Applied Economics

Reducing Product Waste within the Retail Industry: A Post-COVID-19 Era Study on Enhancing Demand Prediction with Hybrid Prediction Models.

Submission ID	247732722
Article Type	Research Article
Keywords	Forecasting, Hybrid demand prediction model, Machine learning, Food waste, Product Waste
Authors	Maider Elorza, Sara Pilar Segura Querol, Eduardo Castellano

Article title

Reducing Product Waste within the Retail Industry: A Post-COVID-19 Era Study on Enhancing Demand Prediction with Hybrid Prediction Models.

Author information

Maider Elorza[a], Sara Segura[b*], Eduardo Castellano[b]

[a]*Advanced Strategies in Business Management Department, Mondragon Unibertsitatea Business Faculty, Oñati, Spain.*

[b]*Data and Innovation Department, Mondragon Unibertsitatea Business Faculty, Oñati, Spain.*

PhD. Maider Elorza.
ORCiD: 0000-0002-7040-2086
Postal address: Ibarra Zelaia, 2, 20560 OÑATI Gipuzkoa, Spain
Email address: melorzai@mondragon.edu

PhD. Sara Segura.
ORCiD: 0000-0002-5854-1024
Postal address: Ibarra Zelaia, 2, 20560 OÑATI Gipuzkoa, Spain
Email address: spsegura@mondragon.edu *corresponding author

PhD. Eduardo Castellano.
ORCiD: 0000-0002-8533-1874
Postal address: Ibarra Zelaia, 2, 20560 OÑATI Gipuzkoa, Spain
Email address: ecastellano@mondragon.edu

Abstract

The retail sector encounters significant challenges, prominently among them being the European Union's sustainable policies aimed at diminishing waste from food and products. In this context, the objective of this paper is to introduce a framework that addresses two pivotal goals: i) Introducing cutting-edge machine learning models tailored to forecast demand within the context of a post-COVID environment, thereby enhancing the accuracy of optimal inventory level calculations. ii) Evaluating the benefits of integrating these predictive models into operational strategies by measuring the reduction in overstock levels compared to traditional business practices. Regarding demand prediction, two main conclusions are reached. On the one hand, the hybrid Prophet-XGBoost model consistently outperformed other hybridisations in terms of accuracy, achieving the lowest MAPE and WAPE across all seven categories. On the other hand, our results emphasised the significance of historical imputation approaches when dealing with COVID-19 data. More specifically, using the 2019 data as a proxy for the data from 2021 consistently yielded lower errors across categories. In addition to its contributions to the machine learning field, this paper enriches the literature by introducing a novel method for calculating overstock reduction.

Keywords

Forecasting; Hybrid demand prediction model; Machine Learning; Food waste; Product waste; Retail analytics

1. Introduction

Retail stands as the paramount industrial ecosystem within the European Union, constituting a staggering 11.5% of value added and providing employment to approximately 30 million individuals (European Commission, 2024). Within this expansive sector, the significance of food and beverage cannot be overstated, contributing a substantial €185 billion in added value and generating revenues of €1,117 billion in the fiscal year of 2023 (FoodDrinkEurope, 2023).

The retail sector encounters significant challenges, prominently among them being the European Union’s sustainable policies aimed at diminishing waste from food and products (European Commission, 2023). Particularly concerning food waste, the

European Commission in July 2023 proposed legally binding reduction targets for Member States to achieve by 2030.

In this context, addressing the pressing need to mitigate the issue of overstock is paramount, given its repercussions both at the financial and sustainable level, when stored goods vastly exceed actual demand. Product wastage persists at a higher frequency than deemed acceptable. The retail and food distribution sectors alone contribute to a staggering excess of 4 million tonnes of food waste across the European Union, underscoring the urgency for effective solutions (Eurostat, 2023). Given the importance of this issue, it is gaining increasing attention in research (Miguéis et al., 2022; Rodrigues et al., 2024; Tait et al., 2024).

Retail companies work to minimise overstock while simultaneously aiming to prevent out-of-stock situations. The latter, where they fail to meet customer demands, not only diminishes customer satisfaction but also results in significant financial losses (Caro et al., 2020). Therefore, achieving an optimal inventory (OI) balance, stands as the cornerstone for retail operations (Deng & Liu, 2021; Perez et al., 2021).

Companies conduct daily calculations in advance to determine the OI levels for each day. Demand prediction is the core component of this calculation process. The precision of this forecast significantly impacts the accuracy of the OI calculation, ultimately determining the presence of overstock and consequent product wastage (Miguéis et al., 2022; Rodrigues et al., 2024).

The objective of this paper is to introduce a framework that addresses two pivotal goals: i) Introducing cutting-edge machine learning models tailored to forecast demand within the context of a post-COVID environment, thereby enhancing the accuracy of optimal inventory level calculations. ii) Evaluating the benefits of integrating these predictive models into operational strategies by measuring the reduction in overstock levels compared to traditional business practices.

To validate the efficacy of our framework, we conduct an empirical case study within a prominent European retail company. Our analysis focuses on the most critical product categories housed within the retailer's largest warehouse, encompassing five food categories and two non-food categories.

It is evident from the literature that much researchers have been carried out on the various machine learning algorithms to predict demand. We used four prominent prediction models to accomplish this goal: Prophet, XGBoost, LightGBM, and Random Forest, which have become increasingly popular due to their effectiveness in these contexts. Specifically, we explore hybridisations of the Prophet model with XGBoost (following the approach proposed by Wang et al. in (2022)), with LightGBM (as recommended by Shakeel et al. (2023)), and with Random Forest (as outlined by Srinidhi et al. (2020)).

We possess a dataset spanning from the pre-COVID-19 era through to the post-COVID-19 landscape, crucial for accurate demand prediction. The inherent challenge lies in navigating the fluctuations in demand during the COVID-19 period, a time when customer behaviour changed due to the lockdowns. It is imperative to emphasise research dedicated to addressing this pivotal period, as highlighted by Fildes et al. (2022).

The following are the major contributions of this work:

- Advancement to the emerging literature exploring the efficacy of hybrid prediction models in terms of performance. Through an empirical study, we demonstrate that hybrid models outperform conventional approaches.

- Introduction of various value imputation strategies tailored for the COVID-19 period, evaluating their effectiveness in maintaining prediction accuracy. This ensures predictions remain unbiased by the effects of the pandemic, a unique aspect not addressed in existing literature.
- Expansion of the performance evaluation criteria for product forecasting beyond the typical focus on forecast quality. Our analysis includes quantifying overstock reduction and, consequently, minimising product waste, addressing a notable gap in current literature.

The remainder of this paper is organised as follows. Section 2 presents the literature review. Section 3 describes the employed methodology. Section 4 compares various machine learning models implemented for demand forecasting and quantifies the potential reduction in overstock. Finally, Section 5 provides conclusions and points towards directions for further research.

2. Literature review

The retail sector operates within a context where inventory management (IM) is supremely important, particularly in terms of product waste, due to European regulations. Demand prediction is crucial for making a realistic advance estimation of the optimal stock to hold in the warehouse. The efficacy of this forecast depends primarily on incorporating relevant variables and employing appropriate methodologies.

Factors such as the day of the week, time of year, and prominent holidays like Christmas significantly influence variations in retail demand, underscoring the importance of considering such variables in predictive models (Ehrenthal et al., 2014; Kramar &

Alchakov, 2023). Furthermore, macroeconomic events, like the COVID-19 pandemic, significantly influence the volatility of demand for retail products, impacting their seasonality (Deng & Liu, 2021; Kim et al., 2023). The abnormal demand behaviour during COVID-19 will impact future forecasts because they are based on historical data, so it is essential to work on strategies to deal with data belonging to that period (Roggeveen & Sethuraman, 2020). Madeira (2020) conducted a study that explored the impact of COVID-19 on household consumption patterns throughout the pandemic. Additionally, Kim et al. (2023) pointed out that, in 2020 and 2021, the COVID-19 pandemic had a varying effect on sales for different kinds of retail stores, particularly in the food and beverage (F&B) category. However, as stated by Fildes et al. (2022), nothing has yet been published focusing on the post-COVID-19 era. Our paper begins to shed light on the matter.

In addition to the ever-changing nature of customer demand, the rapid evolution of information technology and the increasing complexity of business operations have surpassed the capabilities of traditional IM methods (Deng & Liu, 2021). The turning point in the evolution from traditional IM to IM lies in the technology for predicting consumer demand. Prediction methods can be divided into two categories. The first relies on statistical analysis, such as traditional time series methodologies, and the second employs machine learning (ML) techniques. Statistics relies heavily on the quantity and quality of historical data, often struggling to keep up with the swift changes in today's customer demand. ML has consistently strong performance results, as reported by various researchers (Kramar & Alchakov, 2023; Srinivas & Katarya, 2022).

In this context, tree-based algorithms excel and are becoming more and more used not only in demand prediction problems (Auppakorn & Phumchusri, 2022; Pavlyshenko,

2019) but also in many fields, such as economics and finance (Y. Li et al., 2022; Yu & Huo, 2022), politics (Funk et al., 2022), biology and the environment (Aswad et al., 2021; Demir & Sahin, 2022), and medicine and healthcare (Ghiasi & Zendehboudi, 2021; Srinivas & Katarya, 2022).

Furthermore, some research studies have shown that ML models yield better outcomes when combined with other models, the so-called hybrid models (Wang et al., 2022). Conventional models rely heavily on stationary data for accurate predictions, and when that data has multiple seasonalities design becomes extremely difficult. Some authors authors, such as Shakeel et al. (2023) in the energy sector, Srinidhi et al. (2020) in the field of IoT communication, and Wang et al. (2022) in civil engineering have been developing hybrid models. Nevertheless, research in this domain in the retail industry is still in its nascent stage, leaving ample room for further enhancements in the accuracy of demand prediction in the field of IM.

Demand prediction serves as a critical factor in determining optimal inventory (OI) levels, a cornerstone of effective inventory management. In this investigation, we derive OI utilising the Economic Order Quantity (EOQ) model pioneered by Ford Harris (1915) and extensively applied by R. H. Wilson (1934), a methodology widely embraced in the field. The proficiency of retail managers in computing OI levels within their warehouses holds the potential to yield multifaceted benefits, including heightened customer satisfaction, minimised waste, increased product turnover, amplified sales revenue, and enhanced operational efficiency (Auppakorn & Phumchusri, 2022).

While numerous studies concentrate on forecasting accuracy, there exists a gap in assessing the concrete reduction in product waste attributed to these predictive models. Notably, while papers such as those authored by D. Li et al. (2023) and Ehrenthal et al.

(2014) quantify the cost savings derived from demand forecasting implementations, they primarily emphasise financial implications rather than delineating the intricacies of waste reduction calculations.

3. Methodology

3.1. Demand prediction

We used an experimental protocol to evaluate the prediction algorithm in which each time series was split into two parts at the same time point. The first part, ordered chronologically, was used as the training dataset for the model. The remaining days were used as the test dataset (the last 120 days for all the options) to assess the quality of the models (Kramar & Alchakov, 2023). Fig. 1 illustrates the flow of the procedure used for demand forecasting.

This procedure was previously utilised by Elorza et al. (2024), however, it only accounted for data within one category (milk) and Prophet and XGBoost model's combination.

[INSERT FIGURE 1 NEAR HERE]

3.1.1 Time series

Prophet is an open-source forecasting procedure for time series data based on an additive regression model where non-linear trends are fit with yearly, weekly, and daily seasonality, including holidays. The main idea of the method was presented by Taylor and Letham (2018). It works best with time series that have strong seasonal effects and several seasons of historical data. Prophet is robust when it comes to missing data and shifts in the trend, and it typically handles outliers well. Predicted parameter values can be found using the prediction model method, with a data frame of time stamps for the prediction horizon as input. At its core is the sum of three time functions plus an error term:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \tag{1}$$

where $g(t)$ is the trend function that models nonperiodic changes in the value of the time series, $s(t)$ represents periodic changes, $h(t)$ represents the effects of holidays that occur on potentially irregular schedules over one or more days, and ε_t represents any idiosyncratic changes that are not accommodated by the model, following a normal distribution with a mean of 0.

3.1.2 Machine learning predictive models

Three different ML models were adopted in this case study to generate the predictive model: XGB, LGBM and RF.

XGB (XGBoost, extreme gradient boosting) is a tree-based ML forecasting method developed by Chen and Guestrin (2016). This model is a more regularised form of gradient boosting decision trees that improves generalisation by adding regularisation terms to prevent overfitting. The objective is to find the optimal value that will minimise total loss (Sibindi et al., 2022). The most common loss function for regression problems is:

$$L^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \tag{2}$$

where $l(\cdot)$ represents a second-order derivable loss function, which measures the difference between the actual value y_i and the predicted value $\hat{y}_i^{(t-1)}$, and $\Omega(\cdot)$ denotes the regularisation term.

Then, by using the Taylor expansion of the loss function to the second order, the objective function can be simplified to a quadratic equation with $f_t(x_i)$.

$$L^{(t)} \cong \sum_{i=1}^{n} [l\ (y_i, \hat{y}_i{}^{(t-1)}) + g_i f_t(x_i) + \ \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \tag{3}$$

where $g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}_i{}^{(t-1)})$ and $h_i = \partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y}_i{}^{(t-1)})$ are the first and second order derivatives of the loss function, respectively. Because g_i and h_i can be used as constants, the equation (3) can be simplified as.

$$\tilde{L}^{(t)} = \sum_{i=1}^{n} [g_i f_t(x_i) + \ \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \tag{4}$$

LGBM (LightGBM), a data algorithm released by Microsoft in 2017, is based on gradient boosting decision trees. LGBM operates as a distributed gradient-boosting framework with open access. Its core concept involves sequentially merging M weak regression trees to form a robust regression tree (Pan et al., 2021). The formula below is used.

$$F(x) = \sum_{m=1}^{M} f_m(x) \tag{5}$$

This approach, in addition to minimising prediction accuracy errors, enhances prediction speed and reduces memory usage. Two notable advances in the LGBM model are the histogram method and the leaf-wise approach with controlled depth (Sibindi et al., 2022). The histogram method efficiently combines distinct features by dividing continuous feature values into M intervals before constructing a histogram. Data is navigated based on the discretised values in the histogram, resulting in improved computational efficiency, particularly with the use of decision trees, which are inherently weaker models. In the LGBM approach, a leaf-wise method is used to identify the leaf with the highest gain in dispersion among all existing leaves and then split it. To prevent overfitting, LGBM adds a maximum depth constraint to the leaf-wise strategy while maximising efficiency. The level-wise tree growth technique grows trees layer by layer, with each node splitting data while giving more weight to nodes closer to the tree's root.

RF (Random Forest), developed by Breiman (2001), is composed of multiple decision trees and performs random feature selection on each tree. Then it averages the output values of all individual trees to obtain the model's output (Lai et al., 2023). During the training process, the RF model constructs multiple decision trees using random subsets of the data. The class is assigned based on the mode. The RF algorithm uses the bootstrap aggregation technique, which enhances algorithm stability and accuracy while reducing variance and the risk of overfitting. Formally, consider a training set $X = x_1$, ..., x_n along with their corresponding outputs $Y = y_1$, ..., y_n giving dataset $D_n = (X_1,\ Y_1)$, ..., (X_n, Y_n). RF is a collection of random base regression trees $\{rn(x, \theta_m, D_n),\ m\ \geq 1\}$, where θ_1, θ_2, … are outputs of a randomising variable θ. The outputs of these random trees are aggregated to get an estimate. Thus, the RF model is given by,

$$\overline{rn}(X, D_n) = E_\theta[rn(X, \theta, D_n)] \tag{6}$$

where E_θ is the expectation for θ. The dataset is partitioned into various subspaces based on a randomising variable theta. This variable also guides the successive splitting of nodes when constructing individual trees. During the learning process the algorithm selects a random subset of features, a technique known as feature bagging. If a particular feature proves to be a strong predictor it may be used again in multiple trees. One of the primary advantages of RF over conventional decision trees is that predictions made by individual trees can be sensitive to noise. However, this limitation is overcome by aggregating the predictions from the entire forest.

3.1.3 Hybrid prediction models

This paper proposes three hybrid models to address the challenges of precise load forecasting and hyperparameter tuning.

The feature extraction property of the Prophet model helps obtain precise findings which are then fed into XGB, LGBM and RF models separately to train the hybrid model. The ML models were tuned using the Optuna hyperparameter approach to evaluate the best hyperparameter approach for the proposed models.

Efficient hyperparameter tuning is crucial for performance improvement. Optuna is a trending method known for achieving better optimisation results by finding the best combination of hyperparameters for a given model (Srinivas & Katarya, 2022). Traditional methods like manual search and random search can be inefficient and time consuming. Optuna stands out as it continuously learns from previous optimisations and efficiently prunes and samples the search space (Lai et al., 2023). There are five essential steps in determining hyperparameters for ML models in Optuna. The first step is to enter the hyperparameters of the machine learning models. The second step is determining the search ranges of hyperparameters and types, including integers and floats. The third step is to set the objective function for Optuna, as provided by the machine learning models (Table 1). Then optimisation directions are determined. Minimising forecasting errors was used as the direction and the objective function of this study. Lastly, the number of trials of Optuna is set. In the research the direction and n_trials are set to the minimum and 100, respectively (Lai et al., 2023).

[INSERT TABLE 1 NEAR HERE]

3.2 Optimal inventory and waste product reduction calculation

3.2.1 Setting up the problem

optimal inventory (OI) represents the exact amount of inventory needed in a warehouse at any given time to satisfy regular customer demand (Zeng et al., 2019) while avoiding product waste and unnecessary costs.

Companies frequently calculate OI in advance to optimise inventory management. Consequently, the accuracy of the forecast directly influences the realism of the OI calculation. In this study, we calculate the OI according to the Economic Order Quantity (EOQ) developed by Ford Harris (1915) and applied extensively by R. H. Wilson (1934), which will be explained in detail in subsection 3.2.2.

First, we calculate the Optimal Inventory (OI) based on actual demand, referred to as *OI_actual*. While this represents the ideal scenario, it is often not realistic due to the challenges of accurately predicting demand. Second, we determine the OI using the demand forecasts from the previous section, named as *OI_forecast*. Additionally, the retailer has provided us with the current inventory data, termed Current Inventory (*CI*). To assess the improvement of using our demand forecast, we measure the deviation between *OI_forecast* and *OI_actual* and compare it to the deviation between *CI* and *OI_actual*.

3.2.2 Optimal Inventory calculation

For month *m* the calculation of the OI level is obtained by summing the variables:

$$OI_m = MinI_m + SS_m + EOQ_m \tag{7}$$

Where $MinI$ is the minimal inventory that ensures that company does not run out of stock before the new order arrives, SS is the safety stock which is an additional quantity above the minimum stock to prevent stockouts. EOQ represent the optimal ordering quantity that minimises the total cost of an inventory system.

$MinI$ and SS are initially calculated on a daily basis. Subsequently, the daily values are aggregated to provide monthly totals.

$$MinI_t = \sum_{i=0}^{LT-1} \hat{d}_{t+i} \tag{8}$$

$MinI$ in day t is calculated by adding the predictions of daily demand ($\hat{d}_t$), depending on the Lead Time (the time between placing an order for a product with a supplier and the product's arrival at the retailer's warehouse), i.e. if the LT is four days, the $MinI$ is the sum of the shipping prediction in day t plus the next three.

$$SS_t = \sum_{i=0}^{LTD-1} \hat{d}_{t+i} - \sum_{i=0}^{LT-1} \hat{d}_{t+i} \tag{9}$$

SS in day t is calculated by taking the sum of the daily demand prediction, which depends on the Lead Time Delay (LTD), minus the $MinI$ in day t. The LTD quantifies the extra days required to mitigate potential harm from delayed delivery times from suppliers. In this study, LTD doubles the LT due to a business decision.

The function for the monthly EOQ model is:

$$EOQ_m = Q_m = \sqrt{\frac{2\hat{d}_m S}{H_m}} = \sqrt{\frac{2\hat{d}_m S}{I_m P}} \tag{10}$$

where $\hat{d}_m$ is the predicted monthly demand in units for the category, S is the ordering cost for each order, H_m is the monthly per-unit holding cost, P is the purchase price per unit, and I_m is the holding cost per unit per month expressed as a percentage of the price.

After calculating both *OI_forecast* and *OI_actual* and, considering the Current Inventory (CI) figure provided by the retailer, we calculate the following ratios, considering that *OI_actual represents* the ideal scenario:

$$Diff_{forecast} = \frac{OI\_forecast_m - OI\_actual_m}{OI\_actual_m} * 100 \tag{11}$$

$$Diff_{current} = \frac{CI_m - OI\_actual_m}{OI\_actual_m} * 100 \tag{12}$$

4. Results and discussion

4.1 Data collection and preprocessing

The European retail company provided historical order data that was recorded daily at one of its largest warehouses. The data shows the units of products shipped from the warehouse to stores depending on daily customer demand. In addition to that data, the retailer also provided additional variables, including the current inventory (CI) in units, the corresponding lead time (LT) for each category and a dummy variable indicating whether the stores were open. We also added dummy variables for both national and local holidays.

The case study retail company separates the products into seven main sections that are split into categories. The empirical dataset includes the top categories of each section with the highest sales for the years analysed in the paper (Table 2). This dataset was also utilised in Elorza et al. (2024) but considering only the Milk category.

[INSERT TABLE 2 NEAR HERE]

Although promotion campaigns are usually included when forecasting demand (Fildes et al., 2022; Sarlo et al., 2023; Wolters & Huchzermeier, 2021), because we are analysing the category level rather than the product level the promotion factor has been excluded in this study. The retailer informed us that implementing promotions on specific products (such as 3 for 2 or a 20% discount on the second unit) leads to increased sales at the product level but not at the category level. If one product within a category sells more, another product in the same category sells less. In other words, while demand is influenced at the product level, consumer behaviour remains unaffected at the category level.

The data preprocessing was done specifically on the daily customer demand variable. That data is stored in the database in a raw form, so preprocessing and filtering were required. Work was done on categorising items. And, among other actions, items that had been blocked or that had not been shipped in the last year were filtered out.

Outlier values were also removed. The outliers were identified using the boxplot method. This method, originally suggested by Tukey (1977), is widely and extensively used as a tool to detect outliers in univariate datasets (Adil & Zaman, 2020; Zhao & Yang, 2019). An observation is considered to be a potentially abnormal data point if its value falls outside the interval defined by $(q_1 - 1.5 * IQR, q_3 + 1.5 * IQR)$, where q_1 and q_3 are the first and third quartiles, respectively, and IQR is the interquartile range, calculated as $IQR = q_3 - q_1$. Once the potential outliers were identified, we checked whether there was an event or holiday that could explain the outlier values for a specific date. If there was a valid reason (Easter, summer campaign, Christmas campaign or food bank campaign), the values were kept unchanged. Otherwise the values were imputed using the mean of the last three months.

The data referenced in our study was recorded between March 2019 and August 2023, however the data from 2020 and 2021 was tainted by the pandemic (Eurostat, 2021). For that reason it makes sense to omit it because it was an extraordinary situation in terms of household consumption (Sleiman et al., 2022). The COVID-19 pandemic caused a significant change in consumer behaviour (Papanagnou et al., 2022). People reduced their spending on hospitality while purchase volumes in retail stores increased significantly. There was also a decrease in purchase frequency but a significant increase in the volume per transaction, which was a fundamental shift in the customer demand pattern (MAPA, 2020, 2021). Evidence of this shift in behaviour can be seen in Table 3, which lists the peak shipments over the past five years. Upon closer examination, it can be seen that the highest peaks occurred in 2020 and 2021. Specifically, there were respective surges of 68.25% and 21% compared to the 2019 figures.

[INSERT TABLE 3 NEAR HERE]

Omitting those two years, breaks the time series. Consequently, four different alternatives have been proposed:

- Option 1: Exclude 2019 data from the analysis and use data starting from 2022 so there is a time series of 608 days: 365 from 2022 and 243 from 2023.
- Option 2: Considering the homogeneous patterns in the data from 2019, 2022 and 2023 (Fig. 2), and assuming that the pattern of the missing data would have followed a similar trend if the pandemic had not occurred (Sleiman et al., 2022), historical imputation could have been done using the 2019 data as a proxy for

the missing data from 2021. A time series can thus be established with 910 daily observations: 302 from 2021, 365 from 2022, and 243 from 2023.

- Option 3: Another approach would be interpolation, where an equation is used to estimate missing values between two points when similar patterns are detected (Eurostat, 2021). In this scenario, interpolation involves calculating the mean of the values for 2019 and 2022. This approach also produces a time series with 910 daily observations: 302 from 2021, 365 from 2022, and 243 from 2023.
- Option 4: The last proposal is interpolation using the historical evolution between 2019 and 2022. First, the variation in outputs for 2019 and 2022 is calculated, day-by-day, and in total. Second, since we lack data for January and February 2019 to obtain the data for these two months the total variation of the values of January and February 2022 is subtracted. Third, to calculate the units for 2020, the values from 2019 are increased by one-third of the day-by-day variation. And, finally, to calculate the output units for 2021, the values from 2022 are reduced by one-third of the day-by-day. In essence, this process involves using available information from 2019 and 2022 to estimate values for the intermediate years (2020 and 2021) while considering how the variable evolved over time. With this approach, the time series includes 1703 daily observations: 365 from 2019, 365 from 2020, 365 from 2021, 365 from 2022, and 243 from 2023.

[INSERT FIGURE 2 NEAR HERE]

4.2 Demand prediction results

After fitting the training data to the all the models, we assessed the accuracy of the performance of the models with different datasets using two key metrics, where n stands for the number of predictions, and y_i and $\hat{y}_i$ are the ground truth of the time series value and the prediction of the models, respectively. These metrics, denoted MAPE (Mean Absolute Percentage Error) and WAPE (Weighted Absolute Percentage Error), help us evaluate the quality of our predictions. MAPE calculates the percentage difference between the predictions of our model and the actual values in the time series (D. Li et al., 2023). In contrast, WAPE calculates the absolute error averaged over the real requested quantity, i.e. it weights the error (Panarese et al., 2022). Comparing the two metrics, with WAPE the impact of the prediction can be estimated in a more balanced way. MAPE provides a useful measure of prediction error, however it may not be suitable when units shipped are intermittent or near zero. Therefore, WAPE, a metric which is gaining relevance but which has not yet achieved widespread usage in scientific literature, emerges as a more robust choice for assessing prediction accuracy in such scenarios (Panarese et al., 2022). There are no values close to or equal to 0 in the data analysed in this study. So, we assessed prediction accuracy using both metrics.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| * 100 \tag{13}$$

$$WAPE = \sum_{t=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| * 100 \tag{14}$$

Table 4 shows the accuracy results between the actual values and the forecast models in each category. Upon comparison of Option 1, Option 2, Option 3 and Option 4, it becomes apparent that choosing a historical imputation approach is well-founded when confronted with a time series break that exhibits a discernible pattern over different years. This assertion is substantiated by the consistently lower values observed in Option 2 across the seven categories. In contrast, it can be seen that making predictions by considering only the last two years (2022 and 2023) leads to a decrease in the model's accuracy. Moreover, the proposed hybrid Prophet-XGB model is the most accurate for the seven categories analysed in the case study, obtaining the best result in the milk category, with a MAPE of 13.37% and a WAPE of 10.58%. Notably, there are high accuracy predictions for four categories, specifically Milk, Chips and snacks, Cookies, and Toilet paper, with MAPEs lower than 14.39% and WAPEs lower than 11.61%.

[INSERT TABLE 4 NEAR HERE]

The time series data and hybrid model that yielded the best accuracy (Option 2, using the hybridisation model of Prophet and XGB) was identified. Fig. 3 illustrates the actual platform output units for each category alongside the model’s predicted values, spanning 120 days. It can be observed from Fig. 3 that the hybrid prediction model effectively captures the weekly pattern of units shipped from the platforms. The week starts (on Monday) with a significant increase. That is followed by a slight decrease during the mid-week, namely Tuesday and Wednesday. Subsequently, another increase can be seen on Thursday that reaches its peak on Friday. There is a downward trend on Saturday that continues into Sunday, where it reaches the lowest point of the week.

The predictions perform correctly in all categories. We have also observed that the accuracy of predictions for downward peaks is better than for upward peaks. The least accurate predictions are in food bank event outliers, as illustrated by the Legume category in May.

[INSERT FIGURE 3 NEAR HERE]

4.2 Optimal inventory and waste product reduction

In this section, we evaluate the benefits of integrating these predictive models into operational strategies by measuring the reduction in overstock levels compared to traditional business practices.

To achieve this goal, our initial step involves computing the ideal scenario, known as the Optimal Inventory based on actual demand (*OI_actual*). This represents the inventory level we would have ideally calculated if we had possessed perfectly accurate demand forecast. Subsequently, we determine the optimal inventory based on our forecast (*OI_forecast*). We also have data of the stock level that actually was in the retail warehouse *current inventory* (CI). Evaluating the efficacy of our demand forecast entails quantifying the deviation between *OI_forecast* and *OI_actual* ($Diff_{forecast}$) and compared it to the variation between CI and *OI_actual* ($Diff_{actual}$). Both ratios measure the percentage difference in product units.

Table 5 provides an overview of $Diff_{forecast}$ and $Diff_{actual}$. For analysis purposes we have grouped product categories based on whether are food or non-food products. Additionally, within the food category, further distinction is made between perishable and durable items. Negative values denote being bellow the ideal level, while positive values indicate overstock situations.

In the case of perishable products, our model yields similar results to those used by the retail. This is because the retail already had a good adjustment (in no case are there deviations greater than 15%), so the margin for improvement by this study was lower. We can deduce that, given the perishable nature of food, the retail conducts deeper and daily demand analyses to avoid overstock.

Regarding durable products, although they have a long shelf life, they do have a best-before date, and after this date, their sale is usually discontinued. Therefore, it is also important for the retail to manage overstock carefully. In the case of legumes, our model significantly improves the situation. For the total of four months, *OI_forecast* is 4.86% below the ideal scenario compared to an overstock rate of 52.28%. With beer, albeit with lower figures, our model also provides a much better adjustment to the actual demand.

In the case of non-food products, the retail experiences considerable overstock (up to 75% in some months) compared to a worst-case scenario of 5.65% below ideal level, following our model. Our model allows for a significantly closer adjustment to reality. Although these are products with a longer shelf life, excessive storage can lead to losses and deterioration, resulting in product wastage. Additionally, for intimate hygiene products, accumulating too many items poses the risk of consumer preferences changing over time, leading to uncertain future demand.

In general, it is evident that using our model would drastically reduce overstock, and ultimately the wastage of products, especially in durable food and non-food items.

[INSERT TABLE 5 NEAR HERE]

5. Conclusion

We present a comprehensive framework designed to assist retailers in minimising their overstock. Our approach begins with a focus on achieving the highest levels of demand prediction accuracy. Secondly, we propose an innovative method for evaluating overstock variance, leveraging adjustments in demand prediction methodologies.

Regarding demand prediction, two main conclusions are reached. On the one hand, the hybrid Prophet-XGBoost model consistently outperformed other hybridisations in terms of accuracy, achieving the lowest MAPE and WAPE across all seven categories. On the other hand, our results emphasised the significance of historical imputation approaches when dealing with COVID-19 data. More specifically, using the 2019 data as a proxy for the data from 2021 consistently yielded lower errors across categories.

In addition to its contributions to machine learning models, this paper enriches the literature by calculating overstock reduction. While accuracy metrics are important for evaluating the performance of demand prediction models, quantifying the actual reduction in product waste can provide a more direct measure of their effectiveness in practical terms. This metric can provide valuable insights into the model's impact on improving operational efficiency.

Our current study opens up several directions for future research. Concerning demand prediction, besides all the variables we have used to analyse customer demand behaviour, it could be interesting to include other variables, such as weather. On the other hand, regarding optimal inventory calculation other fulfilment options could involve collaborating more closely with the demand planner to adjust the lead time delays for each category instead of applying a uniform calculation.

Acknowledgments

The authors would like to thank the retailer (who wishes to remain anonymous) who provided the data used in the study.

Funding

This work was supported by the Bikaintek 2019 for the completion of industrial doctorates and for the incorporation of research personnel [grant number 20AFW2201900003].

Data availability

The data that has been used is confidential.

References

Adil, I. H., & Zaman, A. (2020). Outliers detection in skewed distributions: Split sample skewness based boxplot. *Economic Computation and Economic Cybernetics Studies and Research*, *54*(3), 279–296. https://doi.org/10.24818/18423264/54.3.20.17

Aswad, F. M., Kareem, A. N., Khudhur, A. M., Khalaf, B. A., & Mostafa, S. A. (2021). Tree-based machine learning algorithms in the Internet of Things environment for multivariate flood status prediction. *Journal of Intelligent Systems*, *31*(1), 1–14. https://doi.org/10.1515/JISYS-2021-0179/MACHINEREADABLECITATION/RIS

Auppakorn, C., & Phumchusri, N. (2022). Daily Sales Forecasting for Variable-Priced Items in Retail Business. *Proceedings of the International Conference on Management Science and Industrial Engineering (MSIE)*, 80–86. https://doi.org/10.1145/3535782.3535794

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32.

Caro, F., Gürhan Kök, A., & Martínez-De-Albéniz, V. (2020). The Future of Retail Operations. *Manufacturing & Service Operations Management*, *22*(1), 47–58. https://doi.org/10.1287/MSOM.2019.0824

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Demir, S., & Sahin, E. K. (2022). Comparison of tree-based machine learning algorithms for predicting liquefaction potential using canonical correlation forest, rotation forest, and random forest based on CPT data. *Soil Dynamics and Earthquake Engineering*, *154*, 107130. https://doi.org/10.1016/J.SOILDYN.2021.107130

Deng, C., & Liu, Y. (2021). A Deep Learning-Based Inventory Management and Demand Prediction Optimization Method for Anomaly Detection. *Wireless Communications and Mobile Computing*, *2021*, 1–14. https://doi.org/10.1155/2021/9969357

Ehrenthal, J. C. F., Honhon, D., & Van Woensel, T. (2014). Demand seasonality in retail inventory management. *European Journal of Operational Research*, *238*(2), 527-527–539. https://doi.org/10.1016/j.ejor.2014.03.030

Elorza, M., Castellano, E., & Segura, S. (2024). Prediction of customer demand for perishable products in retail inventory management, using the hybrid prophet-XGBoost model during the post-COVID-19 period. *Applied Economics Letters*, 1–7. https://doi.org/10.1080/13504851.2024.2333995

European Commission. (2023). *Food waste reduction targets*. https://food.ec.europa.eu/safety/food-waste/eu-actions-against-food-waste/food-waste-reduction-targets_en

European Commission. (2024). *Retail - European Commission*. https://single-market-economy.ec.europa.eu/single-market/services/retail_en

Eurostat. (2021). *Guidance on the impact of the Covid-19 crisis on annual business statistics*. https://ec.europa.eu/eurostat/web/covid-19/support-statisticians#Economy

Eurostat. (2023). *Food waste and food waste prevention - estimates*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Food_waste_and_food_waste_prevention_-_estimates#Amounts_of_food_waste_at_EU_level

Fildes, R., Kolassa, S., & Ma, S. (2022). Post-script - Retail forecasting: Research and practice. *International Journal of Forecasting*, *38*(4), 1319–1324. https://doi.org/10.1016/J.IJFORECAST.2021.09.012

FoodDrinkEurope. (2023). *EU Food and Drink Industry 2023 Edition*. https://www.fooddrinkeurope.eu/wp-content/uploads/2023/12/FoodDrinkEurope-Data-Trends-Report-2023-digital.pdf

Funk, K. D., Paul, H. L., & Philips, A. Q. (2022). Point break: using machine learning to uncover a critical mass in women's representation. *Political Science Research and Methods*, *10*(2), 372–390. https://doi.org/10.1017/PSRM.2021.51

Ghiasi, M. M., & Zendehboudi, S. (2021). Application of decision tree-based ensemble learning in the classification of breast cancer. *Computers in Biology and Medicine*, *128*, 104089. https://doi.org/10.1016/j.compbiomed.2020.104089

Harris, F. W. (1915). Operations and cost. *Factory Management Series*, *2*, 48–52.

Kim, H.-J., Kim, J.-H., & Im, J. (2023). Forecasting Offline Retail Sales in the COVID-19 Pandemic Period: A Case Study of a Complex Shopping Mall in South Korea. *BUILDINGS*, *13*(3), 627-627–639. https://doi.org/10.3390/buildings13030627

Kramar, V., & Alchakov, V. (2023). Time-Series Forecasting of Seasonal Data Using Machine Learning Methods. *Algorithms*, *16*(5). https://doi.org/10.3390/A16050248

Lai, J. P., Lin, Y. L., Lin, H. C., Shih, C. Y., Wang, Y. P., & Pai, P. F. (2023). Tree-Based Machine Learning Models with Optuna in Predicting Impedance Values for Circuit Analysis. *Micromachines*, *14*(2). https://doi.org/10.3390/MI14020265

Li, D., Li, X., Gu, F., Pan, Z., Chen, D., & Madden, A. (2023). A Universality–Distinction Mechanism-Based Multi-Step Sales Forecasting for Sales Prediction and Inventory Optimization. *Systems*, *11*(6), 311. https://doi.org/10.3390/SYSTEMS11060311

Li, Y., Stasinakis, C., & Yeo, W. M. (2022). A Hybrid XGBoost-MLP Model for Credit Risk Assessment on Digital Supply Chain Finance. *Forecasting*, *4*(11), 184–207. https://doi.org/10.3390/FORECAST4010011

Madeira, C. (2020). The impact of the COVID public policies on the Chilean households. *Applied Economics Letters*, *28*(18), 1562–1565. https://doi.org/10.1080/13504851.2020.1832194

MAPA. (2020). *Informe del consumo de alimentación en España 2020*. Ministerio de Agricultura, Pesca y Alimentación. https://www.mapa.gob.es/es/alimentacion/temas/consumo-tendencias/informe-anual-consumo-2020-v2-nov2021-baja-res_tcm30-562704.pdf

MAPA. (2021). *Informe del consumo de alimentación en España 2021*. Ministerio de Agricultura, Pesca y Alimentación. https://www.mapa.gob.es/es/alimentacion/temas/consumo-tendencias/informe-consumo-alimentario-2021-baja-res_tcm30-624017.pdf

Miguéis, V. L., Pereira, A., Pereira, J., & Figueira, G. (2022). Reducing fresh fish waste while ensuring availability: Demand forecast using censored data and machine learning. *Journal of Cleaner Production*, *359*. https://doi.org/10.1016/j.jclepro.2022.131852

Pan, Z., Fang, S., & Wang, H. (2021). LightGBM Technique and Differential Evolution Algorithm-Based Multi-Objective Optimization Design of DS-APMM. *IEEE Transactions on Energy Conversion*, *36*(1), 441–455. https://doi.org/10.1109/TEC.2020.3009480

Panarese, A., Settanni, G., Vitti, V., & Galiano, A. (2022). Developing and Preliminary Testing of a Machine Learning-Based Platform for Sales Forecasting Using a

Gradient Boosting Approach. *Applied Sciences (2076-3417)*, *12*(21), 11054-11054–11068. https://doi.org/10.3390/app122111054

Papanagnou, C., Seiler, A., Spanaki, K., Papadopoulos, T., & Bourlakis, M. (2022). Data-driven digital transformation for emergency situations: The case of the UK retail sector. *International Journal of Production Economics*, *250*, 108628. https://doi.org/10.1016/J.IJPE.2022.108628

Pavlyshenko, B. M. (2019). Machine-Learning Models for Sales Time Series Forecasting. *Data*, *4*(1), 15. https://doi.org/10.3390/DATA4010015

Perez, H. D., Hubbs, C. D., Li, C., & Grossmann, I. E. (2021). Algorithmic Approaches to Inventory Management Optimization. *Processes*, *9*(1), 102. https://doi.org/10.3390/PR9010102

Rodrigues, M., Miguéis, V., Freitas, S., & Machado, T. (2024). Machine learning models for short-term demand forecasting in food catering services: A solution to reduce food waste. *Journal of Cleaner Production*, *435*. https://doi.org/10.1016/j.jclepro.2023.140265

Roggeveen, A. L., & Sethuraman, R. (2020). How the COVID-19 Pandemic May Change the World of Retailing. *Journal of Retailing*, *96*(2), 169–171. https://doi.org/10.1016/J.JRETAI.2020.04.002

Sarlo, R., Fernandes, C., & Borenstein, D. (2023). Lumpy and intermittent retail demand forecasts with score-driven models. *European Journal of Operational Research*, *307*(3), 1146–1160. https://doi.org/https://doi.org/10.1016/j.ejor.2022.10.006

Shakeel, A., Chong, D., & Wang, J. (2023). District heating load forecasting with a hybrid model based on LightGBM and FB-prophet. *Journal of Cleaner Production*, *409*, 137130. https://doi.org/10.1016/j.jclepro.2023.137130

Sibindi, R., Mwangi, R. W., & Waititu, A. G. (2022). A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices. *Engineering Reports*. https://doi.org/10.1002/ENG2.12599

Sleiman, R., Mazyad, A., Hamad, M., Tran, K. P., & Thomassey, S. (2022). Forecasting Sales Profiles of Products in an Exceptional Context: COVID-19 Pandemic. *International Journal of Computational Intelligence Systems*, *15*(1). https://doi.org/10.1007/s44196-022-00161-x

Srinidhi, N., Sagar, C., Deepak Chethan, S., Shreyas, J., & Dilip Kumar, S. (2020). An improved PRoPHET - Random forest based optimized multi-copy routing for opportunistic IoT networks. *Internet of Things*, *11*, 100203. https://doi.org/10.1016/J.IOT.2020.100203

Srinivas, P., & Katarya, R. (2022). hyOPTXg: OPTUNA hyper-parameter optimization framework for predicting cardiovascular disease using XGBoost. *Biomedical Signal Processing and Control*, *73*, 1746–8094. https://doi.org/10.1016/j.bspc.2021.103456

Tait, P., Saunders, C., Dalziel, P., Rutherford, P., Driver, T., & Guenther, M. (2024). How much less? Estimating price discounts for suboptimal food with environmental and social credence attributes. *Applied Economics*, *56*(13), 1581–1594. https://doi.org/10.1080/00036846.2023.2176460

Taylor, S. J., & Letham, B. (2018). Forecasting at Scale. *The American Statistician*, *72*(1), 37–45. https://doi.org/10.1080/00031305.2017.1380080

Tukey, J. W. (1977). Exploratory Data Analysis. In *Addison-Wesley*.

Wang, J., Du, X., & Qi, X. (2022). Strain prediction for historical timber buildings with a hybrid Prophet-XGBoost model. *Mechanical Systems and Signal Processing*, *179*, 109316. https://doi.org/10.1016/j.ymssp.2022.109316

Wilson, R. H. (1934). A Scientific Routine For Stock Control. *Harvard Business Review*, *13*(1), 116–129.

Wolters, J., & Huchzermeier, A. (2021). Joint In-Season and Out-of-Season Promotion Demand Forecasting in a Retail Environment. *Journal of Retailing*, *97*(4), 726–745. https://doi.org/10.1016/J.JRETAI.2021.01.003

Yu, T., & Huo, Y. (2022). Complexity Analysis of Consumer Finance following Computer LightGBM Algorithm under Industrial Economy. *Mobile Information Systems*, *2022*. https://doi.org/10.1155/2022/2865959

Zhao, C., & Yang, J. (2019). A Robust Skewed Boxplot for Detecting Outliers in Rainfall Observations in Real-Time Flood Forecasting. *Advances in Meteorology*, *2019*. https://doi.org/10.1155/2019/1795673

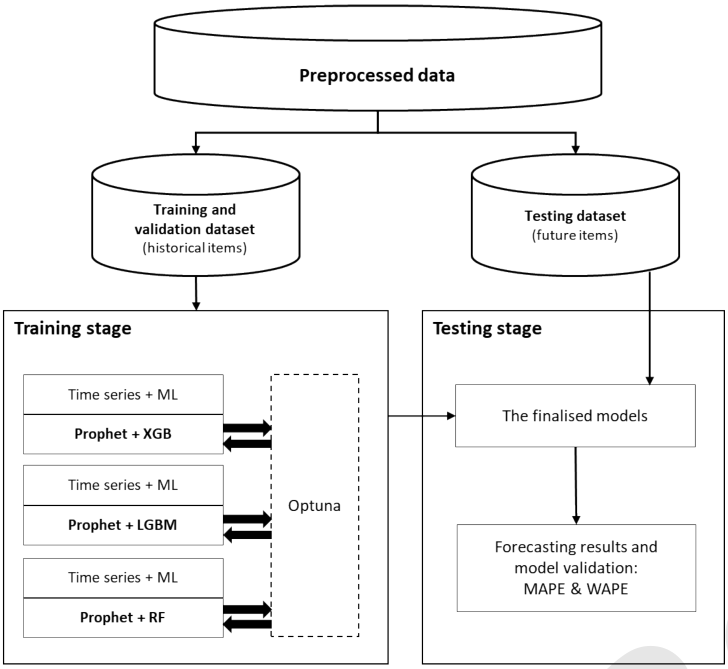

Fig. 1. The procedure flow used for demand prediction.

Model	Hyperparameters	Types	Search Ranges
XGB	n_estimators	Integer	[50, 10000]
	max_depth	Integer	[3, 10]
	subsample	Float	[0, 1]
	reg_alpha	Float	[0, 180]
	reg_lambda	Float	[0, 1]
	learning_rate	Float	[0.01, 0.3]
	colsample_bytree	Float	[0.3, 1]
	min_child_weight	Float	[0.3, 10]
LGBM	n_estimators	Integer	[50, 10000]
	max_depth	Integer	[3, 10]
	subsample	Float	[0, 1]
	reg_alpha	Float	[0, 180]
	reg_lambda	Float	[0, 1]
	learning_rate	Float	[0.01, 0.3]
	colsample_bytree	Float	[0.3, 1]
	min_child_weight	Float	[0.3, 10]
	num_leaves	Integers	[10, 120]
RF	n_estimators	Integer	[50, 10000]
	max_depth	Integer	[3, 10]
	min_samples_split	Integer	[2, 10]

Table 1. List of hyperparameters optimised using the Optuna library in the XGB, LGBM and RF models.

Section	Category	Description	Mean units shipped per week
Dairy	Milk	A white liquid produced by mammals	168,333.06
Drinks	Beer	An alcoholic beverage made from fermented grains	119,856.17
Sweets	Cookies	Sweet baked treats made from dough	31,412.71
Salty	Chips and snacks	Savoury and crunchy snack foods	20,739.94
Pharmacy	Toilet paper	A paper product used for personal hygiene	13,596.48
Staple foods	Legume	A type of plant that produces seeds in a pod	12,225.46
Perfumery	Protection and intimate hygiene	Products related to personal care and protection	5,975.94

Table 2. Description of the data sample.

Year	Max units shipped
2019	351,842
2020	592,086
2021	445,939
2022	386,886
2023	341,345

Table 3. Shipped units maximum peak per year.

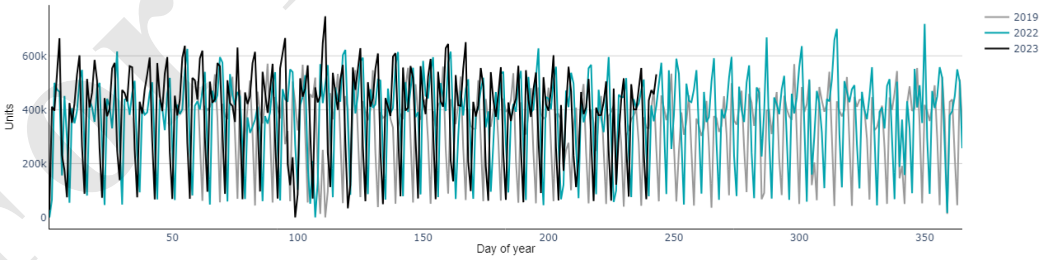

Fig. 2. Evolution of units shipped by year: 2019, 2022 and 2023.

Milk

Models	**Option 1**		**Option 2**		**Option 3**		**Option 4**	
	MAPE	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***
Prophet - XGB	19.28%	12.70%	**13.37%**	**10.58%**	17.29%	11.62%	17.49%	11.84%
Prophet - LGBM	23.48%	14.58%	20.24%	12.79%	19.23%	12.79%	18.80%	12.80%
Prophet - RF	26.64%	21.11%	21.35%	13.96%	21.00%	13.68%	20.81%	13.37%

Beer

Models	**Option 1**		**Option 2**		**Option 3**		**Option 4**	
	MAPE	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***
Prophet -XGB	27.38%	17.76%	**17.27%**	**14.35%**	22.58%	17.32%	23.63%	17.25%
Prophet - LGBM	44.90%	23.91%	35.10%	20.05%	38.81%	22.49%	39.61%	21.78%
Prophet - RF	41.92%	27.36%	38.82%	23.03%	36.00%	25.21%	32.20%	23.34%

Cookies

Models	**Option 1**		**Option 2**		**Option 3**		**Option 4**	
	MAPE	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***
Prophet - XGB	24.60%	13.28%	**14.39%**	**11.61%**	20.78%	12.98%	19.49%	12.06%
Prophet - LGBM	26.66%	14.17%	27.09%	13.55%	27.86%	14.31%	20.49%	12.40%
Prophet - RF	29.69%	21.14%	25.45%	15.05%	29.47%	17.50%	20.93%	12.69%

Chips and snacks

Models	**Option 1**		**Option 2**		**Option 3**		**Option 4**	
	MAPE	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***
Prophet - XGB	28.42%	15.23%	**13.71%**	**11.53%**	23.90%	14.04%	25.48%	15.00%
Prophet - LGBM	30.29%	17.13%	35.35%	19.74%	33.91%	18.02%	40.09%	20.25%
Prophet - RF	29.24%	19.45%	59.45%	40.95%	44.34%	28.17%	64.54%	35.91%

Toilet paper

Models	**Option 1**		**Option 2**		**Option 3**		**Option 4**	
	MAPE	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***
Prophet - XGB	21.94%	13.95%	**13.95%**	**11.46%**	17.76%	13.04%	19.63%	13.36%
Prophet - LGBM	25.11%	15.70%	24.93%	15.49%	22.07%	15.31%	22.63%	15.65%
Prophet - RF	26.25%	16.67%	31.17%	17.83%	24.60%	17.21%	23.89%	15.78%

				Legume				
Models	**Option 1**		**Option 2**		**Option 3**		**Option 4**	
	MAPE	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***
Prophet - XGB	29.67%	21.87%	**18.80%**	**17.19%**	28.43%	19.68%	27.12%	19.29%
Prophet - LGBM	34.26%	26.15%	27.37%	20.79%	28.69%	21.72%	27.16%	21.62%
Prophet - RF	31.15%	25.00%	27.10%	19.86%	27.21%	20.39%	26.28%	20.85%
				Protection and intimate hygiene				
Models	**Option 1**		**Option 2**		**Option 3**		**Option 4**	
	MAPE	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***	***MAPE***	***WAPE***
Prophet - XGB	31.05%	16.46%	**19.64%**	**14.79%**	28.99%	16.46%	27.44%	16.36%
Prophet - LGBM	37.44%	19.82%	33.95%	18.95%	32.98%	18.79%	30.30%	18.71%
Prophet - RF	35.65%	21.09%	36.38%	20.14%	35.90%	19.95%	32.77%	19.14%

Table 4. Performance evaluation summary with the best results highlighted in bold.

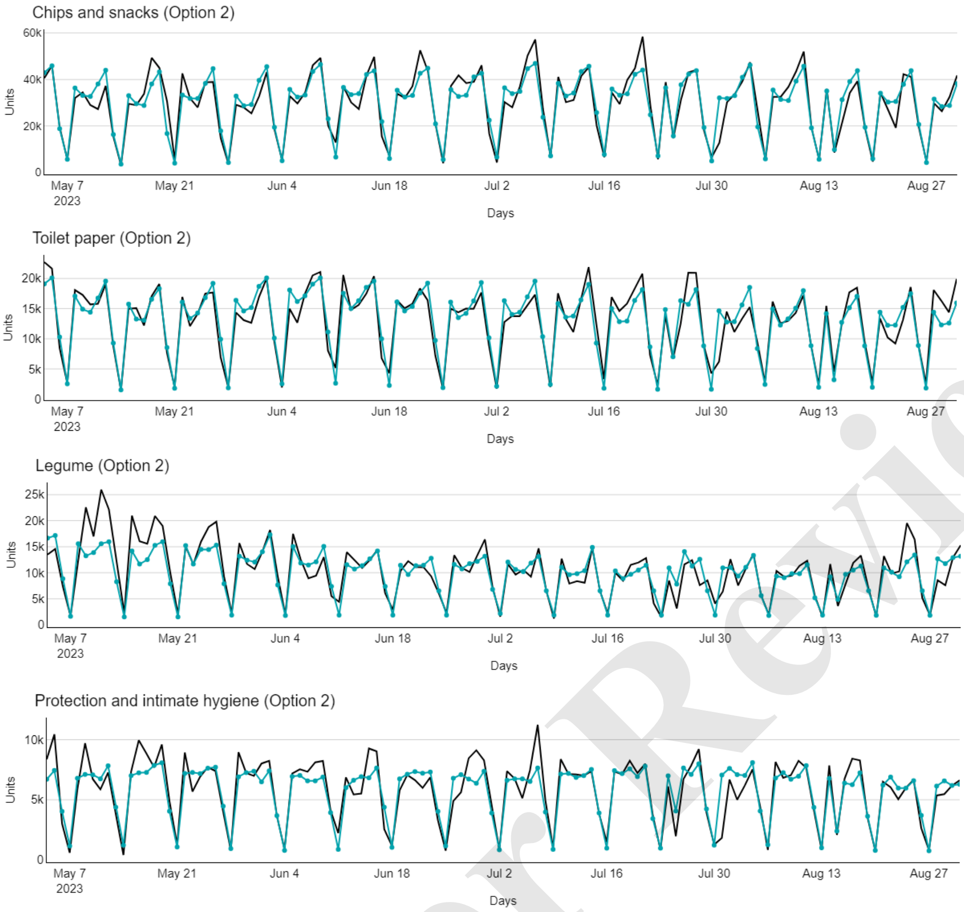

Fig. 3. Based on Prophet-XGB hybrid shipped units forecasts modelled from May and August (120 days).

	Food – Perishable products					
Month	**Milk**		**Chips & snacks**		**Cookies**	
	Diff_forecast	*Diff_current*	*Diff_forecast*	*Diff_current*	*Diff_forecast*	*Diff_current*
May	-8.03	-12.51	1.60	-4.69	3.91	-5.57
June	-4.51	4.53	-14.87	-13.49	-0.95	11.49
July	-5.68	6.68	-15.46	-16.44	-4.29	5.29
August	-2.79	1.99	-8.28	-7.90	-6.81	4.78
TOTAL	**-5.32**	**-0.09**	**-9.75**	**-10.95**	**-2.06**	**-1.46**

	Food – Durable products			
Month	**Legume**		**Beer**	
	Diff_forecast	*Diff_current*	*Diff_forecast*	*Diff_current*
May	-15.24	13.74	3.74	23.47

June	-0.58	85.35	5.70	32.86
July	6.47	92.13	4.64	38.46
August	-5.95	33.56	-7.15	35.13
TOTAL	**-4.86**	**52.28**	**1.70**	**32.68**
Non Food products				
Month	**Toilet**		**Protection & intimate hygiene**	
	Diff_forecast	*Diff_current*	*Diff_forecast*	*Diff_current*
May	3.64	73.97	-5.65	29.18
June	3.32	66.61	-2.87	43.74
July	-1.84	59.47	2.60	57.14
August	-4.93	57.00	-1.48	37.07
TOTAL	**0.01**	**64.11**	**-1.85**	**41.82**

Note: Negative values denote being below the ideal level, while positive values indicate overstock situations.

Table 5. Deviation to the ideal scenario from our forecast $(Diff_{forecast})$ and current inventory $(Diff_{current})$. Percentage units.