, the waste data does not show any immediately obvious patterns when compared to the annual seasonality of sales which can be seen in Figure 2. For the sales, there were decreases and increases during July and December, respectively. The decrease in sales during the summer (particularly July) is likely what is causing the increase in waste. Thus, using sales and the summer months as explanatory variables for the *Regression* model appeared as an obvious implication after initially reviewing the data.

*Figure 1*

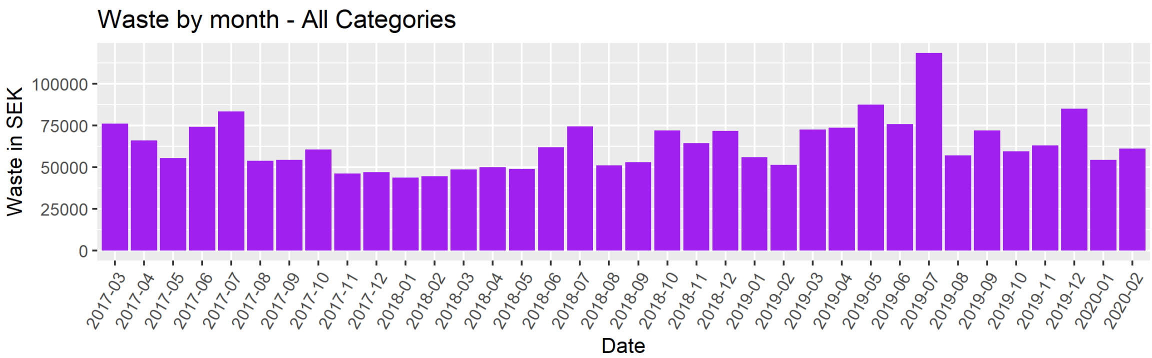
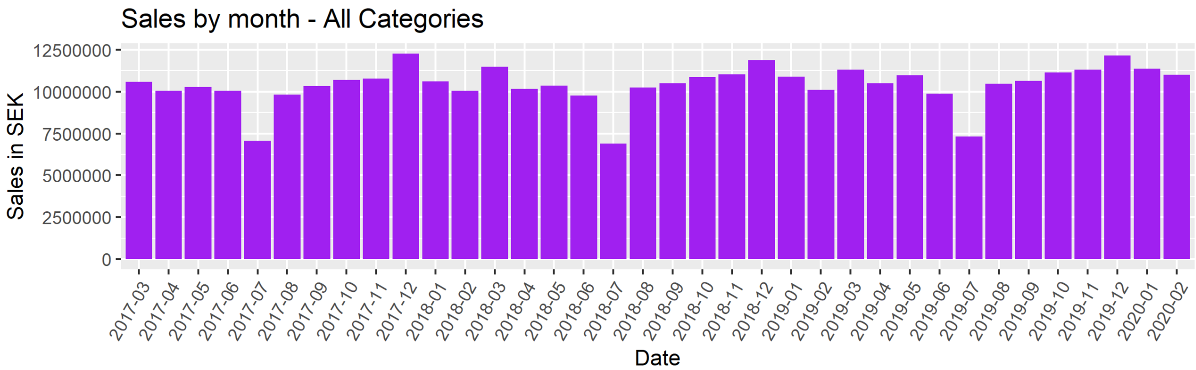



*Figure 2*




## 3.2 Selection methodology

Before conducting the forecasts, we needed to make a number of decisions. The sections below explain what selections were made and why. In sections 3.2.1, selections regarding aggregation dimensions and levels are explained. Secondly, the selection of which product and categories to forecasts are presented. Which error measures to use for evaluating the forecasts are presented in 3.2.3. Lastly, the selection of which method to use for creating a forecasting model which can explain food waste are described in section 3.2.4.

### 3.2.1 Selection of aggregation dimensions and levels

When selecting the data there are choices to be made about the different dimensions of aggregation, as well as which model that the forecasts can be based on. Fildes et al. (2019) divide these choices into three different dimensions: product, time and supply chain. The product dimension can further be divided into three aggregated levels: SKU, brand, and category. The product hierarchies in the data for this thesis correspond quite well with the product dimensions as described by Fildes (2019). The data in this thesis has SKU level, several hierarchies in category levels but misses the brand level because no exported variable from Axfood contained the SKU's brand. The time dimension can be divided into seven aggregated levels: hour, day, week, month, quarter season and year. Given that our data consists of daily observations and spans three years, almost all of these hierarchies are viable options with year being somewhat out of bounds, since that would give us only three complete periods. Moreover, the supply chain dimension can be divided into three aggregated levels: store, distribution center and chain level. Our data only contains the store level aggregation, as the entire chain and distribution centers are not included in the dataset.

When choosing between the dimensions and the aggregated levels there is a trade-off between either forecasting on more descriptive data but with less data points, or less descriptive data but with more data points. Song (2015) suggests that it is beneficial to forecast and model data on more aggregated levels since stronger and more seasonal information can be collected. This general rule was used when choosing between the dimensions and hierarchies.

As our purpose is to look for effects that impact waste, and not improve the forecasting for a specific outlet, it was deemed better to do all this on an aggregate level. Therefore, we chose to aggregate the data for all three stores and not do distinct forecasts for the separate stores. Regarding the time dimension, it is more relevant to retain the daily observations because the stores have daily replenishments. Also, retaining the daily observations makes it possible to capture weekly seasonality, if that exists in the data. If the data has been reported from a distribution center, it would be more appropriate to use a weekly aggregated level instead. The horizon set for every forecast was therefore seven, which denotes seven days, or a week, making up one period of the forecast. For a dataset of 1096 daily observations, this amounts to 155 full periods.

In order to improve the modelling accuracy and capture more of the different promotional, corporate and seasonal effects, we decided to conduct forecasts in both a top-down and bottom-up fashion. Traditionally, forecasts are completed with only one of these two methods (Fildes et al., 2019). Bottom up entails building forecasts at the lowest level (in our case *A1*) and then compiling those in order to explain more aggregated levels. Top down works by doing one forecast on the highest (*A5*) level (overall waste) and then disseminating that forecast based on the percentages of the total that each SKU contributes. An important concern with both of these approaches is that aggregating or disaggregating the models may lead to misspecifications and estimation errors. The *Combination* approach instead computes forecast for every aggregated level, and therefore bypasses the issues of incorrectly specified models and the errors introduced when trying to aggregate up or down. This approach allows for incorporating more detailed information on products and categories (Kourentzes & Athanasopoulos, 2019a). Further reasoning behind conducting the modelling in this fashion emanates from the purpose of the thesis. As the focus of this paper is on food waste and not the technicalities of forecasting, the decision was made to focus on the effects and correlations of explanatory variables of food waste, such as promotions and holiday events.

### 3.2.2 Product and category selection

Having determined that the forecasts were to be focused on daily observations and on the product dimension, further decisions had to be made regarding which aggregated levels the product dimension forecasts should be focusing on. In the dataset, there are five different levels of aggregation for the product dimension. As can be seen in Table 5, the first aggregated level is at SKU level (*A1*). The second is the subcategory 2 level (*A2*) and the third is at the subcategory 1 level (*A3*). The fourth is the aggregated data per category (*A4*). The last and most aggregated level is for all data (*A5*). The aggregated levels will subsequently be called: *A1, A2, A3, A4,* and *A5*. The levels are numbered in orders of aggregation, with *A1* being the least aggregated and *A5* being the most aggregated level.

*Table 5 Naming convention for the different aggregated levels*

| **Aggregated level** | **Variable name in the data set** |
|---|---|
| *A1* | SKU |
| *A2* | Subcategory 2 |
| *A3* | Subcategory 1 |
| *A4* | Category |
| *A5* | All data |

While forecasting at the *A1* level could be informative, the low number of observations at the SKU:s limits the possibility to conduct proper forecasts. Although one forecast was made on the most wasted SKU in SEK (grilled chicken) in *A1*, focus was put on higher aggregated levels.

Furthermore, in some instances there is no difference between *A2* and *A3*, as they both feature the same data and description for certain categories which would have made some forecasts redundant. Therefore, it was decided that a focus on forecasts on level *A3* would be more beneficial, since it sometimes has more underlying SKU:s, and therefore *A2* was subsequently abandoned. *A3* contains 177 separate categories. Out of these, the 30 categories that had the most waste measured in SEK were chosen. As mentioned in section 3.1.1, the choice of categories is partly based on the financial importance they have for the retailer. Thus, basing the selection on value of waste in monetary terms for *A1* and *A3* follows from the fact that the forecasts on these categories possess the most potential to reduce monetary losses from food waste. For *A5* and *A4*, all categories available were chosen. *A5* consists of only one category that contains all waste data of the dataset. *A4* is made up of five categories, that is: Cheese, Dairy, Fish, Meat and Refrigerated Vegetarian products.

For *A1*, we choose the one product that had the most waste. Initially, we planned to conduct more forecasts at the *A1* level, but due to lack of data in most of the SKU:s, only one SKU was chosen. This was done to prove that the model still worked for this aggregated level, provided there are enough data points and observations to construct a forecast. As such, the *Combination* approach to modelling is still applied as the levels selected are all forecasted individually. The only limitation is that *A2* is disregarded and only one SKU from *A1* is forecasted.

### 3.2.3 Error value selection

In order to evaluate forecasting methods, different error values were chosen for examining the different aspects of the models. By using different residuals, more aspects of the performance of the forecast methods can be derived. Athanasopoulosa & Kourentzes (2020) suggest that in terms of error metrics *Mean Error* (*ME*) and *Root Mean Squared Error* (*RMSE*) should be used when making hierarchical forecasts and benchmarking a model's performance against the *Naïve* forecast in order to attain a relative value.

Mean absolute error (*MAE*) and root mean square error (*RMSE*) were used as evaluation criteria regarding the predictive power of the methods. *MAE* is the mean value of the errors produced when comparing the actual values to those produced by the forecast. Because of this, *MAE* is particularly robust against outliers. These errors are called residuals. *RMSE* is the square root of the measure *MSE* or mean square error. As the name suggest *MSE* squares the mean errors, whereas *RMSE* gives us the square root. Both measures are quite similar, but *RMSE* is generally more intuitive as it is not displayed in squares which is why it was chosen. The formula for these residuals can be found in Formula A 1 for *RMSE*, Formula A 2 for *MSE* and Formula A 3 for *MAE*. The *RMSE* and *MAE* are used in Chu & Zhang (2003) as error measures. They state that since there is no best-practice measure for the performance of every forecast, it is better to use multiple criteria for a comprehensive assessment.

*Akaike information criterion* (*AIC*) was used in order to assess which model performed best within a forecasting method, a selection criteria also used by Kourentzes & Athanasopoulos (2019) and recommended by Beier et al. (2001) as a criteria for choosing the best performing model. *AIC* is a relative measure, meaning that it is not fit for comparisons between different *methods*. It only serves as a relative measure between different models of the same type, with a lower value being more favorable, implying a lower level of uncertainty projected. Furthermore, *AIC* is used to combat multicollinearity, more about this can be found in section 3.2.3.1. The formula for *AIC* can be found in Formula A 4.

*Mean error* (*ME*) was used to validate that the forecasts had been calculated correctly. *ME* is used as a basic estimator of errors for a forecast. It is useful in detecting a systematic bias, as the value of the error measure will be large and either positive or negative when the forecast is consistently under- or overestimating the actual outcome. Equations used for all error values are available in the appendix. The formula for *ME* can be found in Formula A 5.

#### *3.2.3.1 Combating multicollinearity with AIC*

*AIC* is also employed to combat multicollinearity that may arise in the *Regression* model. Some of the variables which are based on dates, such as holidays events or the weekly seasonality, are similar because they occur close to each other or simultaneously, but there is no strong evidence of them being overtly correlated with each other. *Stepwise Regression* eliminates variables that are deemed superfluous. The method bases this decision on whether or not the variable negatively impacts the *AIC*. If the variable affects the *AIC* positively, the *Stepwise Regression* keeps the variable in the model. When, for instance, adding lags for up to seven days, *Stepwise Regression* removes lag 7, since weekly seasonality is already encoded. While not a perfect solution, it suits the purposes of this thesis as the goal is not focused on creating a perfect forecast (Beier et al., 2001). The goal of the *AIC* is to improve predictive power while limiting model complexity. Hence, variables that do not improve the *AIC* are removed. This is contrastive to more traditional modelling based on statistical tests, which are more biased towards overfitting, which would produce better residuals but ultimately not yield valid results.

Using the *AIC* as a selection criterion means that the models may not pass traditional statistical tests, since their evaluation is based on a different modeling methodology. However, traditional tests are usually based on in-sample data which does not guarantee predictive capability, something that the *AIC* rectifies to some extent, by basing evaluations on out of sample errors. Ultimately, the *AIC* is an imperfect measure (like all measures) albeit a useful one in combating multicollinearity, especially since the *Stepwise Regression* model is utilized either way. Homoscedasticity, i.e., if all its random variables have the same finite variance, was checked using scatter plots, and there was no identifiable presence of this phenomena.

#### *3.2.3.2 Regression coefficient to explain explanatory variables*

To understand how an explanatory variable impacts waste, the *Regression Coefficient*, or estimate, of the variables is used. The estimate measures the effect that a unit-increase of an explanatory variable has (while other explanatory variables remain the same) on the dependent variable (Davies & Newbold, 1986). In our case, the dependent variable is waste. By using the methods explained by Bill Evans (2010), the estimates were interpreted and converted into more discernable values. Both the dependent variable of waste (Y) and Sales ($yS$) were converted into logarithms in order to make the values absolute rather than relative. This procedure makes it highly difficult to give absolute values for the conversions. Using the aforementioned formulas, the estimates are converted thusly. For non-logarithmic values, such as campaigns, the estimate one units change in the independent variable gives a $\beta$ unit change in the dependent. For the logarithmic variables, such as sales and the sales lags, a 100% change in the independent gives a $\beta$ change in the dependent. Finally, for the binary dummy variables, a change from 0 to 1 in the independent generates a $\beta$ change in the dependent.

### 3.2.4 Forecasting method selection

In order to explain what causes food waste a model was built using a regression method where the dependent variable is food waste and a set of different explanatory variables were utilized as independent variables. The model was expanded upon by using lags for both sales and waste along with constructed explanatory variables (dummy variables). Additionally, promotional effects were added to see if waste could be explained by such events. Lags are constructed through postponing figures and duplicating data from origin t onto t+1. This is done in order to see if the chronology and the time at which, for instance, sales and waste occur help improve the forecast. Constructed explanatory variables are manually computed variables that coincide with a specific effect or instance that would impact the data and thus the forecast. The constructed explanatory variables used in this thesis are all binary, meaning that there either is or isn't an effect present, signified by either 1 or 0. For example, the explanatory variable for the first week of August (*AugD1*) consists of zeroes (0) for all days that are not the first week of august, and ones (1) for all days that are. Such variables are often called *dummy variables* but will be referred to as explanatory. All the 65 explanatory variables used in the *Regression* model can be found in Table A 1. On A1 level the promotion variables indicate if a promotion is there or not, much like a binary variable. But on higher levels, the promotion variables show how big share of the products in a category or subcategory that have a promotion.

Since no study where forecasting is applied to food waste was found, the selection methodology for which variables to use are based on common variables used in models for forecasting retail sales. The variables used in the model can be divided into lags (of waste and sales), holiday events, promotions and weekly seasonality, effects which have impact on retail sales (Arunraj & Ahrens, 2015a). The sales and waste lags were used in order to assess if an increase/decrease in food waste or sales from previous days affects food waste. Certain holiday events and promotions usually entail a higher ordering amount because of the spike in demand, which could result in food being wasted because of a wrongly predicted demand forecast. Weekly seasonality is interesting because it could be used to identify if the waste routines have commonalities in terms of certain days waste is logged in the system.

The purpose of including these variables is to see if they can help explain why waste occurs and to what extent they help explain. The *Stepwise Regression* model was determined to be most fitting for the purpose of determining this[3]. *Stepwise Regression* selects between different variables by combining them back and forth in different constellations in order to find the optimal combination based on the error metric *AIC*. When the inclusion of a variable gives a higher *AIC* value, and thus makes the model worse at predicting the actual outcome, it is

[3] Alternative approaches to forecasting are available. However, the Stepwise approach allows us to efficiently produce a forecast and since the purpose of this thesis is to determine the effects that help explain waste and not to produce the most accurate forecast, it was determined to be appropriate.

removed. When the remaining variables removal worsens the model's *AIC* value, *Stepwise Regression* stops its selection process and provides a finished model. The variables that are left assist the forecast in becoming more accurate to a greater or lesser extent and are therefore said to be explanatory (Ord et al., 2017). While a SARIMAX model, as used by Arunraj & Ahrens (2015) might have been more powerful in its predictive power, it was quickly deemed too difficult to produce. Given the time constraints, and the technical challenges of producing a proficient SARIMA and ARIMAX models for the levels and then combining them. Such methods can be powerful but run the risk of being poorly optimized and having lesser fit simply because they are more demanding. Thus, we made the decision to use the Stepwise Regression as the main approach instead.

The model for the *Regression*, with all possible variables included, is presented below. For a detailed description of the variables see Table A 1. As described in section 3.2.4, not all variables are utilized in each run of the *Regression*. Information on which ones were included for specific forecasts can be found in Table A 5. The *ES* and *Naïve* use the same data as the *Regression*, including the variables but they do not impact the forecast as the methods do not retrieve information from them. The model for the *Regression* is:

$$
\begin{aligned}
\ln(Y_{it)} = {} & \beta_0 + \ln(\sum_{k=1}^{6} \beta_{1k} Lag_{itk}) + \ln(\sum_{k=1}^{6} \beta_{2k} Slag_{itk}) + \beta_3 Central_{it} + \beta_4 Local_{it} + \beta_5 Normal_{it} + \beta_6 Week_{it} + \beta_7 MDF_{it} + \beta_7 EDF_{it} + \beta_8 CDF_{it} \\
& + \beta_9 EDt_{it} + \beta_{10} EDt_{1_{it}} + \beta_{11} EDt_{2_{it}} + \beta_{12} EDt_{3_{it}} + \beta_{13} EDt_{4_{it}} + \beta_{14} EDt1_{it} + \beta_{15} EDt2_{it} + \beta_{16} EDt3_{it} + \beta_{17} EDt4_{it} \\
& + \beta_{18} MDt_{it} + \beta_{19} MDt_{1_{it}} + \beta_{20} MDt_{2_{it}} + \beta_{21} MDt_{it} + \beta_{22} MDt_{it} + \beta_{22} JDt_{it} + \sum_{k=1}^{6} \beta_{23k} JDt_{k_{itk}} + \sum_{k=1}^{4} \beta_{24k} JDtk_{itk} \\
& + \beta_{25} NyåaraftD_{it} + \sum_{k=1}^{10} \beta_{26k} JanDk_{itk} + \beta_{27} JuliD1_{it} + \beta_{28} JuliD2_{it} + \beta_{29} JuliD3_{it} + \beta_{30} JuliD4_{it} + \beta_{31} AugD1_{it} \\
& + \beta_{32} AugD2_{it} + \beta_{33} AugD3_{it} + \beta_{34} AugD4_{it} + \ln(\beta_{35} yS_{it}) + \varepsilon
\end{aligned}
$$

All variables included above are the explanatory ones used for the *Regression*. The equation shows us that the dependent variable Y, which is the waste data in the origin, is correlated to and explained by the variables included. For instance, "*Central*", stands for the central promotional campaigns, while *EDt3* stands for the third day after the middle of the Easter Holiday.

## 3.3 Forecasting method evaluation methodology

As Athanasopoulosa & Kourentzes (2020) recommend, by comparing the residuals to a benchmark, a better assessment can be made regarding of how well the *Regression* method can explain food waste. Therefore, three other forecasting methods were used as benchmarks. A description of the comparison between models are found in the section 3.3.1. Additionally, the method *Rolling Origin* was applied to further enhance the evaluation basis (Kourentzes & Athanasopoulos, 2019; Ord et al., 2017). The section 3.3.2 further explains how the *Rolling Origin* was conducted[4].

### 3.3.1 Comparing between forecasting methods

In order to evaluate the performance of the *Regression* model, three other forecasting methods were used as benchmarks. By comparing the *Regression* model against other methods, it is possible to determine if the *Regression* model is the best viable method to use in order to explain food waste. This reasoning stems from the fact that by using other models as benchmarks, the error metrics become relative, making them more easily comparable (Athanasopoulosa & Kourentzes, 2020; Arunraj & Ahrens, 2015b).

First, the exponential smoothing (*ES*) model with the best performing *AIC* value was used. The *ES* method uses weights in order to make proportional adjustments to latter and earlier observations. *ES* works by prioritizing observations that happened more recently to the forecast origin and weighs older ones as less impactful. The aim of a *ES* model is to describe some kind of average of the recent behavior of a time series (Ord et al., 2017). The *ES* model is proficient at finding structural patterns inside the data, while the *Regression* finds patterns by using external variables. Thus, comparing the performance of the two models can help explain how explanatory the *Regression* model's variables are, relatively to the structure of the waste. *The Error Trend Seasonality*, or *ETS* framework as defined by Rob J. Hyndman & Yeasmin Khandakar, (2008) is used in order to categorize the models produced by the *ES*. A total of 12 different models exist. *ES* and *ETS* is further explained in the appendix. Formula A 8 display the formula for the *ES* method.

The second method used was a *Naïve* forecast (the true *Naïve* forecast). The *Naïve* forecast only contains information from the latest data point of a time series, "t", and is therefore fully biased on what happened in that instance. The *Naïve* assumes that an event will proceed in line with the latest time series and will continue along the same trajectory without deviation. The *Naïve* is a very simple form of forecasting and assumes that what happens in its origin will continue

[4] The forecasting is done through RStudio[4], which provides sufficient tools to produce a reliable model. RStudio is a bridging tool for the programming language R and allows users to import data packages designed by other users or dedicated developers. The packages used for the purposes of this thesis can be found in the end of the Appendix. The data used is exported from Axfoods central database as a converted CSV file and then restructured in RStudio.

indefinitely. As such, it serves as a good basic benchmark, since any proper forecast should at the very least outperform the *Naïve* based on the above criteria (Makridakis et al., 2018; Ord et al., 2017). The formula for the *Naïve* forecast can be found in Formula A 6.

Last, a *Combination* model between the *Regression* model and the *ES* model was used. The *Combination* method takes the average of two or several individual forecasts. In general, the *Combination* between a set of forecasts often produces more accurate forecasts compared to individual methods, because two, or more models, with different capabilities are integrated rather than a single specific model with limited capability (Ord et al., 2017; Arunraj & Ahrens, 2015a). As mentioned previously, while Arunraj & Ahrens (2015) inspired the choice to produce a combination forecast, their use of ARIMAX and SARIMA was not followed up in this study due to limitations and the technical difficulties of producing such models. The *Combination* model that was used had a weight of 0.5 for both the *Regression* model and the *ES* model, in order to give even parts to both models . The formula for the *Combination* model forecast can be found in Formula A 7. The decision to use a 50/50 split of the different methods for the combination was based on the lack of knowledge of how the models would interact with one another and how they would perform on the error values individually. While more optimal combination ratios might exist for the different individual categories and aggregated levels, such testing would have become both arbitrary and highly time consuming. Therefore, it was deemed more fitting to use a common approach for all forecasts.

### 3.3.2 Methodology for producing error values

Using multiple error windows is a well-accepted practice and *Rolling Origin* is a common method for achieving this (Athanasopoulosa & Kourentzes, 2020). Therefore, that method is used in order to improve the error values. *Rolling Origin* is a method to expand the data for which models are trained and to generally improve forecast accuracy. Forecasts starts at origin t, with t+m (where m is the total amount of periods ahead) observations available. Forecasts are then generated successively (one-step-ahead) continuously until the data set runs out. This effectively expands the training set while simultaneously decreasing the test set. The *Training Set* refers to the data which the model is trained on, and the *Test Set* refers to the data which the model's predictive power is tested on.

The *Regression* model is based on iterative forecasting, meaning that as the forecast is produced for each subsequent observation, more and more residuals will start to accumulate as the origin time-series becomes more distant. *Rolling Origin* was partly introduced to circumvent this. The time frame for the *Rolling Origin* was set to two periods, meaning that 14 days passed between each forecasting instance. When two weeks elapse, the model expands the training for the same number of days and retrieves new information for which to build forecasts upon based on the new training set. This process continues until the entire time series of 1096 days ends. In order to make comparisons fair for both the *ES* model and the *Naïve* forecast the same *Rolling Origin* conditions are applied on them as well. This approach eliminated the issues introduced by iterative forecasting while also reducing overall errors and producing better predictions.

For the *Rolling Origin*, 26 forecasts were produced along the test set for each forecasted category and method. Each of the 26 forecasts spans 14 days, apart from the 26$^{th}$ forecast, which is capped by the number of days remaining in the test set. Each forecast produces error values for each of the 14 days. The error values are compiled into the error measures we have chosen. Furthermore, each error value for each of the 26 forecasts is also compiled into a final average error value, making the error values robust towards extreme outliers. The waste and sales variables were converted into a logarithmic scale before fitting the model and running the *Rolling Origin*. This procedure avoids exponential growth and instead allows for linear growth, which stabilizes the variance of the data. Before producing the residuals, the variables were converted back into original units and in doing so all residuals are produced in normal scale.

After *Rolling Origin* was complete, the error values *MAE*, *RMSE* and *ME* were produced for all models, following studies like Aye et al. (2015) and Chu & Zhang (2003). Different conclusions about the accuracy of the forecast and the implication of the model's construction were then made based on how the models performed compared to one another.

# 4 Results

First, a compiled result for all the forecasts will be presented and after that the results will be divided up per aggregated level. Lastly, the result of the explanatory variables' estimates will be presented. In total, 37 forecasts were conducted. Table 6 below gives an overview of each forecast. The forecast ID contains the aggregated level that the forecast is conducted on as well as a number to identify for each forecast on each aggregated level. In Table A 3 and Table A 4 in the appendix, the underlying data for the A1.1 and the *A3* forecasts is shown, respectively. Table 3 shows summarizing data for the *A4* and *A5* forecasts.

*Table 6 All forecasts conducted*

| Forecast ID | Description | Belonging category | Forecast ID | Description | Belonging category |
|---|---|---|---|---|---|
| A1.1 | Grilled chicken | Meat | A3.19 | Cottage cheese natural | Cheese |
| A3.1 | Grilled chicken | Meat | A3.20 | Imported pork | Meat |
| A3.2 | Swedish chicken | Meat | A3.21 | Dessert mold cheese | Cheese |
| A3.3 | Fresh fish manual | Fish | A3.22 | Milk low fat | Dairy |
| A3.4 | Swedish beef central packaged | Meat | A3.23 | N/A | Dairy |
| A3.5 | Fresh shell fish | Fish | A3.24 | Cream | Dairy |
| A3.6 | Fresh fish packaged | Fish | A3.25 | Egg from free-range hens | Meat |
| A3.7 | Swedish beef | Meat | A3.26 | NFC Drink later | Dairy |
| A3.8 | Fresh Swedish CPK | Meat | A3.27 | NFC Drink now | Dairy |
| A3.9 | Sour milk high fat | Dairy | A3.28 | Brine products | Meat |
| A3.10 | Swedish pork | Meat | A3.29 | Mozzarella | Cheese |
| A3.11 | Smoked/cured fish | Fish | A3.30 | Sour milk low fat | Dairy |
| A3.12 | Fresh hamburger meat | Meat | A4.1 | Category Meat | |
| A3.13 | Swedish pork central packaged | Meat | A4.2 | Category Dairy | |
| A3.14 | Milk high fat | Dairy | A4.3 | Category Fish | |
| A3.15 | Big package flavored | Dairy | A4.4 | Category Cheese | |
| A3.16 | Fresh Swedish | Meat | A4.5 | Category Vegetarian | |
| A3.17 | Cooled drink to go | Dairy | A5.1 | All categories | |
| A3.18 | Milk medium fat | Dairy | | | |

The best performing model within its method is shown, based on the AIC residual. The actual AIC residual is not shown because, as already mentioned, its comparative power between methods is relatively bad, and the result is mainly dedicated to compare between methods, and not models within methods. The results for the *ME* metric are more indicative of overall bias to predict overtly positive or negative values rather than having a forecast that fits well with the actual out of sample values. Therefore, this residual will not be touched upon in the results but displayed for the purpose of transparency.

## 4.1 Result for all forecasts

In Table 7, the distribution of which forecast that had the best performing residual can be seen for all 37 forecasted categories. The *Naïve* forecast has the worst residuals in most cases, which indicates that it is quite possible and fruitful to conduct more efficient forecasts on the waste data. Further, performing better than the *Naïve* proves that there is structure in the waste data, meaning that the data possess either seasonal or trended structure. This also indicates that the other methods used are predicting future values more efficiently. Table 7 further displays the mean value for each error metric for the different methods chosen in order to give a sense of which one on average performs the best. Those with the best performance are marked in bold. Again, the values show that the *Naïve* never performs the best overall, meaning that the chosen forecasting models works.

For the *MAE metric*, the *Regression* is highly dominant. The only other method that performs best in some instances is the *Combination* model. *MAE* is highly robust towards outliers, which is something that the *Regression* is most proficient at explaining, hence its dominant performance for that residual. The prevalence of the *Combination* outperforming the other methods, however, suggest that there are events that the *Regression* cannot explain with its current set of variables. Furthermore, as can be seen in Table 7, the *Combination*'s performance is very close to the *Regression's* performance, meaning that whilst it may not outperform as often, it is not far off in its predictability. There are many reasons as to why this may occur. The *Regression* models are fairly simple, and mainly carry constructed explanatory variables - apart from the promotion variables and the waste and sales lags. As such, it lacks information regarding seasonality which can be important to help explain some of the waste and improve model performance. The *ES*, which captures such effects automatically, complements the *Combination* model with factors that the *Regression* cannot explain.

For the *RMSE* residual, the *ES* model performs the best 23 times, followed by the *Combination* which performed best 11 times. However, in terms of the average residual value for the *RMSE*, the *Combination* model actually attains the lowest value, although admittedly by a small margin. Given the circumstances of the data, and the simplicity of the *Regression* model, the results are not that surprising. The goal of the forecasting was to prove that the variables linked to certain events and effects would impact waste, not to make the most accurate forecast.

*Table 7 Compiled result of all the 37 forecasts*

| **Distribution of which error value performed best between the four models, compiled for all 37 forecasts** | | | | | **The average of each error value for all 37 forecasts** | | | |
|---|---|---|---|---|---|---|---|---|
| | **Regression** | **ES** | **Naive** | **Combination** | **Regression** | **ES** | **Naive** | **Combination** |
| **MAE** | **29** | 0 | 0 | 8 | **0.990378** | 1.230459 | 1.564865 | 1.060432 |
| **RMSE** | 2 | **23** | 0 | 12 | 2.354892 | 2.239622 | 2.932243 | **2.235000** |
| **ME** | 1 | **22** | 7 | 7 | 0.789189 | **-0.04883** | -0.19818 | 0.370135 |

*Table 8 Different type of cases for all the 37 forecasts**

| Type of Case | Explanation | Count of occurrences |
|---|---|---|
| *Case 1* | Regression performs best on MAE, ES performs best on RMSE, ES performs best on ME | 14 |
| *Case 2* | Regression performs best on MAE, ES performs best on RMSE, Naive performs best on ME | 2 |
| *Case 3* | Combination performs best on MAE, ES performs best on RMSE, ES performs best on ME | 6 |
| *Case 4* | Regression performs best on MAE, ES performs best on RMSE, Combination performs best on ME | 1 |
| *Case 5* | Combination performs best on MAE, RMSE and ME | 2 |
| *Solid Forecast* | Regression performs best on MAE, Combination or Regression performs best on RMSE; any model performs best on ME | 12 |

*Here we display/explain which models performed best based on what metric and overall how the different methods perform based on our errors.

Based on how the forecasts performed on the error values (residuals), different *Cases* were identified, as can be seen in Table 8. Forecasts where the *Regression* was performing best, together with the *Combination* model for the *RMSE* and *MAE* residuals, are considered as *Solid Forecasts*, as it means that the explanatory variables of the *Regression* help improve the predictive power of the forecast. For the *Solid Forecasts*, the explanatory variables can be said to make a significant impact, since the variables help improve the *Regression* models' predictive power. The other *Cases,* where the *Regression* is not a top performer for *RMSE* or *MAE*, are considered lost causes, with the exception of *Case 5*, wherein the *Combination* outperforms all others. In *Case 5*, many of the explanatory variables can still help explain waste effects and are therefore still counted as valid, even if it must be done more cautiously. In *Case 1, 2* and *4*, the *Regression* outperforms on the *MAE* residual, while the *ES* model outperforms on the *RMSE* residual. For these cases, assumptions can still be made about the explanatory variables, but they should be considered less valid. The reasoning behind this is that they help explain outliers, although their explanatory power overall is less significant. It is likely that adding more variables would further improve the *Regression* model's explanatory power. Therefore, the explanatory variables present still have merit.

## 4.2 Results per aggregated level

The performance of the forecasts per aggregated level is presented in order to identify if the results vary depending on the aggregation of the data. First are the results of the *A5*, followed by *A4*, *A3* and lastly *A1*. What needs to be considered is that the number of forecasts differ for each aggregated level, where levels *A5* and *A1* only have one forecast each.

### 4.2.1 Result for the A5 forecast

The A5.1 forecast is considered a *Solid Forecast* since the *Combination* performed best on all the residuals. The *ES MNA* model also indicates that there are multiplicative errors, meaning that there is a lot of *noise* present in the data. The noise is non structured errors which the model cannot explain, and likely accumulate with the higher levels of aggregation, as the data contains many more categories that are affected by different external events While the *Regression* model performed worse than the *ES* model on the *MAE* residual, the *Combination* outperformed both of them. The *Combination* manages to leverage the explanatory power of the *Regression* model's variables and the seasonality from the *ES* model in order to improve the overall predictive power. All forecasts on *A5* perform relatively closely on the residuals apart from the *Naïve*, something which indicates that the *Regression* is not far off in terms of predictive power. Due to the *Regression* model's relatively simple construction, it is safe to assume that more efficient explanatory variables would have improved the *Regression* model's predictive performance.

*Table 9 Result for the A5.1 forecast*

| **A5.1** | **All categories** | | | |
|---|---|---|---|---|
| | **Regression** | **ES(M,N,A)** | **Naive** | **Combination** |
| **MAE** | 0.449 | 0.432 | 0.615 | **0.427** |
| **RMSE** | 0.579 | 0.547 | 0.797 | **0.546** |
| **ME** | 0.097 | -0.04 | -0.226 | **0.029** |

### 4.2.2 Result for the A4 forecasts

The five forecasts for level *A4* differ in their results, with two being *Solid Forecasts*, two *Case 3* and one *Case 6*. As can be seen in Table 10, the *Combination* performs the best on *MAE* with the *Regression* coming second, whilst the *ES* method is better on both *RMSE* and *ME*. The fact that the *Regression* is losing its explanatory power compared to *A3* can be linked to the higher aggregation and hierarchical level. Because more observations than at the lower levels are being forecasted at once, there is more seasonal structure, which the *Regression* deals with less proficiently than the *ES*. Still, the fact that the *Combination* performs better suggests that there are factors that the *Regressions* variables help deal with. Much like *A5*, the *Combination* uses aspects of both methods to perform better. As can be seen for the A4.3 forecast in Table A 6, the reason why *ES* has lower residual values on the *RMSE* is mostly due to the *Combination* being significantly worse at predicting the Fish category, while otherwise performing better on this measure overall as well. The two *Solid Forecasts* further indicate that the *Regression* variables still play a large role in the causes of food waste at this aggregated level. The *Solid Forecasts* come from the categories Refrigerated Vegetarian products and Dairy, suggesting that these are less bound to seasonality and overall structure, alternatively it could be due to these categories being more affected by outliers. Still, it is clear that the structural/seasonal effects play a large role as evidenced by the fact that the *Combination* and *ES* perform best overall.

In terms of seasonality, the *ES* gives us a good indication of how the more aggregated data is structured. Four *ANA* (seasonality present) models and one *ANN* (no seasonality) model are produced by the *ES* method, with Cheese being the only one that performs better when not accounting for seasonality. This phenomenon likely caused by cheese itself not having as much seasonality as the other categories, since consumption always remains fairly stable year-round, along with the fact that cheese products possess a comparatively long shelf life. Still, the model selection is indicative that there are many seasonal/structural factors at play.

*Table 10 Compiled results for A4.1 to A4.5 forecasts*

| **Count of which error value performed best between the four models, compiled for forecasts A4.1-A4.5** | | | | | **The average of each error value for forecasts A4.1-A4.5** | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Regression** | **ES** | **Naive** | **Combination** | | **Regression** | **ES** | **Naive** | **Combination** |
| **MAE** | 2 | 0 | 0 | **3** | **MAE** | 0.782200 | 0.834800 | 1.147800 | **0.755800** |
| **RMSE** | 0 | 2 | 0 | **3** | **RMSE** | 1.322600 | **1.205200** | 1.624200 | 1.212800 |
| **ME** | 0 | **4** | 0 | 1 | **ME** | 0.540200 | **-0.05580** | -0.21600 | 0.242200 |

### 4.2.3 Result for the A3 forecasts

Level *A3*, which contains the most forecasts out of all hierarchies, naturally also contains a lot of different model combinations. As can be seen in Table 11, the *Regression* model performs best on *MAE*, both in terms of average residuals and frequency of best performance. The *ES* model performs the best in terms of frequency on *RMSE*, while the *Combination* model performs better in terms of average residual value. The *Solid Forecasts* A3.22, A3.21 and A3.8 are the only ones, including other aggregated levels, where the *Regression* performs best on the *RMSE* and *ME*, with the average residuals on *RMSE* being fairly close to the *ES* and the *Combination*. The collective values indicate that forecasting on this aggregated level is solid, with elements of the *Regression* being key in explaining waste. Moreover, *RMSE* displays an additional instance of the *Combination* model leveraging strengths from other developed forecast models, similar to level *A4* & *A5*.

*A3* contains nine *Solid Forecasts* in total, along with 14 *Case 1*, one *Case 4*, two *Case 2*, and four *Case 3* occurrences, for a grand total of 30. The *Solid Forecasts* are produced for the categories Dairy, Meat and Cheese, with six, two, and one being made for each category. Meat and Dairy are the categories with the largest amount of waste, which impacts how much these categories are included in *A3*, as waste in SEK formed the basis of selection. This then could lead to them being overrepresented as solid forecasts. Nevertheless, the results may be indicative of these categories being impacted by the explanatory variables in the *Regression* model, rather than systematic and structured seasonality.

The *ES* models for *A3* vary between *ANA* and *ANN,* with 16 and 14 incidents respectively. This suggests that there is a split between the different types of product, with some having strong seasonal effects that explain waste while others are more tied to special events or actions. There is no clear pattern to the division, as the different categories all have instances of seasonality being and not being present. Seasonality was not considered as a factor when selecting the different subcategories that make up *A3*. The model selection would suggest that for the selected categories, which were chosen for their high amount of waste, there is little pattern in terms of seasonality, even if there is some other structure in the data that makes the *ES* perform better than the *Naïve*.

*Table 11 Compiled results for A3.1 to A3.30 forecasts*

| Count of which error value performed best between the four models, compiled for forecasts A3.1-A3.30 | | | | | The average of each error value for forecasts A3.1-A3.30 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Regression** | **ES** | **Naive** | **Combination** | | **Regression** | **ES** | **Naive** | **Combination** |
| **MAE** | **26** | 0 | 0 | 4 | **MAE** | **1.045133** | 1.323933 | 1.668367 | 1.133433 |
| **RMSE** | 4 | **21** | 0 | **5** | **RMSE** | 2.600433 | 2.481500 | 3.242233 | **2.475000** |
| **ME** | 1 | **18** | 6 | 5 | **ME** | 0.855767 | **-0.04810** | -0.19390 | 0.403767 |

### 4.2.4 Result for the A1 forecast

The results for the *A1* level are seen in Table 12, and shows a *Solid Forecast* for the chosen article, as defined previously. While the *Regression* performs worse than the *ES* on *RMSE*, the *Combination* proves that values from the *Regression* are still good at reducing residuals. As level *A1* contains only one SKU, due to complications described in section 3.2.2, the results from the forecasts are somewhat inconclusive. Nevertheless, the *Solid Forecast* functions as an indicator of how the *Regression* helps explain the waste for the food article that contained the most waste data in the entire dataset. The *ES* model *ANA* also suggests that there is evidence of seasonality for the SKU.

*Table 12 Results for the A1.1 forecast*

| **A1.1** | **Grilled chicken, 1040G** | | | |
|---|---|---|---|---|
| | **Regression** | **ES(A,N,A)** | **Naive** | **Combination** |
| **MAE** | **0.93** | 1.20 | 1.50 | 1.03 |
| **RMSE** | 1.926 | 1.848 | 2.308 | **1.835** |
| **ME** | 0.729 | **-0.045** | -0.21 | 0.342 |

## 4.3 The *Regression* model's explanatory variables

Section 4.2.1 to 4.2.4 suggests that the *Regression* model does have explanatory power to some extent. Hence, it was deemed of interest to explore the Regression's explanatory variables. The most interesting variables to look at are those from the *Solid Forecasts* as well as *Case 5* since the models' performance on out of sample data proves that the variables used can help explain and predict food waste accurately. These forecasts sum up to 14 in total. Therefore, results from, and eventual assumptions made based on these variables, are more robust than those where the *Regression* does not perform as well. The paragraphs below are referencing to the results in Table A 2.

All variables which had an average significance above 0.01 (Pr > 0.01) were removed in order to get statistically reliable results. Furthermore, all the lag variables of waste were removed due to the fact that these variables were dependent on the actual waste data, making it hard to interpret the result of these variables. Given how variables are defined or dependent errors in the database, the below results are most likely only indicative, and should not be interpreted as exact, but they give some reference as to what impact the variables have. The percentage levels are included in order to give a comprehensive picture of their impact. Although some of the estimates are too negative to convert into accurate percentage points, their size still indicates to what extent they affect waste.

The *A1* forecast is a *Solid Forecast*. *Central* campaigns and *AugD3* (August week 3) are the most significant and strongest variables for this category. As the results are limited to one type of SKU, the results are somewhat inconclusive. Regardless, the two variables are still indicative that common effects such as promotions and summers also impact single products. The exact amount the variables impact grilled chicken is difficult to determine but both effects are shown to strongly reduce the amount of waste. *Sales lags 2 & 6* are also present, with lag 2 giving an increase in waste by approximately 50%, and lag 6 reducing waste by approximately 60%. This is likely due to the fact that sales figures are directly tied to waste, with more sold goods leading to less waste and vice versa. Therefore, historic sales would naturally lead to different effects.

The *A3* level consists of nine *Solid Forecasts*. The most significant variables are *Sales lag 5* along with the *MDt2 (two days after Midsummer)*. Both of these have highly negative estimate values, meaning that they reduce waste. Furthermore, *Normal* and *Central* has the most negative estimates, indicating that an increase in central campaigns and normal sales reduces waste. To what exact extent is difficult to say, but the size of the estimators suggests that both variables have a sizable effect. The holiday variables most often have positive estimates which indicates that waste is increased by these events. Negative estimates are dominant for the days after Christmas such as *JDt1 & JDt2* meaning that these days lead to less waste. Meanwhile *JDt3* has highly positive values, leading to a rise of circa 280%. In fact, *JDt3* is one of the most positive values, along with *MDt1* which suggests an increase of 380%. The estimates of the sales lags are largely negative, indicating that an increase in sales has a decreasing effect of waste.

The *A4* level contains two *Solid Forecasts* and one *Case 5* for Meat. This entails that the *Regression* did not perform with the highest accuracy on the *Case 5*, but some of its explanatory power are inherited in the *Combination*. The most significant estimates are *Sales lag 1* along with *CDF*. *CDF* is positive and indicates an increase of 55 %, which suggests that Christmas leads to more waste to some extent. This is presumably related to customers substituting certain products for others and bulking up on products for the holiday. *Sales lag 1* on the other hand is highly negative, along with the day after Christmas, so much so that their percental values cannot be established. Other sales lags give varying levels of either positive or negative estimates, while yS (sales) overall contribute approximately to 80% more waste given that the sales increase by 100%. This again is due to how waste is tied to the amount of sales for a given period. The variables with the highest positive estimates are *MDt_1* and the first week of July (*JuliD1*), with the third (*JuliD3*) and fourth (*JuliD4*) week also being positive. The variables give an increase of roughly 139.9%, 115%, 70%, and 64% respectively when these periods occur.

The *A5* forecast is a *Case 5*. Because of this, the *Regression* is not entirely trustworthy, as the other methods perform better. Still, the *Combination* carries values from the *Regression* in order to produce a more accurate forecast. This means that the variables and their impact are still relevant, even if the results are not as reliable. The *lag 5* variable of sales have a positive estimate. The sales lag variable gives an indication that a 100% increase of sales 5 days prior leads to a 25% decrease in waste in the future. But the sales variable, *yS*, has a positive estimate, which entails an indication that a 100% increase in sales will lead to 28% increase in waste the same day. This effect is present at levels *A4* and *A3* (with different estimates) as well but in those cases, it is largely overshadowed by other more dominant variables. The holiday variables are largely positive in the case of *MDF* and *JDt2*, with *AugD2* being an exception. *JDt2* is most significant amongst all the variables of the *A5* forecast and its estimates indicate that waste increases by 80% two days after Christmas. *AugD2* decreases waste by 32% and could be explained by the fact that the Swedish summer holiday comes to an end at roughly the same period, which increases sales and leads to less waste. Contrastive to this is *MDF,* which increases overall waste by approximately 33%. This holiday occurs when the Swedish summer vacation usually begins, which decreases sales and therefore could lead to more waste. It should once again be noted that these effects are more indicative than they are precise. Even so, the estimates of the variables give insight into the fact that the explanatory variables identify large outliers present within the dataset.

# 5 Discussion

Tying back to the opening statement of this thesis, our purpose is to evaluate how food waste could be forecasted and how food waste could be explained, as well as to provide indications for how retailers can use forecasting in order to reduce food waste. As stated in the introduction, there is little research regarding the forecasting of food waste. However, some key findings from related literature do exist, such as Van Donselaar et al. (2016), which investigates saturation levels and substitution effects for promotions on perishable products, and Arunraj & Ahrens (2015a) who developed and applied two models to forecast the daily sales of bananas. In this section, we will elaborate on the results of this study. This is followed by a discussion of the methodology, along with implications for practitioners. Lastly, we will touch upon the research obstacles encountered during the writing of the thesis.

## 5.1 Discussion of the results

With regard to the forecasting aspect, it is clear that fairly simple modeling still has strong predictive capabilities, based on how the models outperforms the *Naïve*. Much has been made of the difficulties of forecasting food waste and how the seemingly erratic nature of the waste process makes forecasting seem somewhat useless. However, as evidenced by the fact that the *Naïve* never achieves the best average values on any level, there clearly is structure in the food waste data. Furthermore, the *ES* method likewise suggests the presence of structured seasonal patterns for a large share of the categories, especially on the higher hierarchical levels. Therefore, it is safe to assume that for a substantial part of all the food waste, there are definite patterns that can be countered, thereby diminishing food waste.

However, since the *Regression* model is often performing best for the *MAE*, it suggests that there are still many outliers that cannot be explained simply through structure by methods such as the *ES*. The *Regression* is a more advanced method of forecasting compared to the *ES*, since it takes more factors into account. As stated many times before, the goal of this thesis is not to produce the most accurate forecast possible, but instead to provide insight into how to forecast food waste, and to investigate if explanatory variables can explain the food waste. In order to provide a model with higher predictive power, more advanced forms of regression, more variables or other methods like *ARIMA* would be required. However, the *Regression* model's explanatory variables provide guidance in determining what actions can be taken, besides taking the seasonal effects into account when determining orders and sales.

Again, we see evidence of seasonal and error effects that are captured by the *ES*, which are important to the performance of the forecast. However, the fact that the *Combination* attains the lowest residuals in many of the forecasts indicates that the explanatory power of *ES* and regression is needed. This is however not surprising, as Arunraj & Ahrens (2015b), along with Ord et al. (2017), states that hybrid models do have stronger predictive power compared to single models. Additionally, for the forecasts where the *Regression* did not perform the best, it still performed relatively close to the other models' residuals. This fact, in combination with how the *Regression* performs on the *MAE*, suggests that the *Regression* model's variables have explanatory power. While ultimately a more sophisticated form of regression with a greater number of, and more efficient explanatory variables, would perform better than the *Combination* model, as the structural effects from the *ES* would no longer be missing. Currently however, the *Combination* proves that both seasonality and outliers are needed in order to most accurately forecast food waste.

Regarding the variables in the Regression models for the Solid Forecasts and the Case 5 forecasts, there are certain variables that are common on multiple levels. Variables such as Christmas (JDt2) along with Sales (yS) appear often, are highly significant, and are therefore closely tied to waste. Furthermore, the sales lag variables show that an increase in sales will most likely lead to a decrease of waste in the future. On the other hand, a majority of the holiday variables indicate that waste is increased due to these events. These results are not surprising, as holiday events creates larger fluctuations in demand, making it harder for demand forecasts to accurately predict the order numbers. The same negative impact derived from holidays may not necessarily be present in other regions, where other consumer patterns are present. Therefore, such effects need to be explored on a case by case basis in order to establish a pattern. The same is true for the summer vacation period, which most leads to a decrease in demand in the three stores located in Stockholm, as can be seen in Figure 2, while other areas of Sweden could potentially experience an upturn due to an influx of customers during the same period.

One of the more important observations comes in the form of *Central* campaigns. For single categories, the results suggest that the promotions reduce waste for the product, as evidenced by the *A1* forecast. When moving to higher levels of aggregation, however, it appears as if the central promotions actually drive waste higher instead. This is most likely due to the substitution effects and cannibalization caused by promotions. While promotions may lead to a reduction in waste for a single item, on an overall level, the campaigns increase the total waste level, potentially due to decreased demand of non-promoted products. Due to the limited number of forecasts and data available for the different levels, such conclusions have to be explored further in order to be fully validated. Still, the fact that such effects are visible in this study, and at such a high level of significance, indicate that they are highly likely to reappear for other products.

It is safe to assume that some of the explanatory variables used in the *Regression* model had a consistently enhancing effects on the forecasts' performances, and it can be derived whether a variable's impact is decreasing or increasing the level of food waste. Therefore, it can be concluded that there does exist certain variables that explain food waste and by implementing these in retail organizations' forecasts would improve the performances.

## 5.2 Implications for practitioners

There are different implications for the different hierarchical levels, but also a number of commonalities. On every level, effects relating to above all sales, lags and the weeks of summer are both prevalent and significant. Results in *A1* are somewhat inconclusive since they rely on a single forecasted article, which cannot be said to be representative of the entire dataset.

Tying back to section 5.1, *Central* campaigns seem to be a variable that affects the lower aggregated levels (*A1, A3*) more substantially. Therefore, a recommendation when managing waste in *A1* and *A3* is to try to manage their sales campaigns more in line with the waste forecasts, and match these with forecasts made for demand. For single products, central campaigns are an effective means of reducing the economic impact of waste. However, as can be seen in the results for higher aggregated levels, central campaigns seemingly increase waste. Price elasticity of demand and customers' substitution patterns need to be taken into account. Although adjusting order numbers in regard to the cross-products effects of promotions is a complex problem, it is an effective mean of combating waste and reducing cost of goods. Relationships between products and to what extent the promotion of one product leads to waste later on for that particular product, or for other products, should be a key concern for retailers. This is applicable for *A4* and *A5* to some extent as well, but since it is easier to manage campaigns on specific products or subcategories it is advisable to focus promotion management on the lower aggregated levels.

Prevalent for all aggregated levels is the impact of sales and lags, which are highly significant. As such, matching the demand forecast with the waste forecast and improving the waste forecast, is another recommendation to reduce the waste level, and by extension reduce cost and increase revenue. To combat waste more efficiently, improving the coordination between promotional activities and order numbers, along with joint forecasting of demand and waste, must become integrated parts of any retailer's core strategy. These recommendations are also supported by Thyberg & Tonjes (2016) and Buzby & Hyman (2012), which state that poor forecasting is a leading cause of food waste. Both holiday variables during the summer and Christmas are recurring themes for explaining waste. For the summer weeks, where sales decreases and waste increase, potentially due to less demand[5], it becomes important to match demand forecasting with waste forecasting in order to foresee decreases in demand and increases in waste. While the demand forecasts might mitigate overstocking to some extent, the fact that the models still find the holiday variables significant, and that the data displays clear spikes in waste, indicates that there are still unresolved issues and unnecessary losses in revenue. These issues could be moderated by more efficient forecasting approaches, such as joint forecasting of the demand forecast and the waste forecast.

## 5.3 Research limitations

Not all categories were chosen on the less aggregated levels (*A1, A2, A3*), and the selection methodology used to choose the products and categories was based on those which had the highest amount of waste in SEK. Therefore, the sample can be somewhat skewed toward products/categories with a high amount of waste in SEK. However, from a financial point of view, reducing the food waste for these products/categories generates the greatest financial impact for the retailer. This touches upon the issue that Halloran et al. (2014) discusses, namely that there are few incentives for retailers to really address the issue of food waste. In order to create incentives for retailers to use forecasts to decrease food waste, it was deemed most efficient to use this selection methodology. Due to the selection methodology chosen, Meat was the only category present in *A1,* Meat and Dairy are the most numerous in *A3*, and Refrigerated

[5] This is the case for the stores surveyed in the dataset, which are all located in Stockholm and my not be as accurate for other regions

Vegetarian product only appears in *A4*. Because of this skewness, it was decided to not analyze the differences between the categories and instead focus on the differences between the aggregated levels.

Furthermore, differences between stores could have been explored, as it could have provided insights into any locational differences and possibly be used as an indicator of store performance. However, in order to gather more data for the different hierarchical levels, particularly the lower ones which often contained insufficient number of observations, it was deemed better to compile the stores' observations into a larger data set. The reason for this was both computational as well as time related constraints, as part of reducing the scope of the thesis. Another limitation with the methodology selection was the possibility of observing the seasonal patterns in the sales and waste data mentioned section 3.1.4. This is due to our choice of having daily observations and a forecast horizon of 7 days. However, the holiday variables in the regression model were utilized capture some of the seasonality of the waste data. Utilizing longer horizons could be a way to capture other types of seasonal patterns but was not feasible for this study given the limited amount of historical data.

A difficulty that became clear early on in the process of forecasting was issues with data and data structuring. Waste data is inherently messy, since there are many human factors involved in producing the actual data points (Lebersorger & Schneider, 2014a). As mentioned in section 3.1.2, there are limitations with the data, mostly related to the high risk that waste routines are not followed accurately. First of all, more robust conclusions could likely be drawn if level *A1* would have been more expansive, with more forecasts having been produced. However, as previously discussed, due to issues related to the amount of observations, this was difficult to do. There is a risk that wrongly logged waste data skews the result. Although extreme values could be removed, such as those mentioned in section 3.1.3, other less extreme, but still incorrect values, are still hard to identify. Instances of negative sales and waste were also found in numerous spots within the data set, which further complicated the integrity of the data and diminished the amount of usable observations for the less aggregated hierarchies. While such observations occur naturally through clerical errors or misunderstandings and were dealt with by aggregating the data, they nevertheless complicate the forecasting process.

Other store factors that impact the structure of the data stem from the operational nature of food waste retailing, where waste can be compiled from several dates and are logged as one, or the risk that *TBA*:s diminishes the actual food waste derived from a store (Eriksson et al., 2017). Such actions further skew the sample data. While such data points reflect a more real picture of how the stores operate and log waste, it does not give fully accurate picture of when items are actually expired. A qualifier for producing more accurate forecasts and reducing food waste is well structured data that originates from well managed waste routines. There are many ways to achieve this, but diving deeper into the management of food routines is beyond the scope of this thesis. However, we would like to highlight that mitigating human errors and having uniform and well managed waste routines across all stores of a chain would be highly effective measures in combating messy data and improving the results of the forecasts on food waste data. At the very least, it would most certainly give a more realistic picture of the monetary value of the food that is being thrown away.

It was desirable to use Axfood's demand forecast as a variable in the *Regression* model, however it was not practically feasible to export that data. Hence, it was not used in any forecast. The demand forecast of a product potentially has several interesting uses in explaining food waste. First of all, it is likely that this variable could have a high correlation with food waste because it is the basis of the number of units ordered. An overoptimistic demand forecast would therefore give rise to larger amounts of waste. Secondly, by using the food waste forecast as a moderator for a demand forecast, it could help identify and reduce the over ordering of products. Furthermore, incorporating the products shelf lives could give an even more accurate picture of the food waste caused by a particular order, a variable which Van Donselaar et al. (2016) utilizes. Unfortunately, such information was not made available, and would also not be fully utilized without the access to the actual order data or demand forecast to derive a wasted product to a particular order.

Other residuals, such as R-squared or adjusted R-squared, were considered for the purposes of model evaluation within methods. However, we chose not to use these, due to the AIC residual being better for our purposes. The AIC allows us to measure the models' out of sample fit, which in general is the best forecasting practice, as it is future data points (out of sample data) that is of interest to explain, and not historical data points (the in-sample data). Critique of the AIC can be found in Beier et al. (2001) and Ma et al. (2016), which discusses the potential of shrinkage estimators, such as least absolute shrinkage and selection operators (LASSO), possess. Shrinkage estimators are an alternative approach to effectively select variables, when a greater number of variables are used in a model, which was not deemed to be the case for our models. These estimators have constraints, as they are difficult to interpret and compute, and lack generality; which makes some of them only applicable on regression models. Meanwhile, the AIC is applicable on other methods, apart from regressions. For our purposes, it was therefore deemed more fitting to utilize the AIC, as little added value could be seen for choosing shrinkage estimators or variants of R-squared.

# 6 Conclusion

(1) The first conclusion that can be drawn is that forecasting food waste is feasible with more advanced methods than the *Naïve*, such as *ES* and *Stepwise Regression*. (2) Secondly, food waste data does have structure in it, such as seasonality, which can be explained with forecasting methods. This is indicated by the *ES* model's residual performances, where evidence of seasonal patterns is present for several of the categories used for forecasting. Additionally, it is possible to leverage the knowledge of seasonality to reduce food waste through utilizing proactive measures. (3) Thirdly, it is possible to enhance the performance of the forecasts by incorporating explanatory variables, which help explain outliers, as indicated by the *Regression's* overly dominant performance on the *MAE* residual. (4) Furthermore, the performance of the *Combination* model's residuals proves that explaining both structure and outliers is necessary in order to further improve the forecasts. (5) In addition, it can be concluded that certain explanatory variables have more significant correlations with food waste compared to others, and that the significance as well as which variables that are significant differ between aggregated levels. Several of them, such as the *Central* campaigns, the sales lags, holidays such as Christmas and Midsummer, along with the summer weeks of July and August have a significant economic impact on several aggregated levels. (6) Lastly, it can be identified whether the variables have an increasing or decreasing effect on waste.

By combating the effects that affect waste through statistical forecasting methods, retailers could with a high probability reduce their overall waste. This is advantageous from an economic standpoint, as the retailer can increase their revenues, as the cost of goods reduces (COGS). This will also decrease the indirect costs that arise from COGS, such as transportation and storage costs, as well as having employees spend less time managing food waste. Using these insights in forecasting is also helpful in reaching goals set by the UN, and grocery retailers like Axfood, Coop and ICA. Food waste is an important issue that requires investments in better forecasting in order to combat it. By adopting a joint food waste/demand forecasting methodology and using the insight gained in this study, such as focusing on specific explanatory variables, it is possible to boost profits and lessen environmental impact from retail operations.

Our research contributions regarding forecasting food waste data are plentiful. Foremost, we prove that forecasting food waste is feasible. By our explorative approach to the research question, we contribute to different aspects of what forecasting food waste entails. Findings, such as certain variables that can explain food waste, and the different implications forecasts have on different aggregated levels, can be further expanded on in future research. Additionally, insights into what actions can be taken in order to further enhance model performance, are areas that can be built upon by future researchers.

# 7 Future Research

Our research has explored how to forecast food waste and is likely one of the first research attempts at applying forecasting methods on food waste data in order to combat the problem in retail stores. There is much more that can and should be investigated using scientific methods in this area. To begin, with interviews with the store managers could have been conducted in order to gain an understanding of how well waste routines are being followed. In our case, the reason for not interviewing the store managers was a sheer lack of time. By having more qualitative data regarding the food waste, a more reliable conclusion can be made, since it would increase the likelihood that more of the actual food waste is captured in the data.

The *Stepwise Regression* suited the purposes of our particular research question, as it allowed us to develop many forecasts quickly with different explanatory variables present for different datasets. The use of *AIC* also allowed for combating multicollinearity, to some extent, which otherwise would have required more effort and specified models each time we imported a new dataset. This would have been both time consuming and ultimately, we cannot see how it would lead to much improvement if any. The *Regression's* simplicity in terms of number of variables and predictive power does cause some issues, however. While not the aim of this thesis, a more predictive forecast is, in general, a better one, unless overfitting occurs. Other model choices, such as regressions using an *autoregressive solution* or *difference values* instead, might have proved more accurate. Another way to combat the disadvantages of the *Stepwise Regression* model would have been to use another method altogether. Methods such as *ARIMA*, which Arunraj & Ahrens (2015a) and Arunraj & Ahrens (2015b) utilized, or *Artificial Neural Networks*, as Alon et al. (2001) applied on retail sales data, would likely have produced more accurate forecasts. These methods certainly come with their own issues, chiefly among which is definitely their complexity and the time it takes to produce accurate models. Given the time frame and scope of the study, using these methods would have been impractical. However, doing so is definitely a course of action to consider for both retailers and future researchers wishing to take part in or expand upon what is established in this thesis.

A potential future research area would be to build upon what Van Donselaar et al. (2016) established, by expanding the scope of our study to include price elasticity and the impact promotions have on consumer behavior. This study already proves that promotions have a significant impact on waste. Promotional strategies could be employed in a fashion that fits consumer expectations on prices and deals with perishability by using incremental sales techniques. Utilizing a more proactive approach to promotions could then theoretically dramatically decrease waste. Such assumptions could be tested via data and forecasting models and would definitely help reduce both the economic and environmental impact of food waste.

It should be noted that the *Regression* model does not utilize all possible explanatory variables of waste, which diminishes the potential explanatory power that this method possesses. Effects that would most likely impact food waste include humans factors, as suggested by Lebersorger & Schneider (2014a), such as waste management and human handling, and interaction effects between products, such as cannibalization caused by promotions on related products, as suggested by Van Donselaar et al. (2016). Such factors are both important, interesting and deserve to be explored, but can also be difficult to compute or construct. Arunraj & Ahrens (2015a) uses weather effects as a variable to explain retail sales, a factor which could also impact food waste, meriting further research. Furthermore, incorporating the demand forecasts and the product's shelf life in the food waste forecast, as well as incorporating the food waste forecast as a variable in the demand forecast in order to moderate over-ordering, are thought-provoking future areas of research.

Conducting forecasts on waste data, but on other dimensions such as from separate stores and even higher aggregated levels (more stores, distribution centers and entire food chains), are also areas for future research. In addition, by having data covering a longer time period, forecasts can be made on several time dimensions, such as weekly, quarterly, and yearly aggregated data. Durable products, such as rice and chocolate, are also subject to further research. Although the models used in this study could be directly applicable to durable goods, the nature of those goods are potentially different from non-durable goods. Therefore, using additional variables and more data could enhance the efficiency of the forecast for durable goods. Lastly, conducting forecasts by using units, instead of waste in monetary terms, is an alternative way of conducting the forecasts and could be necessary if a longer time span is used, since price's inflation could impact the result.

# Bibliography


Alon, I., Qi, M., & Sadowski, R. J. (2001). Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods. *Journal of Retailing and Consumer Services*, *8*(3), 147–156. https://doi.org/10.1016/s0969-6989(00)00011-4

Arunraj, N. S., & Ahrens, D. (2015a). A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting. *International Journal of Production Economics*, *170*, 321–335. https://doi.org/10.1016/j.ijpe.2015.09.039

Arunraj, N. S., & Ahrens, D. (2015b). A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting. *International Journal of Production Economics*, *170*, 321–335. https://doi.org/10.1016/j.ijpe.2015.09.039

Athanasopoulosa, G., & Kourentzes, N. (2020). On the evaluation of hierarchical forecasts. *Edinburgh Mathematical Notes*, *32*, xiii–xiv. https://doi.org/10.1017/s0950184300800014

Axfood. (2019). *Års- och hållbarhets redovisning*.

Aye, G. C., Balcilar, M., Gupta, R., & Majumdar, A. (2015). Forecasting aggregate retail sales: The case of South Africa. *International Journal of Production Economics*, *160*, 66–79. https://doi.org/10.1016/J.IJPE.2014.09.033

Beier, P., Burnham, K. P., & Anderson, D. R. (2001). Model Selection and Inference: A Practical Information-Theoretic Approach. In *The Journal of Wildlife Management* (Vol. 65, Issue 3). https://doi.org/10.2307/3803117

Brancoli, P., Rousta, K., & Bolton, K. (2017). Life cycle assessment of supermarket food waste. *Resources, Conservation and Recycling*, *118*, 39–46. https://doi.org/10.1016/j.resconrec.2016.11.024

Buzby, J. C., & Hyman, J. (2012). Total and per capita value of food loss in the United States. *Food Policy*, *37*(5), 561–570. https://doi.org/10.1016/j.foodpol.2012.06.002

Chu, C. W., & Zhang, G. P. (2003). A comparative study of linear and nonlinear models for aggregate retail sales forecasting. *International Journal of Production Economics*, *86*(3), 217–231. https://doi.org/10.1016/S0925-5273(03)00068-9

Cicatiello, C., & Franco, S. (2020). Disclosure and assessment of unrecorded food waste at retail stores. *Journal of Retailing and Consumer Services*. https://doi.org/10.1016/j.jretconser.2019.101932

Coop. (2019). *Årsredovisning 2019 Coop Nord*.

Davies, N., & Newbold, P. (1986). Statistics for Business and Economics. In *The Mathematical Gazette* (Vol. 70, Issue 453). https://doi.org/10.2307/3615710

Eriksson, M., Ghosh, R., Mattsson, L., & Ismatov, A. (2017). Take-back agreements in the perspective of food waste generation at the supplier-retailer interface. *Resources, Conservation and Recycling*. https://doi.org/10.1016/j.resconrec.2017.02.006

Evans, B. (2010). How to Interpret Regression Coefficients. *Agrekon*, *3*(34), 2–5.

Fildes, R., Ma, S., & Kolassa, S. (2019). Retail forecasting: Research and practice. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2019.06.004

Halloran, A., Clement, J., Kornum, N., Bucatariu, C., & Magid, J. (2014). Addressing food waste reduction in Denmark. *Food Policy*, *49*(P1), 294–301. https://doi.org/10.1016/j.foodpol.2014.09.005

HUI. (2018). *Dagligvarukartan 2018 ™, HUI Research*.

ICA. (2018). *Årsredovisning ICA 2018*.

Kourentzes, N., & Athanasopoulos, G. (2019a). Cross-temporal coherent forecasts for Australian tourism. *Annals of Tourism Research*, *75*(February), 393–409. https://doi.org/10.1016/j.annals.2019.02.001

Kourentzes, N., & Athanasopoulos, G. (2019b). Elucidate structure in intermittent demand series. *Monash University, Work Paper 27/19*.

Lebersorger, S., & Schneider, F. (2014a). Food loss rates at the food retail, influencing factors and reasons as a basis for waste prevention measures. *Waste Management*, *34*(11), 1911–1919. https://doi.org/10.1016/j.wasman.2014.06.013

Lebersorger, S., & Schneider, F. (2014b). Food loss rates at the food retail, influencing factors and reasons as a basis for waste prevention measures. *Waste Management*, *34*(11), 1911–1919. https://doi.org/10.1016/j.wasman.2014.06.013

Ma, S., Fildes, R., & Huang, T. (2016). Demand forecasting with high dimensional data: The case of SKU retail sales forecasting with intra- and inter-category promotional information. *European Journal of Operational Research*, *249*(1), 245–257. https://doi.org/10.1016/j.ejor.2015.08.029

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLOS ONE*, *13*(3), e0194889. https://doi.org/10.1371/journal.pone.0194889

Naturvårdsverket. (2020). *Matavfall i Sverige*.

Ord, K., Fildes, R., & Kourentzes, N. (2017). *Principles of Business Forecasting, 2e* (T. Burgard (Ed.); Second). Wessex Press Inc.

Parfitt, J., Barthel, M., & Macnaughton, S. (2010). *Food waste within food supply chains: quantification and potential for change to 2050*. https://doi.org/10.1098/rstb.2010.0126

Rob J. Hyndman, & Yeasmin Khandakar. (2008). Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software*, *27*(3), 22. http://www.jstatsoft.org/v27/i03/paper

Rosa, W. (2018). Transforming Our World: The 2030 Agenda for Sustainable Development. In *A New Era in Global Health*. Springer Publishing Company. https://doi.org/10.1891/9780826190123.ap02

Song, Q. (2015). Lessons learned and challenges encountered in retail sales forecast. *Industrial Engineering and Management Systems*, *14*(2), 196–209. https://doi.org/10.7232/iems.2015.14.2.196

Stenmarck, Å., Jensen, C., Quested, T., Moates, G., Buksti, M., Cseh, B., Juul, S., Parry, A., Politano, A., Redlingshofer, B., Scherhaufer, S., Silvennoinen, K., Soethoudt, H., Zübert, C., & Östergren, K. (2016). Estimates of European food waste levels. Reducing food waste through social innovation. In *Fusions*. https://www.eu-fusions.org/phocadownload/Publications/Estimates of European food waste levels.pdf%5Cnhttps://phys.org/news/2016-12-quarter-million-tonnes-food-logistics.html#nRlv

Teller, C., Holweg, C., Reiner, G., & Kotzab, H. (2018). Retail store operations and food waste. *Journal of Cleaner Production*, *185*, 981–997. https://doi.org/10.1016/j.jclepro.2018.02.280

*This Campaigner Wants to End the Global Food Waste Scandal - VICE*. (n.d.). Retrieved April 9, 2020, from https://www.vice.com/en_us/article/43q9mb/this-campaigner-wants-to-end-the-global-food-waste-scandal

Thyberg, K. L., & Tonjes, D. J. (2016). Drivers of food waste and their implications for sustainable policy development. In *Resources, Conservation and Recycling* (Vol. 106, pp.

110–123). Elsevier. https://doi.org/10.1016/j.resconrec.2015.11.016

Van Donselaar, K. H., Peters, J., De Jong, A., & Broekmeulen, R. A. C. M. (2016). Analysis and forecasting of demand during promotions for perishable items. *International Journal of Production Economics*. https://doi.org/10.1016/j.ijpe.2015.10.022

# Appendix

*Formula A 1 RMSE*

$$RMSE = \sqrt{MSE}$$

*Formula A 2 MSE*

$$MSE = \sum_{i=1}^{m} \frac{(Y_{t+i} - F_{t+i})^2}{m} = \sum_{i=1}^{m} e_{t+i}^2/m$$

*Formula A 3 MAE*

$$MAE = \frac{\sum_{i=1}^{m} |Y_{t+i} - F_{t+i}|}{m} = \frac{\sum_{i=1}^{m} |e_{t+i}|}{m}$$

*Formula A 4 AIC*

$$AIC = -2logL + 2q$$

*Formula A 5 ME*

$$ME = \frac{\sum_{i=1}^{m} (Y_{t+i} - F_{t+i})}{m} = \frac{\sum_{i=1}^{m} e_{t+i}}{m}$$

*Formula A 6 Naive forecast*

$$F_t = F_{t-1}$$

*Formula A 7 Combination model*

The *Combination* model (CM) that were used had a weight of 0.5 for the *Regression* model (R) and the Exponential Smoothing model (ES).

$$CM_t = 0.5R_{1t} + 0.5ES_{1t}$$

ETS:

Error: Additive (A), Multiplicative (M) * Note that None (N) does not exist for error, as there are always errors present in data.

Trend: None (N), Additive (A), Multiplicative (M), or Damped (D)

Seasonal: None (N), Additive (A), or Multiplicative.

With this framework, there are 12 possible models that can be produced. Only 3 were utilized in this paper, as determined by AIC and the ETS function. The basic form of ES model is Simple Exponential Smoothing, or SES. This forms the basis for all other model combinations.

**ANN – SES**

Equation:

$$L_{t+1} = L_t + \alpha(Y_{t+1} - L_t) = L_t + \alpha e_{t+1}$$

*Formula A 8 ES (A,N,N)*

$$\boxed{F_{t+1} = \alpha A_t + (1 - \alpha)F_t}$$

Table A 1 presents the variables used for the *Regression* model. There are 65 variables in total that are included in the *Regression* model. "t+1" indicates that the variable occurs t days after a holiday event. Whereas "t-1" indicates that the variable occurs before a holiday event. Lags are restructured variables of another variable, where each lag occurs x days after the original occurrence of a data point. For example, when conducting a forecast, the waste lag 3 variable occurs three days after the original waste data was logged in the system.

*Table A 1 Variables used in the Regression model*

| Variable name | Description | Variable name | Description | Variable name | Description |
|---|---|---|---|---|---|
| Central | Central Campaign | JDt_1 | Dummy For Christmas, t-1 | JuliD1 | Dummy For 1st Week in July |
| Local | Local Campaign | JDt_2 | Dummy For Christmas, t-2 | JuliD2 | Dummy For 2nd Week in July |
| Normal | Normal sales | JDt_3 | Dummy For Christmas, t-3 | JuliD3 | Dummy For 3rd Week in July |
| Week | Dummy For Weekly Seasonality | JDt_4 | Dummy For Christmas, t-4 | JuliD4 | Dummy For 4th Week in July |
| MDF | Dummy For Midsummer | JDt_5 | Dummy For Christmas, t-5 | AugD1 | Dummy For 1st Week in August |
| EDF | Dummy For The Easter Week | JDt_6 | Dummy For Christmas, t-6 | AugD2 | Dummy For 2nd Week in August |
| CDF | Dummy For The Christmas Week | JDt_7 | Dummy For Christmas, t-7 | AugD3 | Dummy For 3rd Week in August |
| EDt | Dummy For Easter, Mid-Week (t) | JDt1 | Dummy For Christmas, t+1 | AugD4 | Dummy For 4th Week in August |
| EDt_1 | Dummy For Easter, t-1 | JDt2 | Dummy For Christmas, t+2 | lag1 | Waste lag 1 |
| EDt_2 | Dummy For Easter, t-2 | JDt3 | Dummy For Christmas, t+3 | lag2 | Waste lag 2 |
| EDt_3 | Dummy For Easter, t-3 | JDt4 | Dummy For Christmas, t+4 | lag3 | Waste lag 3 |
| EDt_4 | Dummy For Easter, t-4 | NyåraftD | Dummy For New Year's Day | lag4 | Waste lag 4 |
| EDt1 | Dummy For Easter, t+1 | JanD1 | Dummy For January 1st | lag5 | Waste lag 5 |
| EDt2 | Dummy For Easter, t+2 | JanD2 | Dummy For January 2nd | lag6 | Waste lag 6 |
| <EDt3 | Dummy For Easter, t+3 | JanD3 | Dummy For January 3rd | Ys | Sales |
| EDt4 | Dummy For Easter, t+4 | JanD4 | Dummy For January 4th | Slag1 | Sales lag 1 |
| MDt | Dummy For Midsummer, Mid-Week (t) | JanD5 | Dummy For January 5th | Slag2 | Sales lag 2 |
| MDt_1 | Dummy For Midsummer, t-1 | JanD6 | Dummy For January 6th | Slag3 | Sales lag 3 |
| MDt_2 | Dummy For Midsummer, t-2 | JanD7 | Dummy For January 7th | Slag4 | Sales lag 4 |
| MDt1 | Dummy For Midsummer, t+1 | JanD8 | Dummy For January 8th | Slag5 | Sales lag 5 |
| MDt2 | Dummy For Midsummer, t+2 | JanD9 | Dummy For January 9th | Slag6 | Sales lag 6 |
| JDt | Dummy For Christmas, Mid-Week (t) | JanD10 | Dummy For January 10th | | |

*Table A 2 Estimate values of the significant variables for the Solid Forecasts and Case 5 forecasts*

| Aggregated level | Variable | Average of Estimate | Average of Correlation | Average of Pr(>\|t\|) | Number of Variables |
|---|---|---|---|---|---|
| A1 | AugD2 | -1.88998 | -0.11031 | 0.00101 | 1 |
| A1 | Central | -1.24918 | -0.06068 | 0.00004 | 1 |
| A1 | Slag2 | 0.56903 | 0.06875 | 0.00116 | 1 |
| A1 | Slag6 | -0.61328 | -0.0661 | 0.00004 | 1 |
| A3 | AugD1 | 0.98681 | 0.05723 | 0.00956 | 1 |
| A3 | CDF | -0.72258 | -0.07613 | 0.00903 | 1 |
| A3 | Central | -14.460695 | -0.04319 | 0.00042 | 2 |
| A3 | JanD1 | 2.90768 | 0.08471 | 0.00305 | 1 |
| A3 | JanD4 | 3.29731 | 0.07731 | 0.00157 | 1 |
| A3 | JanD5 | 3.64448 | 0.07687 | 0.00223 | 1 |
| A3 | JDt_4 | -0.10136 | 0.00767 | 0.00334 | 2 |
| A3 | JDt1 | -3.32982 | -0.00165 | 0.0095 | 1 |
| A3 | JDt2 | -3.54989 | 0.00854 | 0.0056 | 1 |
| A3 | JDt3 | 2.84498 | 0.06079 | 0.00523 | 1 |
| A3 | JuliD1 | 1.203246667 | 0.07956 | 0.00527 | 3 |
| A3 | JuliD2 | 1.04909 | 0.05353 | 0.00695 | 1 |
| A3 | JuliD3 | 1.30226 | 0.07427 | 0.00106 | 1 |
| A3 | JuliD4 | 1.22604 | 0.051 | 0.00382 | 1 |
| A3 | MDt1 | 3.83069 | 0.1026 | 0.00118 | 1 |
| A3 | MDt2 | -4.1766 | -0.01076 | 0.00002 | 1 |
| A3 | Normal | -9.114734 | -0.07564 | 0.00199 | 5 |
| A3 | Slag1 | -1.286653333 | -0.10443 | 0.00056 | 3 |
| A3 | Slag3 | -0.91586 | -0.0687 | 0.00028 | 1 |
| A3 | Slag4 | -1.25188 | -0.15704 | 0.00043 | 2 |
| A3 | Slag5 | -1.61457 | -0.13094 | 0.00009 | 1 |
| A3 | Slag6 | 0.8449133333 | 0.02811 | 0.00182 | 3 |
| A3 | yS | 1.07028 | 0.04987 | 0.00308 | 4 |
| A4 | CDF | 0.55357 | 0.08735 | 0.00006 | 1 |
| A4 | JDt2 | -1.73823 | 0.02002 | 0.00911 | 1 |
| A4 | JuliD1 | 0.63736 | 0.08543 | 0.0071 | 1 |
| A4 | JuliD3 | 1.15255 | 0.06078 | 0.00632 | 1 |
| A4 | JuliD4 | 0.70428 | 0.09266 | 0.00235 | 1 |
| A4 | MDt_1 | 1.39913 | 0.02874 | 0.00698 | 1 |
| A4 | Slag1 | -1.38316 | -0.23051 | 0 | 1 |
| A4 | Slag2 | 0.68435 | -0.10827 | 0.0047 | 1 |
| A4 | Slag4 | -0.623265 | -0.17662 | 0.00178 | 2 |
| A4 | Slag5 | -0.83762 | -0.20076 | 0.00336 | 1 |
| A4 | Slag6 | 0.78322 | -0.07002 | 0.00145 | 1 |
| A4 | yS | 0.81254 | 0.03526 | 0.00051 | 3 |
| A5 | AugD2 | -0.31998 | -0.05862 | 0.00427 | 1 |
| A5 | JDt2 | 0.80025 | 0.02794 | 0.00001 | 1 |
| A5 | MDF | 0.33864 | 0.07678 | 0.00298 | 1 |
| A5 | Slag5 | -0.25607 | -0.15393 | 0.00173 | 1 |
| A5 | yS | 0.2854 | 0.00328 | 0.00011 | 1 |

*Table A 3 Summarizing data for A1.1 forecast*

| Forecast ID | Total waste amount (SEK) | Total sales amount (SEK) | Number of observations | Standard deviation of sales | Standard deviation of waste | Coefficient of variation of sales | Coefficient of variation of waste |
|---|---|---|---|---|---|---|---|
| A1.1 | 153,903 | 3,560,651 | 2956 | 861.28 | 149.87 | 0.72 | 2.88 |

*Table A 4 Summarizing data for A3 forecasts*

| Forecast ID | Total waste amount (SEK) | Total sales amount (SEK) | Number of observations | Distinct SKU: s | Distinct Subcategories 2 | Standard deviation of sales | Standard deviation of waste | Coefficient of variation of sales | Coefficient of variation of waste | Number of waste observations > 0 | Average amount of waste per observation >0 (SEK) | Share of observations containing waste data |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A3.1 | 298,875 | 6,868,992 | 13,263 | 69 | 1 | 654.98 | 86.82 | 1.265 | 3.85 | 2595 | 115.17 | 19.57% |
| **A3.2** | 172,803 | 23,091,321 | 81,305 | 99 | 1 | 388.93 | 25.30 | 1.369 | 11.90 | 1891 | 91.382 | 2.33% |
| **A3.3** | 148,563 | 16,898,345 | 37,541 | 118 | 1 | 999.35 | 44.03 | 2.220 | 11.13 | 996 | 149.160 | 2.65% |
| **A3.4** | 88,829 | 9,510,930 | 35,327 | 28 | 1 | 281.80 | 21.25 | 1.047 | 8.45 | 959 | 92.627 | 2.71% |
| **A3.5** | 96,746 | 8,958,957 | 18,607 | 45 | 1 | 1008.18 | 45.33 | 2.094 | 8.72 | 604 | 160.175 | 3.25% |
| **A3.6** | 61,802 | 8,166,113 | 10,679 | 54 | 1 | 1452.03 | 57.59 | 1.899 | 9.95 | 271 | 228.052 | 2.54% |
| **A3.7** | 82,203 | 13,186,916 | 31,086 | 92 | 1 | 702.48 | 30.48 | 1.656 | 11.53 | 489 | 168.104 | 1.57% |
| **A3.8** | 63,406 | 13,261,015 | 27,407 | 14 | 1 | 545.21 | 27.22 | 1.127 | 11.77 | 739 | 85.800 | 2.70% |
| **A3.9** | 69,658 | 4,279,205 | 57,836 | 45 | 4 | 71.95 | 9.13 | 0.972 | 7.58 | 2373 | 29.354 | 4.10% |
| **A3.10** | 51,336 | 4,122,920 | 17,113 | 55 | 1 | 295.88 | 34.66 | 1.228 | 11.55 | 402 | 127.701 | 2.35% |
| **A3.11** | 51,293 | 8,108,016 | 32,628 | 67 | 1 | 354.61 | 19.79 | 1.427 | 12.59 | 459 | 111.749 | 1.41% |
| **A3.12** | 46,535 | 1,668,831 | 9,463 | 7 | 1 | 192.33 | 30.25 | 1.091 | 6.15 | 637 | 73.053 | 6.73% |
| **A3.13** | 49,132 | 2,854,701 | 23,710 | 34 | 1 | 125.96 | 14.12 | 1.046 | 6.81 | 813 | 60.433 | 3.43% |
| **A3.14** | 39,505 | 5,374,309 | 30,378 | 21 | 2 | 204.54 | 11.42 | 1.156 | 8.78 | 1217 | 32.461 | 4.01% |
| **A3.15** | 47,838 | 5,676,582 | 108,150 | 91 | 6 | 45.87 | 5.30 | 0.874 | 11.97 | 1649 | 29.010 | 1.52% |
| **A3.16** | 40,880 | 5,930,026 | 15,225 | 16 | 1 | 403.58 | 20.38 | 1.036 | 7.59 | 472 | 86.610 | 3.10% |
| **A3.17** | 47,334 | 2,848,252 | 44,181 | 37 | 4 | 49.08 | 11.14 | 0.761 | 10.40 | 921 | 51.394 | 2.08% |
| **A3.18** | 35,438 | 10,884,747 | 28,937 | 20 | 2 | 377.50 | 20.59 | 1.004 | 16.81 | 971 | 36.496 | 3.36% |
| **A3.19** | 37,853 | 2,748,878 | 20,007 | 7 | 1 | 106.24 | 15.41 | 0.773 | 8.15 | 813 | 46.560 | 4.06% |
| **A3.20** | 33,034 | 1,141,419 | 4,209 | 7 | 1 | 218.06 | 35.46 | 0.804 | 4.52 | 397 | 83.209 | 9.43% |
| **A3.21** | 28,340 | 2,226,778 | 17,392 | 28 | 3 | 146.35 | 32.18 | 1.143 | 19.75 | 285 | 99.439 | 1.64% |
| **A3.22** | 28,039 | 2,267,216 | 14,208 | 9 | 2 | 178.94 | 13.65 | 1.121 | 6.92 | 976 | 28.728 | 6.87% |
| **A3.23** | 24,542 | 1,850,295 | 42,306 | 37 | 5 | 31.59 | 5.98 | 0.722 | 10.31 | 846 | 29.009 | 2.00% |
| **A3.24** | 25,270 | 8,149,683 | 63,628 | 44 | 2 | 181.62 | 6.78 | 1.418 | 17.07 | 712 | 35.492 | 1.12% |
| **A3.25** | 24,692 | 8,727,517 | 23,610 | 20 | 3 | 376.54 | 12.94 | 1.019 | 12.37 | 830 | 29.749 | 3.52% |
| **A3.26** | 23,184 | 10,715,082 | 86,159 | 92 | 5 | 123.12 | 5.46 | 0.990 | 20.30 | 456 | 50.842 | 0.53% |
| **A3.27** | 22,682 | 2,229,241 | 32,694 | 40 | 6 | 73.70 | 8.29 | 1.081 | 11.95 | 551 | 41.165 | 1.69% |
| **A3.28** | 22,306 | 2,495,884 | 14,611 | 21 | 1 | 229.92 | 24.56 | 1.346 | 16.09 | 228 | 97.833 | 1.56% |
| **A3.29** | 19,992 | 4,193,579 | 31,002 | 27 | 1 | 186.47 | 8.53 | 1.379 | 13.23 | 539 | 37.091 | 1.74% |
| **A3.30** | 17,647 | 543,264 | 10,880 | 7 | 3 | 36 | 9.26 | 0.723 | 5.71 | 621 | 28.417 | 5.71% |

*Table A 5 Variables' significance for each forecast*

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

| Forecast | Variable | Estimate | Std. Error | Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|---|---|
| A1.1 | (Intercept) | | 0.33 | 1.66 | 0.20 | 0.842876 |
| A1.1 | lag1 | | 0.10 | 0.03 | 3.30 | 0.000991*** |
| A1.1 | lag3 | | 0.04 | 0.03 | 1.49 | 0.137431 |
| A1.1 | lag4 | | 0.05 | 0.03 | 1.49 | 0.135851 |
| A1.1 | lag6 | | 0.08 | 0.03 | 2.54 | 0.011217* |
| A1.1 | yS | | 0.29 | 0.19 | 1.56 | 0.120121 |
| A1.1 | Slag2 | | 0.57 | 0.17 | 3.26 | 0.001159** |
| A1.1 | Slag6 | | -0.61 | 0.15 | -4.12 | 0.0000414*** |
| A1.1 | Central | | -1.25 | 0.30 | -4.15 | 0.0000368*** |
| A1.1 | MDF2 | | 1.02 | 0.69 | 1.47 | 0.143323 |
| A1.1 | EDt22 | | -2.18 | 1.49 | -1.47 | 0.143082 |
| A1.1 | MDt2 | | -2.79 | 1.26 | -2.21 | 0.027039* |
| A1.1 | JDt2 | | -2.02 | 0.87 | -2.31 | 0.021256* |
| A1.1 | JDt_72 | | -2.31 | 1.51 | -1.53 | 0.127072 |
| A1.1 | AugD22 | | -1.89 | 0.57 | -3.30 | 0.001007** |
| A3.1 | (Intercept) | | 5.89 | 2.32 | 2.54 | 0.011296* |
| A3.1 | lag2 | | 0.05 | 0.03 | 1.66 | 0.097743. |
| A3.1 | lag4 | | 0.06 | 0.03 | 1.99 | 0.047177* |
| A3.1 | lag5 | | 0.08 | 0.03 | 2.55 | 0.010987* |
| A3.1 | lag6 | | 0.06 | 0.03 | 2.09 | 0.036727* |
| A3.1 | Slag2 | | 0.54 | 0.18 | 3.03 | 0.002492** |
| A3.1 | Slag4 | | 0.44 | 0.18 | 2.40 | 0.016525* |
| A3.1 | Slag6 | | -1.34 | 0.17 | -7.70 | 3.19E-14*** |
| A3.1 | Normal | | 0.66 | 0.44 | 1.51 | 0.130357 |
| A3.1 | EDt_32 | | -3.85 | 1.30 | -2.95 | 0.003208** |
| A3.1 | EDt32 | | -2.67 | 1.31 | -2.04 | 0.041503* |
| A3.1 | MDt2 | | -2.68 | 0.92 | -2.90 | 0.003802** |
| A3.1 | JDt_12 | | -2.58 | 1.30 | -1.98 | 0.048003* |
| A3.1 | JDt12 | | 2.20 | 1.36 | 1.62 | 0.105402 |
| A3.1 | JanD22 | | -2.23 | 1.33 | -1.68 | 0.092941. |
| A3.1 | JanD32 | | -3.15 | 1.31 | -2.41 | 0.016329* |
| A3.1 | JanD52 | | -3.14 | 1.31 | -2.40 | 0.016405* |
| A3.1 | JanD62 | | -3.00 | 1.32 | -2.27 | 0.023278* |
| A3.1 | AugD12 | | -1.41 | 0.50 | -2.81 | 0.005074** |
| A3.1 | AugD22 | | -0.91 | 0.51 | -1.80 | 0.072973. |
| A3.1 | AugD32 | | -1.69 | 0.50 | -3.37 | 0.000791*** |
| A3.2 | (Intercept) | | -5.24 | 3.85 | -1.36 | 0.173925 |
| A3.2 | lag1 | | 0.10 | 0.03 | 3.45 | 0.000584*** |
| A3.2 | lag2 | | 0.10 | 0.03 | 3.32 | 0.000944*** |
| A3.2 | lag3 | | 0.05 | 0.03 | 1.52 | 0.128793 |
| A3.2 | lag4 | | 0.06 | 0.03 | 1.98 | 0.047964* |
| A3.2 | lag5 | | 0.05 | 0.03 | 1.53 | 0.125878 |
| A3.2 | yS | | 1.86 | 0.32 | 5.82 | 0.00000000789*** |
| A3.2 | Slag4 | | -1.23 | 0.31 | -4.03 | 0.0000606*** |
| A3.2 | Slag6 | | 0.47 | 0.31 | 1.50 | 0.132996 |
| A3.2 | Normal | | -3.69 | 2.32 | -1.59 | 0.112945 |
| A3.2 | Central | | -4.79 | 3.15 | -1.52 | 0.129159 |

| | | | | | |
|---|---|---|---|---|---|
| A3.2 | Week2 | 0.29 | 0.15 | 2.01 | 0.044666* |
| A3.2 | MDF2 | 1.48 | 0.58 | 2.56 | 0.01073* |
| A3.2 | CDF2 | 0.68 | 0.32 | 2.14 | 0.032867* |
| A3.2 | EDt12 | -2.27 | 1.30 | -1.74 | 0.082558. |
| A3.2 | MDt_22 | -1.89 | 1.27 | -1.49 | 0.136507 |
| A3.2 | JDt_12 | 3.17 | 1.44 | 2.21 | 0.027462* |
| A3.2 | JDt_52 | -2.20 | 1.34 | -1.65 | 0.100294 |
| A3.2 | JDt_62 | -4.34 | 1.34 | -3.25 | 0.001201** |
| A3.2 | JDt12 | -3.17 | 1.42 | -2.23 | 0.026067* |
| A3.2 | JanD42 | -2.45 | 1.35 | -1.82 | 0.069245. |
| A3.2 | JanD92 | -3.79 | 1.34 | -2.84 | 0.004641** |
| A3.2 | JanD102 | -3.73 | 1.34 | -2.78 | 0.00552** |
| A3.2 | JuliD12 | 1.09 | 0.51 | 2.13 | 0.033132* |
| A3.2 | JuliD22 | 1.76 | 0.52 | 3.36 | 0.000804*** |
| A3.2 | JuliD42 | 1.70 | 0.54 | 3.17 | 0.001584** |
| A3.3 | (Intercept) | 0.10 | 2.18 | 0.05 | 0.96393 |
| A3.3 | lag2 | 0.07 | 0.03 | 2.27 | 0.02335* |
| A3.3 | lag5 | 0.06 | 0.03 | 1.82 | 0.06887. |
| A3.3 | yS | 0.78 | 0.20 | 3.95 | 0.0000841*** |
| A3.3 | Slag1 | -0.66 | 0.20 | -3.27 | 0.00112** |
| A3.3 | Normal | 0.92 | 0.52 | 1.77 | 0.07639. |
| A3.3 | Week2 | 0.34 | 0.17 | 2.01 | 0.04456* |
| A3.3 | MDF2 | 4.57 | 1.84 | 2.49 | 0.013* |
| A3.3 | EDt_12 | -2.56 | 1.49 | -1.72 | 0.08607. |
| A3.3 | EDt12 | 2.28 | 1.49 | 1.52 | 0.12801 |
| A3.3 | MDt2 | -5.42 | 2.13 | -2.55 | 0.01099* |
| A3.3 | MDt_12 | -3.39 | 2.23 | -1.52 | 0.12892 |
| A3.3 | MDt_22 | -4.09 | 2.25 | -1.82 | 0.06921. |
| A3.3 | MDt12 | -4.30 | 2.60 | -1.66 | 0.09806. |
| A3.3 | MDt22 | -4.04 | 2.59 | -1.56 | 0.11908 |
| A3.3 | JDt2 | 1.41 | 0.88 | 1.60 | 0.10906 |
| A3.3 | JDt_42 | 2.57 | 1.49 | 1.72 | 0.08543. |
| A3.3 | JanD12 | -3.14 | 1.51 | -2.08 | 0.03803* |
| A3.3 | JuliD22 | 1.23 | 0.57 | 2.14 | 0.03271* |
| A3.4 | (Intercept) | 15.33 | 4.40 | 3.48 | 0.000518*** |
| A3.4 | lag4 | 0.05 | 0.03 | 1.70 | 0.089003. |
| A3.4 | lag5 | 0.06 | 0.03 | 2.04 | 0.041911* |
| A3.4 | lag6 | 0.07 | 0.03 | 2.32 | 0.020464* |
| A3.4 | yS | 0.44 | 0.17 | 2.54 | 0.011275* |
| A3.4 | Slag3 | -0.37 | 0.18 | -2.10 | 0.036423* |
| A3.4 | Slag5 | -0.51 | 0.16 | -3.19 | 0.001479** |
| A3.4 | Normal | -9.70 | 3.02 | -3.22 | 0.001345** |
| A3.4 | Central | -8.77 | 3.32 | -2.64 | 0.008347** |
| A3.4 | MDF2 | 1.35 | 0.63 | 2.16 | 0.03122* |
| A3.4 | EDt2 | 1.49 | 0.65 | 2.30 | 0.021861* |
| A3.4 | EDt12 | -2.20 | 1.34 | -1.64 | 0.10242 |
| A3.4 | EDt22 | 2.03 | 1.35 | 1.51 | 0.131282 |
| A3.4 | MDt_12 | -2.65 | 1.32 | -2.02 | 0.044083* |
| A3.4 | MDt12 | -3.73 | 1.75 | -2.13 | 0.033568* |
| A3.4 | JDt12 | 2.09 | 1.34 | 1.56 | 0.119025 |
| A3.4 | JDt32 | 1.92 | 1.36 | 1.42 | 0.157428 |
| A3.4 | JanD32 | 2.19 | 1.35 | 1.62 | 0.105931 |

| | | | | | |
|---|---|---|---|---|---|
| A3.4 | JanD82 | 2.74 | 1.34 | 2.04 | 0.04209* |
| A3.4 | JuliD12 | 1.26 | 0.52 | 2.43 | 0.01527* |
| A3.4 | JuliD32 | 1.47 | 0.54 | 2.74 | 0.006286** |
| A3.4 | AugD12 | 1.29 | 0.53 | 2.44 | 0.014706* |
| A3.5 | (Intercept) | -5.37 | 1.26 | -4.25 | 0.0000238*** |
| A3.5 | yS | 0.41 | 0.08 | 5.14 | 0.000000324*** |
| A3.5 | Slag2 | 0.17 | 0.10 | 1.73 | 0.08405. |
| A3.5 | Slag3 | 0.39 | 0.12 | 3.20 | 0.00142** |
| A3.5 | Slag4 | -0.30 | 0.12 | -2.44 | 0.01504* |
| A3.5 | Slag5 | 0.18 | 0.10 | 1.84 | 0.06566. |
| A3.5 | MDF2 | 2.55 | 0.77 | 3.30 | 0.00101** |
| A3.5 | CDF2 | -0.87 | 0.32 | -2.72 | 0.00669** |
| A3.5 | EDt_22 | 2.26 | 1.40 | 1.61 | 0.10676 |
| A3.5 | EDt12 | 3.44 | 1.40 | 2.46 | 0.01412* |
| A3.5 | MDt2 | -4.14 | 1.26 | -3.29 | 0.00102** |
| A3.5 | MDt_12 | -3.29 | 1.43 | -2.30 | 0.02189* |
| A3.5 | JDt_52 | 2.47 | 1.43 | 1.72 | 0.08504. |
| A3.5 | JDt_62 | 2.27 | 1.43 | 1.58 | 0.11393 |
| A3.5 | JDt12 | 2.54 | 1.44 | 1.76 | 0.07877. |
| A3.5 | JDt22 | -2.12 | 1.45 | -1.47 | 0.14287 |
| A3.5 | JanD32 | 2.25 | 1.46 | 1.55 | 0.12273 |
| A3.6 | (Intercept) | 0.78 | 0.89 | 0.88 | 0.3782 |
| A3.6 | lag1 | 0.06 | 0.03 | 1.89 | 0.059017. |
| A3.6 | lag2 | 0.09 | 0.03 | 2.95 | 0.003282** |
| A3.6 | yS | 0.19 | 0.09 | 2.21 | 0.027393* |
| A3.6 | Normal | -1.78 | 0.48 | -3.73 | 0.000201*** |
| A3.6 | Central | -1.32 | 0.56 | -2.38 | 0.017634* |
| A3.6 | Week2 | -0.26 | 0.13 | -1.94 | 0.052238. |
| A3.6 | EDF2 | 2.72 | 0.59 | 4.61 | 0.00000458*** |
| A3.6 | CDF2 | -0.46 | 0.25 | -1.83 | 0.068146. |
| A3.6 | EDt2 | -2.59 | 0.81 | -3.19 | 0.001481** |
| A3.6 | EDt_22 | -2.54 | 1.31 | -1.94 | 0.053039. |
| A3.6 | EDt_32 | -3.85 | 1.31 | -2.93 | 0.003494** |
| A3.6 | EDt12 | -4.00 | 1.31 | -3.05 | 0.002362** |
| A3.6 | EDt32 | -3.72 | 1.31 | -2.83 | 0.004694** |
| A3.6 | MDt2 | 2.81 | 0.83 | 3.37 | 0.000768*** |
| A3.6 | JDt_32 | 2.46 | 1.20 | 2.05 | 0.040774* |
| A3.6 | JDt32 | 2.82 | 1.20 | 2.35 | 0.019228* |
| A3.6 | JuliD22 | 0.90 | 0.45 | 1.97 | 0.048958* |
| A3.6 | JuliD42 | 0.78 | 0.45 | 1.72 | 0.086455. |
| A3.7 | (Intercept) | -0.86 | 1.82 | -0.47 | 0.63745 |
| A3.7 | lag2 | 0.04 | 0.03 | 1.44 | 0.149119 |
| A3.7 | lag3 | 0.08 | 0.03 | 2.49 | 0.013085* |
| A3.7 | yS | 0.57 | 0.15 | 3.83 | 0.000138*** |
| A3.7 | Slag1 | -0.45 | 0.18 | -2.48 | 0.013438* |
| A3.7 | Slag2 | 0.37 | 0.15 | 2.47 | 0.01376* |
| A3.7 | Slag4 | -0.23 | 0.12 | -1.94 | 0.052342. |
| A3.7 | CDF2 | -0.90 | 0.32 | -2.86 | 0.004381** |
| A3.7 | JDt_42 | 2.82 | 1.41 | 2.00 | 0.046288* |
| A3.7 | JDt_52 | 2.81 | 1.41 | 1.99 | 0.046363* |
| A3.7 | JDt_62 | 2.12 | 1.41 | 1.50 | 0.133841 |
| A3.7 | NyåraftD2 | 2.29 | 1.44 | 1.59 | 0.11289 |

| | | | | | |
|---|---|---|---|---|---|
| A3.7 | JanD62 | 2.94 | 1.41 | 2.08 | 0.037598* |
| A3.7 | AugD32 | -1.33 | 0.53 | -2.51 | 0.01224* |
| A3.8 | (Intercept) | 12.51 | 3.27 | 3.83 | 0.000136*** |
| A3.8 | lag1 | 0.13 | 0.03 | 4.21 | 0.0000274*** |
| A3.8 | lag2 | 0.07 | 0.03 | 2.40 | 0.01668* |
| A3.8 | lag3 | 0.07 | 0.03 | 2.39 | 0.016881* |
| A3.8 | lag4 | 0.06 | 0.03 | 2.08 | 0.037769* |
| A3.8 | Slag4 | -1.10 | 0.23 | -4.83 | 0.00000159*** |
| A3.8 | Slag6 | 0.47 | 0.22 | 2.12 | 0.034262* |
| A3.8 | Normal | -5.53 | 2.12 | -2.60 | 0.009337** |
| A3.8 | Central | -3.98 | 2.27 | -1.76 | 0.07957. |
| A3.8 | CDF2 | -0.72 | 0.28 | -2.62 | 0.009029** |
| A3.8 | JDt_32 | 1.86 | 1.28 | 1.46 | 0.145314 |
| A3.8 | JDt_62 | 3.17 | 1.28 | 2.47 | 0.013598* |
| A3.8 | JDt12 | 2.21 | 1.33 | 1.66 | 0.097356. |
| A3.8 | JuliD12 | 1.27 | 0.49 | 2.61 | 0.009148** |
| A3.8 | JuliD42 | 0.97 | 0.50 | 1.95 | 0.051066. |
| A3.9 | (Intercept) | 33.71 | 5.17 | 6.52 | 0.000000000109*** |
| A3.9 | lag1 | -0.08 | 0.03 | -2.54 | 0.011328* |
| A3.9 | lag5 | 0.06 | 0.03 | 2.06 | 0.039444* |
| A3.9 | lag6 | -0.05 | 0.03 | -1.54 | 0.123464 |
| A3.9 | yS | 1.32 | 0.36 | 3.67 | 0.000255*** |
| A3.9 | Slag1 | -1.47 | 0.35 | -4.17 | 0.0000324*** |
| A3.9 | Slag3 | -0.79 | 0.36 | -2.21 | 0.027653* |
| A3.9 | Slag4 | -1.40 | 0.42 | -3.34 | 0.000865*** |
| A3.9 | Slag5 | -1.61 | 0.41 | -3.94 | 0.0000885*** |
| A3.9 | Slag6 | 1.19 | 0.35 | 3.37 | 0.000789*** |
| A3.9 | Normal | -7.73 | 3.41 | -2.27 | 0.023686* |
| A3.9 | Central | -8.64 | 3.47 | -2.49 | 0.012963* |
| A3.9 | MDF2 | -1.35 | 0.54 | -2.50 | 0.012642* |
| A3.9 | MDt_12 | 1.72 | 1.14 | 1.50 | 0.133934 |
| A3.9 | MDt12 | 2.65 | 1.51 | 1.76 | 0.078651. |
| A3.9 | JDt_22 | 1.94 | 1.12 | 1.74 | 0.08285. |
| A3.9 | JDt_42 | -3.17 | 1.12 | -2.84 | 0.00462** |
| A3.9 | JDt12 | -3.33 | 1.28 | -2.60 | 0.009498** |
| A3.9 | JDt22 | -3.55 | 1.28 | -2.78 | 0.005604** |
| A3.9 | JanD32 | -2.64 | 1.13 | -2.34 | 0.019726* |
| A3.9 | JuliD22 | -0.71 | 0.46 | -1.55 | 0.121866 |
| A3.9 | JuliD32 | -1.11 | 0.47 | -2.33 | 0.020169* |
| A3.9 | JuliD42 | -0.68 | 0.48 | -1.42 | 0.154628 |
| A3.9 | AugD22 | -0.81 | 0.44 | -1.83 | 0.067123. |
| A3.10 | (Intercept) | -5.46 | 1.36 | -4.01 | 0.0000657*** |
| A3.10 | lag1 | 0.09 | 0.03 | 3.01 | 0.00266** |
| A3.10 | lag2 | 0.09 | 0.03 | 2.85 | 0.00451** |
| A3.10 | lag3 | 0.07 | 0.03 | 2.14 | 0.03237* |
| A3.10 | lag4 | 0.07 | 0.03 | 2.36 | 0.01855* |
| A3.10 | lag5 | 0.08 | 0.03 | 2.60 | 0.00938** |
| A3.10 | lag6 | 0.06 | 0.03 | 1.84 | 0.06551. |
| A3.10 | yS | 0.43 | 0.14 | 3.09 | 0.00207** |
| A3.10 | Slag2 | 0.33 | 0.14 | 2.35 | 0.01899* |
| A3.10 | EDF2 | -0.79 | 0.43 | -1.84 | 0.06557. |
| A3.10 | EDt2 | 1.59 | 0.72 | 2.22 | 0.02682* |

| | | | | | |
|---|---|---|---|---|---|
| A3.10 | MDt2 | 1.20 | 0.85 | 1.41 | 0.15752 |
| A3.10 | JanD52 | 3.64 | 1.19 | 3.07 | 0.00223** |
| A3.10 | AugD32 | 0.68 | 0.45 | 1.51 | 0.13233 |
| A3.11 | (Intercept) | 3.57 | 1.60 | 2.24 | 0.02537* |
| A3.11 | lag1 | 0.07 | 0.03 | 2.25 | 0.02461* |
| A3.11 | lag2 | 0.05 | 0.03 | 1.49 | 0.13662 |
| A3.11 | Slag1 | -0.27 | 0.18 | -1.48 | 0.14046 |
| A3.11 | Week2 | -0.21 | 0.14 | -1.52 | 0.12855 |
| A3.11 | MDF2 | 4.05 | 1.51 | 2.69 | 0.00735** |
| A3.11 | EDt_12 | 2.58 | 1.23 | 2.11 | 0.03552* |
| A3.11 | MDt2 | -5.25 | 1.73 | -3.03 | 0.00248** |
| A3.11 | MDt_12 | -3.28 | 1.83 | -1.79 | 0.07369. |
| A3.11 | MDt_22 | -2.81 | 1.83 | -1.53 | 0.12624 |
| A3.11 | MDt12 | -3.51 | 2.12 | -1.65 | 0.09868. |
| A3.11 | MDt22 | -5.29 | 2.12 | -2.50 | 0.01265* |
| A3.11 | JDt_22 | 3.87 | 1.27 | 3.06 | 0.00231** |
| A3.11 | JDt_42 | 3.99 | 1.23 | 3.24 | 0.00124** |
| A3.11 | JanD32 | 3.37 | 1.22 | 2.76 | 0.00595** |
| A3.11 | JuliD12 | 1.35 | 0.47 | 2.87 | 0.00413** |
| A3.11 | JuliD42 | 1.36 | 0.47 | 2.90 | 0.00377** |
| A3.12 | (Intercept) | 4.31 | 1.71 | 2.52 | 0.01187* |
| A3.12 | lag1 | 0.15 | 0.03 | 4.91 | 0.00000104*** |
| A3.12 | lag2 | 0.12 | 0.03 | 3.96 | 0.0000802*** |
| A3.12 | lag3 | 0.08 | 0.03 | 2.54 | 0.01116* |
| A3.12 | lag5 | 0.08 | 0.03 | 2.77 | 0.00565** |
| A3.12 | yS | 0.26 | 0.16 | 1.64 | 0.10166 |
| A3.12 | Slag2 | -0.32 | 0.16 | -2.01 | 0.04454* |
| A3.12 | Normal | -3.08 | 1.01 | -3.07 | 0.00222** |
| A3.12 | Central | -3.10 | 1.04 | -2.98 | 0.00295** |
| A3.12 | EDF2 | 0.78 | 0.54 | 1.44 | 0.14986 |
| A3.12 | EDt2 | -1.38 | 0.79 | -1.75 | 0.08028. |
| A3.12 | EDt_22 | -3.07 | 1.32 | -2.33 | 0.02002* |
| A3.12 | EDt12 | -2.56 | 1.32 | -1.94 | 0.05246. |
| A3.12 | EDt32 | -2.10 | 1.31 | -1.60 | 0.11043 |
| A3.12 | MDt2 | 1.85 | 0.85 | 2.18 | 0.02986* |
| A3.12 | JDt2 | 1.18 | 0.71 | 1.65 | 0.09838. |
| A3.12 | JDt_32 | 2.39 | 1.20 | 1.99 | 0.04723* |
| A3.12 | JuliD32 | 1.07 | 0.47 | 2.26 | 0.02422* |
| A3.13 | (Intercept) | -2.90 | 2.23 | -1.30 | 0.19388 |
| A3.13 | lag1 | 0.06 | 0.03 | 1.82 | 0.06905. |
| A3.13 | lag2 | 0.06 | 0.03 | 1.89 | 0.0588. |
| A3.13 | lag5 | 0.12 | 0.03 | 3.98 | 0.0000734*** |
| A3.13 | yS | 0.62 | 0.22 | 2.84 | 0.00461** |
| A3.13 | Slag1 | 0.36 | 0.23 | 1.54 | 0.12375 |
| A3.13 | Slag2 | -0.40 | 0.21 | -1.87 | 0.06188. |
| A3.13 | Central | -2.10 | 0.90 | -2.32 | 0.02038* |
| A3.13 | CDF2 | -0.88 | 0.31 | -2.87 | 0.00414** |
| A3.13 | EDt_32 | 2.33 | 1.24 | 1.87 | 0.06118. |
| A3.13 | JDt2 | 1.77 | 0.79 | 2.23 | 0.02592* |
| A3.13 | JDt_32 | 2.45 | 1.29 | 1.89 | 0.05873. |
| A3.13 | JDt12 | 1.92 | 1.28 | 1.50 | 0.13344 |
| A3.13 | JDt22 | 3.21 | 1.27 | 2.52 | 0.01192* |

| | | | | | |
|---|---|---|---|---|---|
| A3.13 | JanD32 | 2.70 | 1.27 | 2.12 | 0.03429* |
| A3.13 | JanD42 | 1.96 | 1.27 | 1.54 | 0.12371 |
| A3.13 | JuliD12 | 0.95 | 0.47 | 1.99 | 0.04681* |
| A3.13 | JuliD32 | 0.85 | 0.48 | 1.79 | 0.07412. |
| A3.14 | (Intercept) | 11.68 | 3.28 | 3.56 | 0.000382*** |
| A3.14 | lag3 | 0.09 | 0.03 | 2.94 | 0.003374** |
| A3.14 | lag4 | 0.08 | 0.03 | 2.63 | 0.008778** |
| A3.14 | lag5 | 0.05 | 0.03 | 1.71 | 0.088079. |
| A3.14 | lag6 | 0.06 | 0.03 | 2.04 | 0.041983* |
| A3.14 | yS | 1.02 | 0.39 | 2.58 | 0.009979** |
| A3.14 | Slag1 | -1.48 | 0.42 | -3.51 | 0.000476*** |
| A3.14 | Slag3 | 0.86 | 0.38 | 2.24 | 0.025375* |
| A3.14 | Slag5 | -0.76 | 0.40 | -1.90 | 0.058306. |
| A3.14 | Normal | -7.21 | 2.03 | -3.55 | 0.000398*** |
| A3.14 | Central | -5.37 | 2.86 | -1.88 | 0.060967. |
| A3.14 | MDF2 | 1.32 | 0.59 | 2.24 | 0.025076* |
| A3.14 | CDF2 | -0.58 | 0.25 | -2.30 | 0.021554* |
| A3.14 | EDt_12 | 1.49 | 1.05 | 1.42 | 0.15723 |
| A3.14 | EDt_42 | -2.27 | 1.05 | -2.16 | 0.031201* |
| A3.14 | EDt32 | 1.53 | 1.06 | 1.45 | 0.147677 |
| A3.14 | MDt2 | -4.18 | 0.97 | -4.29 | 0.0000197*** |
| A3.14 | MDt_22 | -1.70 | 1.08 | -1.57 | 0.115938 |
| A3.14 | JDt2 | 1.14 | 0.73 | 1.56 | 0.11899 |
| A3.14 | JDt_22 | 2.31 | 1.09 | 2.11 | 0.035085* |
| A3.14 | JDt_72 | 2.19 | 1.08 | 2.03 | 0.042455* |
| A3.14 | JDt12 | 2.85 | 1.13 | 2.52 | 0.011826* |
| A3.14 | JuliD12 | 0.76 | 0.41 | 1.84 | 0.065437. |
| A3.15 | (Intercept) | 18.05 | 5.48 | 3.29 | 0.001025** |
| A3.15 | lag1 | -0.09 | 0.03 | -3.04 | 0.002396** |
| A3.15 | yS | 1.47 | 0.35 | 4.14 | 0.0000375*** |
| A3.15 | Slag1 | -1.19 | 0.36 | -3.34 | 0.000875*** |
| A3.15 | Slag5 | -1.12 | 0.35 | -3.20 | 0.001402** |
| A3.15 | Slag6 | 1.28 | 0.38 | 3.35 | 0.000839*** |
| A3.15 | Normal | -19.53 | 4.24 | -4.60 | 0.00000469*** |
| A3.15 | Central | -19.79 | 4.33 | -4.57 | 0.00000537*** |
| A3.15 | EDt_12 | 1.56 | 1.11 | 1.40 | 0.161763 |
| A3.15 | MDt_12 | 2.10 | 0.99 | 2.12 | 0.034328* |
| A3.15 | MDt22 | 2.16 | 1.37 | 1.58 | 0.113834 |
| A3.15 | JDt32 | 3.25 | 1.26 | 2.58 | 0.010026* |
| A3.15 | JanD32 | -1.89 | 1.12 | -1.70 | 0.089865. |
| A3.15 | JanD62 | 2.39 | 1.13 | 2.12 | 0.03457* |
| A3.15 | JanD72 | 2.51 | 1.11 | 2.25 | 0.024669* |
| A3.15 | JanD82 | 2.66 | 1.12 | 2.38 | 0.01744* |
| A3.15 | JanD92 | 2.03 | 1.12 | 1.82 | 0.069791. |
| A3.15 | JuliD12 | 1.34 | 0.45 | 2.97 | 0.003038** |
| A3.15 | JuliD22 | 0.89 | 0.46 | 1.92 | 0.055345. |
| A3.15 | JuliD32 | 0.97 | 0.48 | 2.03 | 0.042842* |
| A3.15 | JuliD42 | 1.75 | 0.48 | 3.65 | 0.000272*** |
| A3.15 | AugD12 | 1.38 | 0.45 | 3.11 | 0.001948** |
| A3.15 | AugD22 | 1.15 | 0.45 | 2.58 | 0.010116* |
| A3.16 | (Intercept) | 16.05 | 2.46 | 6.51 | 0.000000000116*** |
| A3.16 | lag2 | 0.06 | 0.03 | 1.91 | 0.056856. |

| | | | | | |
|---|---|---|---|---|---|
| A3.16 | lag3 | 0.05 | 0.03 | 1.76 | 0.079458. |
| A3.16 | lag6 | 0.05 | 0.03 | 1.51 | 0.130707 |
| A3.16 | Slag2 | -0.46 | 0.18 | -2.54 | 0.011341* |
| A3.16 | Slag4 | -0.59 | 0.21 | -2.74 | 0.006229** |
| A3.16 | Slag5 | -0.39 | 0.21 | -1.86 | 0.063863. |
| A3.16 | Normal | -2.79 | 0.96 | -2.90 | 0.003832** |
| A3.16 | Central | -2.12 | 1.17 | -1.81 | 0.069967. |
| A3.16 | CDF2 | -1.05 | 0.27 | -3.88 | 0.000112*** |
| A3.16 | EDt_32 | 1.98 | 1.17 | 1.70 | 0.090428. |
| A3.16 | MDt22 | -2.33 | 1.45 | -1.61 | 0.108109 |
| A3.16 | JDt_32 | 2.93 | 1.20 | 2.45 | 0.014458* |
| A3.16 | JDt_42 | 2.09 | 1.20 | 1.74 | 0.081977. |
| A3.16 | JDt32 | 2.89 | 1.21 | 2.39 | 0.017189* |
| A3.16 | JanD22 | 2.46 | 1.19 | 2.06 | 0.040078* |
| A3.16 | JanD32 | 3.85 | 1.20 | 3.20 | 0.001404** |
| A3.16 | JuliD12 | 0.79 | 0.46 | 1.73 | 0.083889. |
| A3.16 | JuliD42 | 0.91 | 0.48 | 1.92 | 0.054621. |
| A3.16 | AugD32 | -0.75 | 0.45 | -1.66 | 0.097438. |
| A3.17 | (Intercept) | 17.00 | 3.56 | 4.78 | 0.00000202*** |
| A3.17 | lag1 | 0.07 | 0.03 | 2.27 | 0.02321* |
| A3.17 | lag3 | 0.08 | 0.03 | 2.58 | 0.01005* |
| A3.17 | lag4 | 0.09 | 0.03 | 2.95 | 0.00325** |
| A3.17 | yS | -0.42 | 0.28 | -1.51 | 0.13152 |
| A3.17 | Slag2 | 0.41 | 0.29 | 1.42 | 0.15547 |
| A3.17 | Slag4 | -0.45 | 0.28 | -1.61 | 0.1071 |
| A3.17 | Normal | -12.17 | 2.50 | -4.87 | 0.00000126*** |
| A3.17 | Central | -12.50 | 2.53 | -4.94 | 0.000000918*** |
| A3.17 | MDF2 | 1.22 | 0.61 | 2.01 | 0.04475* |
| A3.17 | EDt2 | -0.84 | 0.58 | -1.43 | 0.15341 |
| A3.17 | EDt12 | 2.13 | 1.21 | 1.76 | 0.07858. |
| A3.17 | MDt2 | -2.74 | 1.04 | -2.62 | 0.0089** |
| A3.17 | MDt12 | -2.75 | 1.60 | -1.72 | 0.08591. |
| A3.17 | JDt12 | -2.04 | 1.24 | -1.64 | 0.1019 |
| A3.17 | JDt22 | 2.32 | 1.22 | 1.90 | 0.05718. |
| A3.17 | JDt32 | 2.21 | 1.21 | 1.83 | 0.06804. |
| A3.17 | JanD12 | 2.43 | 1.21 | 2.01 | 0.04506* |
| A3.17 | JanD22 | 1.76 | 1.21 | 1.45 | 0.14675 |
| A3.17 | JanD62 | -1.79 | 1.20 | -1.49 | 0.13695 |
| A3.18 | (Intercept) | 23.05 | 2.93 | 7.87 | 8.82E-15*** |
| A3.18 | yS | 0.85 | 0.34 | 2.51 | 0.012131* |
| A3.18 | Slag1 | -0.96 | 0.36 | -2.67 | 0.007735** |
| A3.18 | Slag4 | -0.69 | 0.36 | -1.89 | 0.058587. |
| A3.18 | Slag5 | -0.57 | 0.35 | -1.64 | 0.101201 |
| A3.18 | Normal | -8.75 | 2.30 | -3.81 | 0.000149*** |
| A3.18 | Central | -7.04 | 3.11 | -2.26 | 0.023963* |
| A3.18 | Week2 | -0.20 | 0.11 | -1.77 | 0.077868. |
| A3.18 | MDF2 | 1.62 | 0.47 | 3.43 | 0.000626*** |
| A3.18 | EDt_12 | 2.74 | 1.00 | 2.74 | 0.006218** |
| A3.18 | EDt22 | 1.60 | 1.01 | 1.59 | 0.112951 |
| A3.18 | MDt2 | -2.87 | 0.88 | -3.27 | 0.001106** |
| A3.18 | JDt_62 | 2.42 | 1.00 | 2.42 | 0.015584* |
| A3.18 | JDt12 | -2.33 | 1.09 | -2.13 | 0.033093* |

| | | | | | |
|---|---|---|---|---|---|
| A3.18 | JDt32 | -1.86 | 1.01 | -1.84 | 0.065691. |
| A3.18 | JanD22 | 2.17 | 1.00 | 2.16 | 0.031009* |
| A3.18 | JanD52 | 2.35 | 1.01 | 2.32 | 0.020609* |
| A3.18 | JanD62 | 1.47 | 1.00 | 1.48 | 0.140503 |
| A3.18 | JanD92 | 1.71 | 1.00 | 1.72 | 0.086301. |
| A3.18 | JuliD12 | 0.64 | 0.40 | 1.62 | 0.105862 |
| A3.18 | JuliD22 | 0.99 | 0.41 | 2.42 | 0.015845* |
| A3.18 | AugD22 | -0.84 | 0.39 | -2.16 | 0.030977* |
| A3.19 | (Intercept) | 2.95 | 2.23 | 1.32 | 0.18734 |
| A3.19 | lag4 | 0.11 | 0.03 | 3.55 | 0.000408*** |
| A3.19 | lag5 | 0.07 | 0.03 | 2.26 | 0.023949* |
| A3.19 | lag6 | -0.05 | 0.03 | -1.66 | 0.098222. |
| A3.19 | yS | 0.95 | 0.22 | 4.28 | 0.00002*** |
| A3.19 | Slag1 | -0.67 | 0.25 | -2.69 | 0.007329** |
| A3.19 | Slag2 | -0.52 | 0.22 | -2.38 | 0.01774* |
| A3.19 | Slag4 | -0.44 | 0.19 | -2.36 | 0.018453* |
| A3.19 | Slag6 | 0.47 | 0.20 | 2.42 | 0.015697* |
| A3.19 | EDF2 | -0.77 | 0.43 | -1.81 | 0.070306. |
| A3.19 | EDt2 | 1.40 | 0.69 | 2.04 | 0.04166* |
| A3.19 | EDt_12 | 3.26 | 1.18 | 2.75 | 0.006032** |
| A3.19 | MDt_22 | 1.41 | 0.96 | 1.47 | 0.142802 |
| A3.19 | MDt12 | 2.43 | 1.37 | 1.78 | 0.075928. |
| A3.19 | JDt_52 | 1.61 | 1.11 | 1.45 | 0.146546 |
| A3.19 | JDt_72 | -1.60 | 1.11 | -1.45 | 0.147683 |
| A3.19 | NyåraftD2 | 2.19 | 1.13 | 1.94 | 0.052624. |
| A3.19 | JanD12 | 2.30 | 1.13 | 2.04 | 0.041856* |
| A3.19 | JanD52 | 2.02 | 1.13 | 1.79 | 0.074009. |
| A3.19 | JanD102 | -1.79 | 1.11 | -1.62 | 0.106088 |
| A3.19 | JuliD12 | 0.76 | 0.43 | 1.79 | 0.073394. |
| A3.19 | JuliD22 | 1.16 | 0.43 | 2.66 | 0.0079** |
| A3.19 | JuliD32 | 0.78 | 0.44 | 1.76 | 0.078. |
| A3.19 | JuliD42 | 0.79 | 0.44 | 1.79 | 0.073343. |
| A3.19 | AugD12 | 0.92 | 0.44 | 2.10 | 0.035701* |
| A3.20 | (Intercept) | 1.19 | 1.25 | 0.96 | 0.3384 |
| A3.20 | lag1 | 0.05 | 0.03 | 1.68 | 0.093128. |
| A3.20 | lag2 | 0.09 | 0.03 | 2.83 | 0.004681** |
| A3.20 | lag3 | 0.11 | 0.03 | 3.55 | 0.000406*** |
| A3.20 | lag4 | 0.11 | 0.03 | 3.73 | 0.000202*** |
| A3.20 | lag5 | 0.09 | 0.03 | 2.83 | 0.004821** |
| A3.20 | lag6 | 0.12 | 0.03 | 3.96 | 0.0000796*** |
| A3.20 | yS | 0.53 | 0.12 | 4.62 | 0.00000437*** |
| A3.20 | Slag2 | -0.20 | 0.11 | -1.80 | 0.071445. |
| A3.20 | Slag6 | -0.34 | 0.10 | -3.26 | 0.001165** |
| A3.20 | Normal | -0.67 | 0.28 | -2.40 | 0.016713* |
| A3.20 | EDF2 | -1.37 | 0.44 | -3.09 | 0.002075** |
| A3.20 | CDF2 | 1.70 | 0.37 | 4.56 | 0.00000568*** |
| A3.20 | EDt2 | 1.95 | 0.68 | 2.86 | 0.004382** |
| A3.20 | EDt_12 | 3.98 | 1.15 | 3.46 | 0.000573*** |
| A3.20 | EDt_22 | 2.36 | 1.15 | 2.06 | 0.039551* |
| A3.20 | JDt2 | -1.29 | 0.72 | -1.80 | 0.071836. |
| A3.20 | JDt_12 | -2.06 | 1.13 | -1.82 | 0.068604. |
| A3.20 | JDt_22 | -3.67 | 1.13 | -3.25 | 0.001199** |

| | | | | | |
|---|---|---|---|---|---|
| A3.20 | JDt_32 | -1.86 | 1.13 | -1.64 | 0.101996 |
| A3.20 | JDt_52 | -4.00 | 1.13 | -3.55 | 0.000398*** |
| A3.20 | JDt22 | -1.79 | 1.13 | -1.58 | 0.114027 |
| A3.20 | JDt32 | -3.46 | 1.13 | -3.07 | 0.002232** |
| A3.20 | NyåraftD2 | -3.26 | 1.13 | -2.88 | 0.00402** |
| A3.20 | JanD12 | -2.93 | 1.13 | -2.60 | 0.009538** |
| A3.20 | JanD22 | -3.10 | 1.13 | -2.75 | 0.006039** |
| A3.20 | JanD32 | -2.81 | 1.13 | -2.50 | 0.01263* |
| A3.20 | JanD62 | -2.68 | 1.13 | -2.38 | 0.017643* |
| A3.20 | JanD72 | -1.94 | 1.13 | -1.72 | 0.085715. |
| A3.21 | (Intercept) | 0.67 | 0.06 | 10.71 | <2E-16 |
| A3.21 | lag3 | 0.05 | 0.03 | 1.61 | 0.10751 |
| A3.21 | lag6 | 0.07 | 0.03 | 2.25 | 0.02466* |
| A3.21 | EDF2 | -0.54 | 0.29 | -1.86 | 0.06359. |
| A3.21 | EDt12 | 1.92 | 1.00 | 1.92 | 0.05556. |
| A3.21 | MDt12 | 3.83 | 1.18 | 3.25 | 0.00118** |
| A3.21 | MDt22 | 2.17 | 1.18 | 1.84 | 0.06584. |
| A3.21 | JDt_22 | 2.28 | 0.96 | 2.38 | 0.01755* |
| A3.21 | JDt_42 | 2.97 | 0.96 | 3.09 | 0.00205** |
| A3.21 | JDt12 | 2.17 | 0.96 | 2.26 | 0.02432* |
| A3.21 | JuliD12 | 0.88 | 0.37 | 2.40 | 0.01648* |
| A3.21 | AugD42 | -0.56 | 0.37 | -1.54 | 0.12467 |
| A3.22 | (Intercept) | 10.61 | 2.54 | 4.17 | 0.000033*** |
| A3.22 | lag2 | 0.09 | 0.03 | 2.98 | 0.002997** |
| A3.22 | lag4 | 0.09 | 0.03 | 2.93 | 0.003514** |
| A3.22 | lag5 | 0.11 | 0.03 | 3.74 | 0.000192*** |
| A3.22 | Slag1 | -0.80 | 0.32 | -2.51 | 0.012372* |
| A3.22 | Slag2 | 0.58 | 0.32 | 1.85 | 0.064784. |
| A3.22 | Slag4 | -0.60 | 0.31 | -1.93 | 0.053677. |
| A3.22 | Slag6 | 0.43 | 0.30 | 1.45 | 0.148159 |
| A3.22 | Normal | -6.58 | 1.70 | -3.88 | 0.000111*** |
| A3.22 | Central | -4.79 | 2.14 | -2.24 | 0.025493* |
| A3.22 | EDt_32 | -1.61 | 1.03 | -1.57 | 0.117304 |
| A3.22 | MDt_22 | 1.30 | 0.89 | 1.46 | 0.145998 |
| A3.22 | MDt22 | 1.90 | 1.29 | 1.47 | 0.141855 |
| A3.22 | JDt_72 | -1.81 | 1.03 | -1.76 | 0.078042. |
| A3.22 | JanD12 | 1.65 | 1.03 | 1.60 | 0.110066 |
| A3.22 | JanD42 | -2.06 | 1.03 | -2.00 | 0.045643* |
| A3.22 | JuliD12 | 0.71 | 0.40 | 1.78 | 0.075132. |
| A3.23 | (Intercept) | 10.24 | 3.78 | 2.71 | 0.006917** |
| A3.23 | lag1 | -0.07 | 0.03 | -2.32 | 0.020503* |
| A3.23 | lag2 | -0.05 | 0.03 | -1.60 | 0.110785 |
| A3.23 | yS | 1.52 | 0.26 | 5.75 | 0.0000000121*** |
| A3.23 | Slag1 | -0.91 | 0.28 | -3.25 | 0.001176** |
| A3.23 | Slag2 | 0.44 | 0.28 | 1.57 | 0.117088 |
| A3.23 | Slag3 | -0.92 | 0.25 | -3.64 | 0.000282*** |
| A3.23 | Slag6 | 0.71 | 0.24 | 2.94 | 0.003409** |
| A3.23 | Normal | -14.76 | 3.58 | -4.12 | 0.0000411*** |
| A3.23 | Central | -18.35 | 3.69 | -4.97 | 0.00000079*** |
| A3.23 | MDt_12 | 1.67 | 0.91 | 1.84 | 0.065854. |
| A3.23 | JDt32 | 2.27 | 1.10 | 2.06 | 0.039284* |
| A3.23 | JanD12 | 2.15 | 1.05 | 2.05 | 0.041* |

| | | | | | |
|---|---|---|---|---|---|
| A3.23 | JanD42 | 3.30 | 1.04 | 3.17 | 0.001569** |
| A3.23 | JanD62 | 2.55 | 1.06 | 2.41 | 0.015942* |
| A3.23 | JanD72 | 1.55 | 1.04 | 1.48 | 0.138996 |
| A3.23 | JanD82 | 1.89 | 1.05 | 1.81 | 0.070659. |
| A3.23 | JuliD12 | 1.25 | 0.41 | 3.09 | 0.002081** |
| A3.23 | JuliD22 | 0.65 | 0.40 | 1.62 | 0.106112 |
| A3.23 | JuliD42 | 1.23 | 0.42 | 2.90 | 0.003816** |
| A3.24 | (Intercept) | 17.29 | 3.25 | 5.33 | 0.000000123*** |
| A3.24 | lag3 | 0.07 | 0.03 | 2.23 | 0.026249* |
| A3.24 | yS | 0.65 | 0.26 | 2.47 | 0.013667* |
| A3.24 | Slag1 | -0.62 | 0.32 | -1.92 | 0.055574. |
| A3.24 | Slag2 | 0.54 | 0.34 | 1.59 | 0.111518 |
| A3.24 | Slag3 | -0.56 | 0.33 | -1.67 | 0.094726. |
| A3.24 | Slag4 | -0.52 | 0.28 | -1.89 | 0.059209. |
| A3.24 | Normal | -11.50 | 2.89 | -3.98 | 0.0000749*** |
| A3.24 | Central | -10.57 | 3.16 | -3.35 | 0.000839*** |
| A3.24 | EDt_42 | 1.44 | 1.01 | 1.42 | 0.155196 |
| A3.24 | EDt32 | -1.56 | 1.02 | -1.54 | 0.124173 |
| A3.24 | MDt_22 | 1.65 | 0.89 | 1.85 | 0.065054. |
| A3.24 | JDt2 | 1.24 | 0.64 | 1.93 | 0.053981. |
| A3.24 | JDt_32 | 1.69 | 1.03 | 1.63 | 0.103122 |
| A3.24 | JDt_72 | 1.60 | 1.02 | 1.57 | 0.11597 |
| A3.24 | NyåraftD2 | 2.52 | 1.06 | 2.37 | 0.018081* |
| A3.24 | JanD22 | 2.19 | 1.08 | 2.04 | 0.042132* |
| A3.24 | JanD32 | -1.54 | 1.07 | -1.43 | 0.15185 |
| A3.24 | JanD92 | 1.67 | 1.01 | 1.65 | 0.098556. |
| A3.24 | JuliD12 | 0.81 | 0.41 | 2.00 | 0.045727* |
| A3.24 | JuliD32 | 0.75 | 0.44 | 1.71 | 0.088439. |
| A3.24 | JuliD42 | 0.83 | 0.44 | 1.90 | 0.058186. |
| A3.24 | AugD12 | 1.04 | 0.42 | 2.47 | 0.013821* |
| A3.25 | (Intercept) | -0.08 | 1.99 | -0.04 | 0.9687 |
| A3.25 | yS | 0.93 | 0.19 | 5.00 | 0.000000685*** |
| A3.25 | Slag5 | -0.74 | 0.18 | -4.07 | 0.0000507*** |
| A3.25 | EDt_22 | -1.45 | 1.01 | -1.43 | 0.1518 |
| A3.25 | EDt_42 | -1.50 | 1.01 | -1.48 | 0.1384 |
| A3.25 | JanD22 | 2.32 | 1.01 | 2.30 | 0.0215* |
| A3.25 | JuliD12 | 0.68 | 0.39 | 1.75 | 0.0814. |
| A3.25 | JuliD22 | 0.83 | 0.39 | 2.12 | 0.0339* |
| A3.25 | JuliD32 | -0.56 | 0.40 | -1.41 | 0.1579 |
| A3.26 | (Intercept) | 11.14 | 2.54 | 4.39 | 0.0000127*** |
| A3.26 | lag1 | 0.08 | 0.03 | 2.62 | 0.00895** |
| A3.26 | lag3 | 0.06 | 0.03 | 2.06 | 0.03948* |
| A3.26 | lag5 | 0.05 | 0.03 | 1.70 | 0.09018. |
| A3.26 | lag6 | 0.07 | 0.03 | 2.13 | 0.03316* |
| A3.26 | Slag3 | -0.55 | 0.23 | -2.42 | 0.0156* |
| A3.26 | Slag5 | -0.58 | 0.23 | -2.55 | 0.01096* |
| A3.26 | EDt_12 | 1.78 | 1.04 | 1.71 | 0.08853. |
| A3.26 | EDt_42 | 1.59 | 1.04 | 1.52 | 0.12798 |
| A3.26 | JDt_22 | 1.87 | 1.04 | 1.80 | 0.07188. |
| A3.26 | JuliD12 | 0.72 | 0.40 | 1.80 | 0.07176. |
| A3.26 | JuliD42 | -0.71 | 0.42 | -1.70 | 0.08919. |
| A3.27 | (Intercept) | 3.50 | 2.46 | 1.42 | 0.15599 |

| | | | | | |
|---|---|---|---|---|---|
| A3.27 | yS | 0.43 | 0.19 | 2.28 | 0.0227* |
| A3.27 | Slag4 | -0.47 | 0.20 | -2.31 | 0.02138* |
| A3.27 | Slag5 | -0.40 | 0.23 | -1.75 | 0.0809. |
| A3.27 | Slag6 | 0.44 | 0.22 | 2.03 | 0.04243* |
| A3.27 | Normal | -2.37 | 1.32 | -1.80 | 0.07268. |
| A3.27 | Central | -4.06 | 1.83 | -2.22 | 0.02694* |
| A3.27 | EDt22 | 2.38 | 1.04 | 2.29 | 0.02215* |
| A3.27 | MDt12 | 3.13 | 1.26 | 2.48 | 0.01345* |
| A3.27 | JDt32 | 2.49 | 1.06 | 2.34 | 0.01949* |
| A3.27 | JanD82 | 3.30 | 1.03 | 3.20 | 0.00142** |
| A3.27 | JuliD12 | 0.71 | 0.40 | 1.77 | 0.07784. |
| A3.27 | JuliD32 | 1.34 | 0.41 | 3.26 | 0.00116** |
| A3.27 | JuliD42 | 0.83 | 0.41 | 2.01 | 0.04431* |
| A3.27 | AugD12 | 0.59 | 0.41 | 1.43 | 0.15357 |
| A3.27 | AugD22 | 0.80 | 0.41 | 1.96 | 0.05077. |
| A3.27 | AugD32 | 0.97 | 0.40 | 2.42 | 0.01553* |
| A3.28 | (Intercept) | -1.61 | 1.89 | -0.86 | 0.392688 |
| A3.28 | lag3 | 0.13 | 0.03 | 4.18 | 0.0000316*** |
| A3.28 | yS | 0.24 | 0.13 | 1.78 | 0.07491. |
| A3.28 | Slag5 | 0.47 | 0.13 | 3.56 | 0.000389*** |
| A3.28 | Normal | -3.17 | 1.43 | -2.21 | 0.027284* |
| A3.28 | Central | -3.59 | 1.52 | -2.37 | 0.018106* |
| A3.28 | EDt2 | -0.66 | 0.46 | -1.45 | 0.148809 |
| A3.28 | EDt_22 | 2.14 | 0.95 | 2.27 | 0.023606* |
| A3.28 | JDt_32 | 2.94 | 0.95 | 3.09 | 0.002088** |
| A3.28 | JDt32 | 2.27 | 0.97 | 2.35 | 0.019183* |
| A3.28 | JuliD32 | 0.67 | 0.37 | 1.82 | 0.069738. |
| A3.28 | AugD12 | 0.53 | 0.37 | 1.43 | 0.15441 |
| A3.28 | AugD32 | 0.52 | 0.36 | 1.43 | 0.151784 |
| A3.28 | AugD42 | 0.71 | 0.36 | 1.96 | 0.05038. |
| A3.29 | (Intercept) | 0.88 | 1.94 | 0.45 | 0.65 |
| A3.29 | lag3 | 0.07 | 0.03 | 2.39 | 0.017* |
| A3.29 | lag4 | 0.04 | 0.03 | 1.43 | 0.1546 |
| A3.29 | lag5 | 0.07 | 0.03 | 2.38 | 0.0177* |
| A3.29 | yS | 0.39 | 0.20 | 2.01 | 0.0448* |
| A3.29 | Slag1 | -0.59 | 0.23 | -2.51 | 0.0122* |
| A3.29 | Slag2 | 0.40 | 0.19 | 2.05 | 0.0403* |
| A3.29 | Normal | -1.68 | 1.08 | -1.55 | 0.1207 |
| A3.29 | MDF2 | 1.24 | 0.49 | 2.51 | 0.0122* |
| A3.29 | EDt_22 | 1.43 | 0.98 | 1.46 | 0.1447 |
| A3.29 | MDt_12 | -2.01 | 1.00 | -2.02 | 0.044* |
| A3.29 | MDt12 | -2.40 | 1.30 | -1.85 | 0.0643. |
| A3.29 | MDt22 | -2.51 | 1.30 | -1.93 | 0.0534. |
| A3.29 | JDt2 | 0.93 | 0.59 | 1.58 | 0.1146 |
| A3.29 | JDt_22 | 1.62 | 0.98 | 1.66 | 0.0972. |
| A3.29 | JanD22 | 2.21 | 0.98 | 2.26 | 0.024* |
| A3.29 | JanD52 | 2.27 | 0.98 | 2.31 | 0.0209* |
| A3.29 | JanD62 | 1.47 | 0.98 | 1.50 | 0.1335 |
| A3.30 | (Intercept) | -0.85 | 2.16 | -0.40 | 0.69299 |
| A3.30 | lag1 | -0.08 | 0.03 | -2.66 | 0.00794** |
| A3.30 | yS | 0.42 | 0.19 | 2.16 | 0.03132* |
| A3.30 | Slag4 | -0.39 | 0.19 | -2.05 | 0.0406* |

| | | | | | |
|---|---|---|---|---|---|
| A3.30 | Slag6 | 0.64 | 0.20 | 3.23 | 0.00127** |
| A3.30 | Normal | -1.94 | 1.25 | -1.55 | 0.12164 |
| A3.30 | Central | -2.71 | 1.36 | -2.00 | 0.04623* |
| A3.30 | EDt12 | -1.57 | 0.99 | -1.59 | 0.11184 |
| A3.30 | EDt22 | 1.55 | 0.98 | 1.58 | 0.1151 |
| A3.30 | MDt12 | 2.36 | 1.20 | 1.97 | 0.04895* |
| A3.30 | JDt32 | 2.84 | 1.02 | 2.80 | 0.00523** |
| A3.30 | JanD12 | 2.91 | 0.98 | 2.97 | 0.00305** |
| A3.30 | JanD22 | 1.47 | 0.98 | 1.50 | 0.13287 |
| A3.30 | JanD62 | 1.47 | 0.98 | 1.50 | 0.1347 |
| A3.30 | JanD72 | 2.45 | 0.98 | 2.50 | 0.01256* |
| A3.30 | JanD82 | 1.76 | 0.98 | 1.79 | 0.07308. |
| A3.30 | JuliD12 | 1.09 | 0.38 | 2.84 | 0.00459** |
| A3.30 | JuliD22 | 1.05 | 0.39 | 2.70 | 0.00695** |
| A3.30 | JuliD32 | 1.30 | 0.40 | 3.28 | 0.00106** |
| A3.30 | JuliD42 | 0.97 | 0.41 | 2.37 | 0.01792* |
| A3.30 | AugD12 | 0.99 | 0.38 | 2.60 | 0.00956** |
| A3.30 | AugD22 | 0.80 | 0.39 | 2.07 | 0.0384* |
| A4.4 | (Intercept) | -13.29 | 9.31 | -1.43 | 0.153623 |
| A4.4 | lag3 | 0.08 | 0.03 | 2.70 | 0.007137** |
| A4.4 | lag6 | 0.06 | 0.03 | 1.97 | 0.048992* |
| A4.4 | yS | 0.55 | 0.25 | 2.20 | 0.028088* |
| A4.4 | Slag1 | -0.83 | 0.25 | -3.33 | 0.000906*** |
| A4.4 | Normal | 20.08 | 9.50 | 2.11 | 0.034811* |
| A4.4 | Central | 21.77 | 10.11 | 2.15 | 0.031598* |
| A4.4 | Week2 | 0.22 | 0.11 | 1.90 | 0.058369. |
| A4.4 | MDF2 | 1.17 | 0.51 | 2.29 | 0.022401* |
| A4.4 | EDF2 | -1.96 | 0.46 | -4.31 | 0.0000182*** |
| A4.4 | EDt2 | 1.47 | 0.66 | 2.21 | 0.027286* |
| A4.4 | EDt_22 | 2.67 | 1.11 | 2.41 | 0.016332* |
| A4.4 | EDt_32 | 2.54 | 1.11 | 2.29 | 0.022042* |
| A4.4 | EDt12 | 2.56 | 1.11 | 2.31 | 0.021202* |
| A4.4 | MDt2 | -1.43 | 0.88 | -1.62 | 0.106671 |
| A4.4 | MDt22 | -2.39 | 1.34 | -1.79 | 0.074384. |
| A4.4 | JDt12 | 1.44 | 1.01 | 1.43 | 0.154467 |
| A4.4 | JanD12 | -2.15 | 1.02 | -2.11 | 0.035418* |
| A4.4 | JanD32 | -1.93 | 1.01 | -1.91 | 0.056831. |
| A4.4 | JanD52 | 1.70 | 1.01 | 1.68 | 0.093397. |
| A4.2 | (Intercept) | 25.70 | 5.60 | 4.59 | 0.000005*** |
| A4.2 | lag1 | -0.09 | 0.03 | -2.91 | 0.00365** |
| A4.2 | lag2 | -0.08 | 0.03 | -2.49 | 0.01283* |
| A4.2 | lag3 | -0.05 | 0.03 | -1.49 | 0.13621 |
| A4.2 | yS | 0.78 | 0.25 | 3.19 | 0.00149** |
| A4.2 | Slag1 | -1.38 | 0.26 | -5.29 | 0.000000147*** |
| A4.2 | Slag2 | 0.68 | 0.24 | 2.83 | 0.0047** |
| A4.2 | Slag4 | -0.78 | 0.27 | -2.93 | 0.00351** |
| A4.2 | Slag5 | -0.84 | 0.28 | -2.94 | 0.00336** |
| A4.2 | Slag6 | 0.78 | 0.25 | 3.19 | 0.00145** |
| A4.2 | Normal | -9.72 | 4.89 | -1.99 | 0.04695* |
| A4.2 | Central | -12.51 | 5.46 | -2.29 | 0.02208* |
| A4.2 | CDF2 | 0.55 | 0.14 | 4.02 | 0.0000618*** |
| A4.2 | EDt_32 | -0.86 | 0.57 | -1.51 | 0.13083 |

| | | | | | |
|---|---|---|---|---|---|
| A4.2 | MDt_12 | 1.40 | 0.52 | 2.70 | 0.00698** |
| A4.2 | MDt12 | 1.48 | 0.72 | 2.06 | 0.03931* |
| A4.2 | JDt_42 | -1.01 | 0.59 | -1.73 | 0.08395. |
| A4.2 | JDt12 | -1.49 | 0.68 | -2.20 | 0.02786* |
| A4.2 | JDt22 | -1.74 | 0.67 | -2.61 | 0.00911** |
| A4.2 | JanD22 | 0.92 | 0.60 | 1.54 | 0.12484 |
| A4.2 | JanD32 | -1.18 | 0.59 | -2.00 | 0.04599* |
| A4.2 | JuliD12 | 0.64 | 0.24 | 2.70 | 0.0071** |
| A4.2 | JuliD22 | 0.55 | 0.25 | 2.19 | 0.02897* |
| A4.2 | JuliD32 | 0.44 | 0.26 | 1.72 | 0.08659. |
| A4.2 | AugD12 | 0.57 | 0.24 | 2.36 | 0.01857* |
| A4.1 | (Intercept) | -3.61 | 2.33 | -1.55 | 0.121997 |
| A4.1 | lag1 | 0.08 | 0.03 | 2.78 | 0.005616** |
| A4.1 | lag2 | 0.15 | 0.03 | 4.91 | 0.00000107*** |
| A4.1 | lag3 | 0.13 | 0.03 | 4.18 | 0.0000311*** |
| A4.1 | lag4 | 0.11 | 0.03 | 3.46 | 0.000553*** |
| A4.1 | lag6 | 0.06 | 0.03 | 1.95 | 0.051552. |
| A4.1 | yS | 0.73 | 0.13 | 5.50 | 0.0000000481*** |
| A4.1 | Slag1 | 0.22 | 0.13 | 1.69 | 0.091881. |
| A4.1 | Slag4 | -0.47 | 0.12 | -4.04 | 0.0000585*** |
| A4.1 | Normal | 1.35 | 0.90 | 1.50 | 0.13494 |
| A4.1 | EDt_12 | -0.86 | 0.56 | -1.55 | 0.121667 |
| A4.1 | MDt2 | 0.89 | 0.40 | 2.24 | 0.025588* |
| A4.1 | JanD42 | -0.88 | 0.56 | -1.57 | 0.116245 |
| A4.1 | JanD92 | -0.79 | 0.56 | -1.43 | 0.152985 |
| A4.1 | JanD102 | -0.82 | 0.56 | -1.47 | 0.141108 |
| A4.1 | JuliD12 | 0.50 | 0.22 | 2.28 | 0.022573* |
| A4.1 | JuliD22 | 0.56 | 0.22 | 2.50 | 0.012715* |
| A4.1 | JuliD42 | 0.70 | 0.23 | 3.05 | 0.002346** |
| A4.1 | AugD22 | -0.38 | 0.22 | -1.73 | 0.083214. |
| A4.5 | (Intercept) | 1.32 | 2.67 | 0.49 | 0.62152 |
| A4.5 | lag6 | -0.05 | 0.03 | -1.68 | 0.09369. |
| A4.5 | yS | 0.93 | 0.23 | 4.07 | 0.0000497*** |
| A4.5 | Slag4 | -0.46 | 0.23 | -2.01 | 0.04445* |
| A4.5 | Normal | -3.27 | 1.97 | -1.66 | 0.09743. |
| A4.5 | Central | -4.25 | 2.41 | -1.77 | 0.07785. |
| A4.5 | MDF2 | 1.05 | 0.42 | 2.50 | 0.01275* |
| A4.5 | EDF2 | -0.60 | 0.33 | -1.85 | 0.06454. |
| A4.5 | EDt12 | 2.36 | 1.12 | 2.11 | 0.03489* |
| A4.5 | JDt12 | 2.41 | 1.08 | 2.23 | 0.02628* |
| A4.5 | JDt22 | -1.53 | 1.08 | -1.41 | 0.15759 |
| A4.5 | JDt32 | -1.89 | 1.08 | -1.76 | 0.07925. |
| A4.5 | JanD12 | -1.76 | 1.07 | -1.64 | 0.10059 |
| A4.5 | JanD42 | -1.89 | 1.08 | -1.76 | 0.07949. |
| A4.5 | JuliD32 | 1.15 | 0.42 | 2.74 | 0.00632** |
| A4.3 | (Intercept) | -3.47 | 2.24 | -1.55 | 0.121192 |
| A4.3 | lag4 | 0.05 | 0.03 | 1.71 | 0.086827. |
| A4.3 | lag5 | 0.14 | 0.03 | 4.88 | 0.00000125*** |
| A4.3 | lag6 | 0.05 | 0.03 | 1.52 | 0.127964 |
| A4.3 | yS | 0.79 | 0.19 | 4.25 | 0.0000237*** |
| A4.3 | Slag1 | -0.47 | 0.19 | -2.51 | 0.012285* |
| A4.3 | Normal | 4.11 | 1.09 | 3.77 | 0.00017*** |

| | | | | | |
|---|---|---|---|---|---|
| A4.3 | MDF2 | 3.74 | 1.43 | 2.62 | 0.008912** |
| A4.3 | CDF2 | 0.51 | 0.29 | 1.79 | 0.07391. |
| A4.3 | EDt_42 | 2.05 | 1.17 | 1.75 | 0.080973. |
| A4.3 | MDt2 | -3.58 | 1.65 | -2.18 | 0.029622* |
| A4.3 | MDt_12 | -3.13 | 1.74 | -1.79 | 0.073388. |
| A4.3 | MDt_22 | -2.98 | 1.75 | -1.70 | 0.089183. |
| A4.3 | MDt12 | -3.86 | 2.02 | -1.91 | 0.05688. |
| A4.3 | MDt22 | -4.30 | 2.02 | -2.13 | 0.033376* |
| A4.3 | JDt_12 | -4.14 | 1.22 | -3.40 | 0.000703*** |
| A4.3 | NyåraftD2 | 1.75 | 1.24 | 1.41 | 0.160317 |
| A4.3 | JanD12 | -2.56 | 1.20 | -2.13 | 0.033669* |
| A4.3 | JanD62 | -3.27 | 1.20 | -2.72 | 0.006658** |
| A4.3 | JanD72 | -1.70 | 1.20 | -1.41 | 0.158033 |
| A4.3 | JanD92 | -5.06 | 1.20 | -4.22 | 0.0000264*** |
| A4.3 | JanD102 | -1.97 | 1.20 | -1.65 | 0.100243 |
| A4.3 | AugD12 | 0.63 | 0.44 | 1.42 | 0.156213 |
| A5.1 | (Intercept) | 2.95 | 2.35 | 1.26 | 0.209396 |
| A5.1 | lag1 | 0.06 | 0.03 | 1.88 | 0.060151. |
| A5.1 | lag2 | 0.06 | 0.03 | 2.06 | 0.039539* |
| A5.1 | lag3 | 0.08 | 0.03 | 2.67 | 0.007737** |
| A5.1 | lag4 | 0.08 | 0.03 | 2.79 | 0.005443** |
| A5.1 | lag5 | 0.10 | 0.03 | 3.45 | 0.000582*** |
| A5.1 | lag6 | 0.08 | 0.03 | 2.61 | 0.009325** |
| A5.1 | yS | 0.29 | 0.07 | 3.88 | 0.000112*** |
| A5.1 | Slag3 | -0.18 | 0.08 | -2.11 | 0.035255* |
| A5.1 | Slag4 | -0.16 | 0.09 | -1.70 | 0.089774. |
| A5.1 | Slag5 | -0.26 | 0.08 | -3.14 | 0.001727** |
| A5.1 | Normal | 5.05 | 2.11 | 2.39 | 0.016881* |
| A5.1 | Central | 4.53 | 2.64 | 1.72 | 0.086564. |
| A5.1 | MDF2 | 0.34 | 0.11 | 2.98 | 0.002977** |
| A5.1 | EDF2 | -0.17 | 0.10 | -1.70 | 0.089006. |
| A5.1 | EDt_12 | 0.68 | 0.30 | 2.24 | 0.025396* |
| A5.1 | EDt_22 | 0.68 | 0.30 | 2.25 | 0.024717* |
| A5.1 | EDt_42 | 0.46 | 0.30 | 1.52 | 0.128526 |
| A5.1 | EDt12 | 0.67 | 0.30 | 2.21 | 0.027132* |
| A5.1 | JDt2 | 0.80 | 0.17 | 4.58 | 0.00000525*** |
| A5.1 | JDt_22 | 0.48 | 0.29 | 1.68 | 0.09421. |
| A5.1 | JuliD12 | 0.25 | 0.11 | 2.25 | 0.024779* |
| A5.1 | JuliD22 | 0.23 | 0.12 | 1.98 | 0.048547* |
| A5.1 | AugD22 | -0.32 | 0.11 | -2.86 | 0.004269** |
| A5.1 | AugD32 | -0.20 | 0.11 | -1.78 | 0.0749. |

*Table A 6 Residuals for all the 37 forecasts*

| **A1.1** | **Grilled chicken - Meat** | | | |
|---|---|---|---|---|
| | **Regression** | **ES(A,N,A)** | **Naive** | **Combination** |
| **MAE** | **0.93** | 1.20 | 1.50 | 1.03 |
| **RMSE** | 1.926 | 1.848 | 2.308 | **1.835** |
| **ME** | 0.729 | **-0.045** | -0.21 | 0.342 |

| **A3.1** | **Grilled chicken - Meat** | | | |
|---|---|---|---|---|
| | **Regression** | **ES(A,N,A)** | **Naive** | **Combination** |
| **MAE** | **1.15** | 1.02 | 1.26 | 1.00 |
| **RMSE** | 1.959 | 1.67 | 2.001 | **1.737** |
| **ME** | 1.019 | **0.052** | 0.466 | 0.535 |

| **A3.2** | **Swedish chicken - Meat** | | | |
|---|---|---|---|---|
| | **Regression** | **ES(A,N,A)** | **Naive** | **Combination** |
| **MAE** | **1.32** | 1.44 | 2.49 | 1.26 |
| **RMSE** | 2.436 | 2.227 | 3.825 | **2.237** |
| **ME** | 0.875 | **-0.156** | -1.325 | 0.359 |

| **A3.3** | **Fresh fish manual - Fish** | | | |
|---|---|---|---|---|
| | **Regression** | **ES(A,N,A)** | **Naive** | **Combination** |
| **MAE** | **1.14** | 1.34 | 1.88 | 1.19 |
| **RMSE** | 2.806 | 2.585 | 3.682 | **2.649** |
| **ME** | 1.089 | **0.138** | -0.147 | 0.613 |

| **A3.4** | **Swedish beef central packaged - Meat** | | | |
|---|---|---|---|---|
| | **Regression** | **ES(A,N,A)** | **Naive** | **Combination** |
| **MAE** | **0.76** | 0.84 | 1.14 | 0.77 |
| **RMSE** | 1.348 | 1.152 | 1.728 | **1.193** |
| **ME** | 0.623 | **-0.039** | -0.304 | 0.292 |

| **A3.5** | **Fresh shell fish - Fish** | | | |
|---|---|---|---|---|
| | **Regression** | **ES(A,N,N)** | **Naive** | **Combination** |
| **MAE** | **1.09** | 1.38 | 1.78 | 1.17 |
| **RMSE** | 2.246 | 2.048 | 3.467 | **2.078** |
| **ME** | 1.006 | **-0.061** | 0.08 | 0.472 |

| **A3.6** | **Fresh fish packaged - Fish** | | | |
|---|---|---|---|---|
| | **Regression** | **ES(A,N,N)** | **Naive** | **Combination** |
| **MAE** | **1.16** | 1.83 | 1.38 | 1.49 |
| **RMSE** | 3.176 | 3 | 3.131 | **3.032** |
| **ME** | 1.092 | **-0.081** | 0.751 | 0.506 |

| A3.7 | Swedish beef - Meat | | | |
|---|---|---|---|---|
| | Regression | ES(A,N,N) | Naive | Combination |
| MAE | **0.81** | 1.09 | 1.38 | 0.94 |
| RMSE | 1.67 | 1.453 | 2.001 | **1.481** |
| ME | 0.724 | **-0.179** | -0.486 | 0.273 |

| A3.8 | Fresh Swedish CPK - Meat | | | |
|---|---|---|---|---|
| | Regression | ES(A,N,N) | Naive | Combination |
| MAE | **1.21** | 1.73 | 1.57 | 1.39 |
| RMSE | 2.929 | 2.976 | 2.999 | **2.849** |
| ME | 1.043 | **-0.198** | 0.192 | 0.423 |

| A3.9 | Sour milk high fat - Dairy | | | |
|---|---|---|---|---|
| | Regression | ES(A,N,A) | Naive | Combination |
| MAE | **0.68** | 0.79 | 1.32 | 0.69 |
| RMSE | 1.023 | 1 | 1.901 | **0.94** |
| ME | 0.37 | **-0.179** | -0.688 | 0.095 |

| A3.10 | Swedish pork - Meat | | | |
|---|---|---|---|---|
| | Regression | ES(A,N,N) | Naive | Combination |
| MAE | **1.60** | 2.13 | 2.40 | 1.79 |
| RMSE | 3.446 | 3.139 | 4.19 | **3.136** |
| ME | 1.484 | **-0.268** | 0.111 | 0.608 |

| A3.11 | Smoked/cured fish - Fish | | | |
|---|---|---|---|---|
| | Regression | ES(A,N,A) | Naive | Combination |
| MAE | **1.46** | 2.00 | 2.36 | 1.70 |
| RMSE | 3.505 | 3.277 | 4.219 | **3.297** |
| ME | 1.369 | **0.056** | -0.017 | 0.712 |

| A3.12 | Fresh hamburger meat - Meat | | | |
|---|---|---|---|---|
| | Regression | ES(A,N,N) | Naive | Combination |
| MAE | **1.77** | 1.87 | 1.97 | 1.73 |
| RMSE | 3.572 | 3.158 | 3.449 | **3.294** |
| ME | 1.676 | **0.3** | 0.448 | 0.988 |

| A3.13 | Swedish pork central packaged - Meat | | | |
|---|---|---|---|---|
| | Regression | ES(A,N,A) | Naive | Combination |
| MAE | **1.36** | 1.41 | 2.67 | 1.30 |
| RMSE | 2.302 | 2.024 | 3.738 | **2.064** |
| ME | 1.043 | **-0.082** | -1.131 | 0.48 |

**A3.14 Milk high fat - Dairy**

| | Regression | ES(A,N,A) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **0.36** | 0.48 | 0.55 | 0.40 |
| **RMSE** | 0.684 | 0.699 | 0.962 | **0.658** |
| **ME** | 0.163 | **-0.085** | -0.045 | 0.039 |

**A3.15 Big package flavored - Dairy**

| | Regression | ES(A,N,A) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **0.74** | 0.89 | 2.10 | 0.75 |
| **RMSE** | 1.245 | 1.213 | 3.289 | **1.149** |
| **ME** | 0.322 | **-0.234** | -1.587 | 0.044 |

**A3.16 Fresh Swedish - Meat**

| | Regression | ES(A,N,A) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **1.13** | 1.40 | 1.65 | 1.22 |
| **RMSE** | 2.568 | 2.338 | 2.968 | **2.394** |
| **ME** | 0.995 | **0.012** | 0.047 | 0.504 |

**A3.17 Cooled drink to go - Dairy**

| | Regression | ES(A,N,A) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **1.27** | 1.52 | 2.10 | 1.31 |
| **RMSE** | 3.402 | 3.258 | 3.764 | **3.268** |
| **ME** | 0.972 | **-0.055** | -0.621 | 0.458 |

**A3.18 Milk medium fat - Dairy**

| | Regression | ES(A,N,N) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **1.68** | 2.04 | 2.22 | 1.78 |
| **RMSE** | 7.018 | 6.935 | 7.282 | **6.956** |
| **ME** | 1.293 | **0.427** | 0.569 | 0.86 |

**A3.19 Cottage cheese natural - Cheese**

| | Regression | ES(A,N,A) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **0.91** | 1.07 | 1.72 | 0.93 |
| **RMSE** | 2.061 | 1.949 | 2.869 | **1.966** |
| **ME** | 0.611 | **0.109** | -0.553 | 0.36 |

**A3.20 Imported pork - Meat**

| | Regression | ES(A,N,N) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **1.39** | 1.64 | 2.21 | 1.42 |
| **RMSE** | 2.427 | 2.191 | 3.282 | **2.195** |
| **ME** | 1.181 | **-0.168** | -0.776 | 0.507 |

**A3.21 Dessert mold cheese - Cheese**

| | Regression | ES(A,N,N) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **0.20** | 1.03 | 0.84 | 0.61 |
| **RMSE** | 0.744 | 1.214 | 2.017 | **0.849** |
| **ME** | 0.123 | **-0.844** | -0.527 | -0.36 |

**A3.22 Milk low fat - Dairy**

| | Regression | ES(A,N,A) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **0.47** | 0.70 | 0.54 | 0.55 |
| **RMSE** | 0.949 | 1.054 | 1.05 | **0.963** |
| **ME** | 0.1 | **-0.08** | 0.043 | 0.01 |

**A3.23 N/A - Dairy**

| | Regression | ES(A,N,A) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **0.74** | 0.94 | 1.63 | 0.80 |
| **RMSE** | 1.46 | 1.43 | 2.601 | **1.381** |
| **ME** | 0.509 | **-0.188** | -0.865 | 0.16 |

**A3.24 Cream - Dairy**

| | Regression | ES(A,N,N) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **0.89** | 1.26 | 1.16 | 1.01 |
| **RMSE** | 1.914 | 1.829 | 2.019 | **1.784** |
| **ME** | 0.736 | **-0.261** | 0.083 | 0.237 |

**A3.25 Egg from free-range hens - Meat**

| | Regression | ES(A,N,N) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **1.42** | 1.70 | 2.28 | 1.49 |
| **RMSE** | 6.636 | 6.516 | 6.993 | **6.546** |
| **ME** | 1.229 | **0.164** | -0.232 | 0.697 |

**A3.26 NFC Drink later - Dairy**

| | Regression | ES(A,N,N) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **0.98** | 1.34 | 1.98 | 1.15 |
| **RMSE** | 2.686 | 2.598 | 4.037 | **2.61** |
| **ME** | 0.788 | **0.011** | -0.569 | 0.399 |

**A3.27 NFC Drink now - Dairy**

| | Regression | ES(A,N,A) | Naive | Combination |
|---|---|---|---|---|
| **MAE** | **0.89** | 1.17 | 1.96 | 0.98 |
| **RMSE** | 1.829 | 1.683 | 3.08 | **1.694** |
| **ME** | 0.756 | **-0.106** | -0.568 | 0.325 |

**A3.28 Brine products - Meat**

| | Regression | ES(A,N,N) | Naive | Combination |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| **MAE** | **1.64** | 2.32 | 2.00 | 1.98 |
| **RMSE** | 6.985 | 6.831 | 7.077 | **6.877** |
| **ME** | 1.562 | **0.264** | 1.143 | 0.913 |

**A3.29 Mozzarella - Cheese**

| | **Regression** | **ES(A,N,N)** | **Naive** | **Combination** |
|---|---|---|---|---|
| **MAE** | **1.16** | 1.43 | 1.69 | 1.26 |
| **RMSE** | 3.106 | 2.941 | 3.249 | **2.988** |
| **ME** | 1.032 | **0.212** | 0.158 | 0.622 |

**A3.30 Sour milk low fat - Dairy**

| | **Regression** | **ES(A,N,A)** | **Naive** | **Combination** |
|---|---|---|---|---|
| **MAE** | **0.82** | 1.07 | 1.25 | 0.91 |
| **RMSE** | 1.551 | 1.51 | 2.398 | **1.466** |
| **ME** | 0.612 | **-0.103** | 0.047 | 0.255 |

**A4.4 Category Cheese**

| | **Regression** | **ES(A,N,N)** | **Naive** | **Combination** |
|---|---|---|---|---|
| **MAE** | **0.82** | 0.86 | 1.20 | 0.76 |
| **RMSE** | 1.311 | 1.148 | 1.644 | **1.161** |
| **ME** | 0.684 | **-0.122** | -0.278 | 0.281 |

**A4.2 Category Dairy**

| | **Regression** | **ES(A,N,A)** | **Naive** | **Combination** |
|---|---|---|---|---|
| **MAE** | **0.49** | 0.55 | 0.89 | 0.49 |
| **RMSE** | 0.836 | 0.822 | 1.176 | **0.801** |
| **ME** | 0.183 | **-0.049** | -0.368 | 0.067 |

**A4.1 Category Meat**

| | **Regression** | **ES(A,N,A)** | **Naive** | **Combination** |
|---|---|---|---|---|
| **MAE** | **0.66** | 0.67 | 0.92 | 0.64 |
| **RMSE** | 0.949 | 0.932 | 1.232 | **0.907** |
| **ME** | 0.172 | **-0.076** | -0.206 | 0.048 |

**A4.5 Category Vegetarian**

| | **Regression** | **ES(A,N,A)** | **Naive** | **Combination** |
|---|---|---|---|---|
| **MAE** | **0.93** | 1.20 | 1.50 | 1.03 |
| **RMSE** | 1.926 | 1.848 | 2.308 | **1.835** |
| **ME** | 0.729 | **-0.045** | -0.21 | 0.342 |

**A4.3 Category Fish**

| | **Regression** | **ES(A,N,A)** | **Naive** | **Combination** |
|---|---|---|---|---|
| **MAE** | **1.01** | 0.89 | 1.23 | 0.87 |

| | | | | |
|---|---|---|---|---|
| **RMSE** | 1.591 | 1.276 | 1.761 | **1.36** |
| **ME** | 0.933 | **0.013** | -0.018 | 0.473 |

**A5.1 All categories**

| | **Regression** | **ES(M,N,A)** | **Naive** | **Combination** |
|---|---|---|---|---|
| **MAE** | **0.45** | 0.43 | 0.62 | 0.43 |
| **RMSE** | 0.579 | 0.547 | 0.797 | **0.546** |
| **ME** | 0.097 | **-0.04** | -0.226 | 0.029 |

**Packages Used in R:**

**Apart from the baseline commands included in R, the below packages were used for analysis and data handling.**

## 1. TSutils

**Description**:

The tsutils package provides functions to support various aspects of time series and forecasting modelling. In particular this package includes: (i) tests and visualizations that can help the modeller explore time series components and perform decomposition; (ii) modelling shortcuts, such as functions to construct lag matrices and seasonal dummy variables of various forms; (iii) an implementation of the Theta method; (iv) tools to facilitate the design of the forecasting process, such as ABC-XYZ analyses; and (v) "quality of life" tools, such as treating time series for trailing and leading values.

**Author**:

Nikolaos Kourentzes

## 2. Forecast

**Description:**

Forecast is a collection of many different functions and packages. The packages are used for methods and tools for displaying and analyzing univariate time series forecasts including exponential smoothing via state space models and automatic ARIMA modelling. Several functions, such as ETS are used in Forecast in order to create the models used in this thesis.

**Author**:

The packages have multiple authors. For specifics visit the information page in Rstudio or the webpage https://cran.r-project.org/web/packages/forecast/forecast.pdf (2020-05-12).

## 3. Tidyverse

**Description**:

Tidyverse is a collection of packages that work in harmony and are used for data exploration, analysis and structuring (forecast and dyplr for example). The Tidyverse package allows these to be easily installed and loaded in one simple step.

**Author**:

The packages have multiple authors, the current maintainer of Tidyverse is Hadley Wickham

## 4. Dplyr

**Description:**

Dplyr is a package that provides flexible grammar of data manipulation. It is an iteration of plyr and is focused on tools for working in data frame, hence the d- at the beginning. It makes data manipulation verbs easier to use in R and speeds up many processes through better programming.

**Author**:

The package is co-authored by Romain François, Lionel Henry and Kirill Müller. Rstudio also contributed. Hadley Wickham is the current maintainer.