Demand Forecasting: A Case Study in the Food Industry

Juliana C. Silva[1], Manuel C. Figueiredo[2(✉)], and Ana C. Braga[2]

[1] DPS, School of Engineering, University of Minho, Braga, Portugal
ju_silva21@hotmail.com
[2] ALGORITMI Centre, University of Minho, 4710-057 Braga, Portugal
{mcf,acb}@dps.uminho.pt

Abstract. The use of forecasting methods is nowadays regarded as a business ally since it supports both the operational and the strategic decision-making processes.

This paper is based on a research project aiming the development of demand forecasting models for a company (designated here by PR) that operates in the food business, more specifically in the delicatessen segment.

In particular, we focused on demand forecasting models that can serve as a tool to support production planning and inventory management at the company.

The analysis of the company's operations led to the development of a new demand forecasting tool based on a combination of forecasts, which is now being used and tested by the company.

Keywords: Forecasting demand · Exponential smoothing · ARIMA · Combining forecasts

1 Introduction

1.1 The Company

PR Company already has more than half a century of history, presenting itself today as a reference company in the production and distribution of delicatessen.

From its foundation to the present, PR has specialized and transformed itself into a modern organization, with a professional management and equipped with the latest technologies, thus managing to anticipate the challenges and specificities of competitiveness and globalization without never losing its roots in Portugal.

It should also be mentioned that PR is a highly certified company by several standards that attest Quality Management Systems (ISO 9001: 2008), Environmental Management System (ISO 14001: 2004) and Food Safety Management Systems (ISO 22000: 2005). It is also internationally recognized by specific standards by the food industry, such as IFS - International Featured Standard, which strengthens its competitiveness in the global market.

The company's strategy is essentially a sustained business growth, always guaranteeing a quality of excellence to the consumer, assuming quality and safety as a business priority. In order to respond to market demand, PR presents a diverse portfolio of products, which can be grouped into six different product families as shown in Fig. 1.
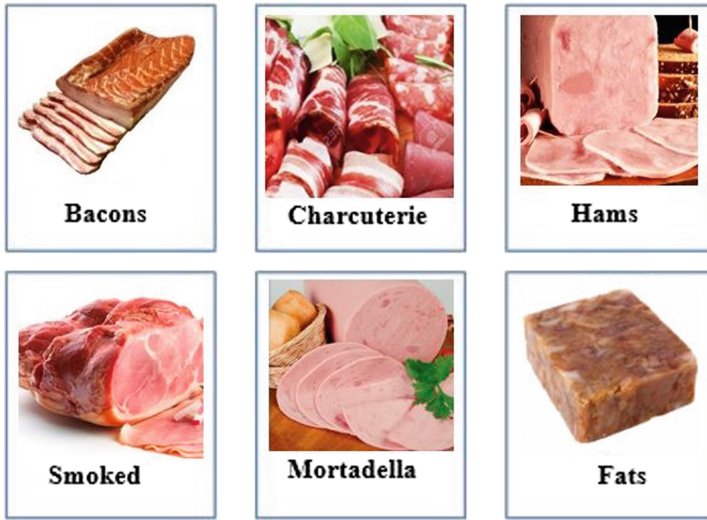

Fig. 1. Product families.

Currently, PR is present in the most diverse sectors of the national distribution: retail, cash & carry, modern distribution and professional channels.

In addition to the domestic market, where it has been gaining strength and holding leading positions, PR has already achieved strong internationalization of its brand in countries such as: Angola, Brazil, France, Germany, Holland, Luxembourg, Mozambique, United Kingdom, Russia, South Africa and Sweden (Fig. 2).

1.2   The Importance of Demand Forecasting in the Food Industry

Due to the increasing level of competitiveness among companies, forecasting plays an important role in supply chain management, and the viability of a company is often dependent on the efficiency and accuracy of forecasts [14]. Demand forecasts are behind all strategic and planning decisions in any retail business as they directly affect the company's profitability and competitive position.

For these reasons, the use of demand forecasting techniques is a fundamental support in planning and managing a company's supply chain [7].

Its importance becomes patent, since its outcome is used for many functions within the organization: they allow the financial department to estimate

Fig. 2. PR company's markets.

costs, levels of profit and capital needs; they enable the sales department to obtain know-how of each product's sales volume; the purchasing department may plan short- and long-term acquisitions; the marketing department can plan their actions and evaluate the impact of different marketing strategies on the sales volume and brand awareness; the logistics department will be able to define the specific logistics needs and finally, the operations department that can manage and plan in advance the purchase of machinery and materials, as well as the hiring of labor.

It is therefore consensual that forecasts are very useful, and even essential for most companies. Accurate demand forecasts have the potential to increase profitability, improving the chain's efficiency and reducing waste.

This paper describes various demand forecasting models for products made by a food company. In the food business, a proper management of inventories involves numerous articles whose particular characteristics, namely perishability, are relevant. Bad decisions in this area can lead to large losses related to excess stock.

2   Demand Forecasting

2.1   Demand Forecasting Methods

Predicting demand is a fundamental activity as it can reveal market trends and contribute to the strategic planning of the company. According to [1], demand forecasting is an essential tool for a faster and safer decision-making process.

There are several techniques available to support analysts in forecasting demand. Although these techniques have substantial differences, there are common characteristics:

– They generally assume that the causes that have influenced demand in the past will continue to act in the future;

– Forecast accuracy decreases as the forecasting horizon increases;
– Aggregated forecasts for product groups are more accurate than individual product forecasts.

Forecasting methods may be divided into quantitative and qualitative methods. Quantitative methods require the construction of mathematical models, using historical data, that describe demand variation over time. These methods include decomposition, moving averages, exponential smoothing, ARIMA, etc.

In general, qualitative methods result from the opinion of process specialists to predict demand. They are frequently questioned as the systematic approach provided by quantitative techniques presents a better performance concerning future estimates. However, in cases of information scarcity, for example in the launching of new products, the experience and know-how of managers may be useful.

2.2 Demand Forecasting Process at PR

The data available consisted of weekly sales from the first week of 2013 through week 17 of 2016 (Fig. 3).
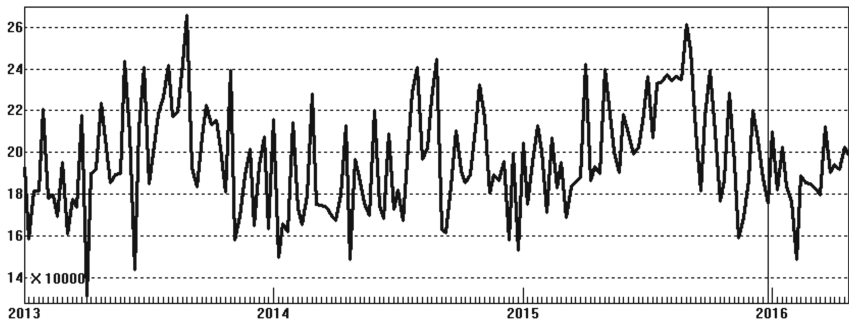

Fig. 3. Sales from week 1 of 2013 to week 17 of 2016.

We start by analysing how the company predicts its sales in order to optimize processes and reduce unnecessary stocks or avoid lost sales.

The demand forecast performed by PR is based on a 4-week moving average. Thus in week t, a demand forecast for week $t+2$ is made for each product based on the actual sales concerning weeks $t-4$, $t-3$, $t-2$ and $t-1$.

Fig. 4 shows the actual sales (black) and the forecasts obtained with the 4-week moving average model (red).

The error autocorrelation function for this model is illustrated in Fig. 5.

We evaluated the normality of the errors using Kolmogorov-Smirnov test, which reveals that the errors could be considered normally distributed ($p = .200$). This result is illustrated in Fig. 6 (a and b).
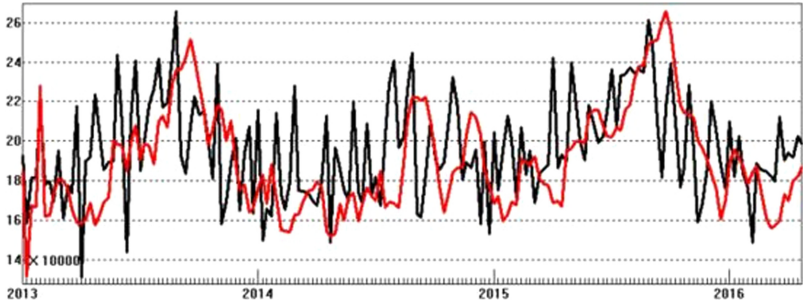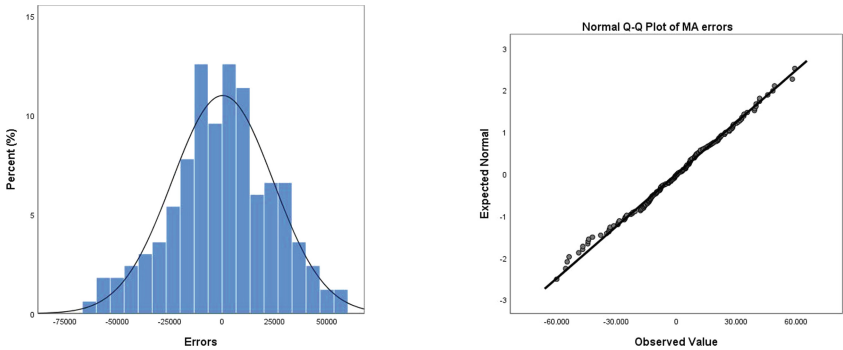
Fig. 4. Actual sales vs Forecasts using the 4-week MA model. (Color figure online)

Fig. 5. 4-week MA error autocorrelation function.

(a) Histogram of MA errors.   (b) QQ Plot of MA errors.

Fig. 6. Normality assumption for MA errors.

Currently, and as a form of support, the company has a file in Excel format, which is updated weekly with the actual sales.

These base forecasts may be adjusted according to qualitative information obtained from its sales force. Salespeople play a key role in obtaining information for forecasts as they work very closely to customers and to the market. It is up to the task manager to generate forecasts, informally collect information from the sales team in order to adjust the base forecasts, obtained from the method above described.

The inputs collected from the commercial team can be of various nature. For example, a sales campaign for a particular product for a particular customer, or the exit from the market of a competing product. In this case, the sensitivity and experience of the task manager comes into play, which can increase or decrease orders to a reasonably weighted value.

Accordingly, the value of expected sales (in kg), calculated on the basis of the 4-week moving average, is adjusted by the person responsible for the forecasts, which includes the inputs that he collects from both past sales and future sales. The aim is mainly to try to decode the demand peaks of certain products and/or their low demand in certain weeks.

In summary, after the calculation of the base forecasts based on the 4-week moving averages (MA), in order to include all the relevant information it becomes necessary to receive inputs from the marketing area (promotion of new products and promotional plans), the commercial area (customer and market data) and operational planning (capacity constraints, logistics, etc.) in order to adjust the values against these inputs.

After collection, processing and analysis phases, the final values obtained are validated once more and communicated to the production planning department.

3   Exponential Smoothing Models

In the moving average method described in the previous section, the forecast is determined by assigning the same weight to each of the last 4 observations available ($t - 1$, $t - 2$, $t - 3$ and $t - 4$) and ignoring all the older observations (prior to week $(t - 4)$).

However, it is reasonable to assume that the most recent observations contain more relevant information about what might happen in the future and therefore should have a greater weight in forecasting than the older observations. Exponential smoothing methods are based on this principle.

Exponential smoothing is one of the most popular forecasting methods. This method applies a weighted average to the observations of a time series, with greater weights being given to the most recent information.

Single exponential smoothing is based on the Eq. 1

$$F_{t+1} = \alpha X_t + (1 - \alpha)F_t \tag{1}$$

where F_{t+1} is the forecast for period $t + 1$, α is the smoothing constant and X_t is the observed value for period t.

Since the smoothing constant varies between 0 and 1, more weight is given to the most recent observations in determining forecasts. Its application in forecasting appears to have been pioneered in 1956 by Robert Brown. In 1957, Holt described double exponential smoothing and, in 1960, Peter R. Winters improved the algorithm by adding seasonality. This algorithm became known as triple exponential smoothing or the Holt-Winters method [10, 13].

Using data from 2013, 2014 and 2015 and the Forecast Pro software, we developed exponential smoothing models to forecast the first 17 weeks of 2016. The model selection criterion was the Mean Absolute Deviation (MAD), defined in Eq. 2.

$$MAD = \frac{\sum_{i=1}^{n} |X_i - F_i|}{n} \tag{2}$$

where X_i and F_i represent, respectively, the actual sales in week i and sales forecast for week i.

The model that minimized within sample MAD is described in Table 1.

Table 1. Exponential smoothing model selected.

Forecast model	Exponential smoothing: No trend, Multiplicative seasonality
Smoothing weights	
Level	$\alpha = 0.05227$
Seasonal	$\gamma = 0.3294$
Mean Absolute Deviation:	$MAD = 14100$

Figure 7 shows the actual sales (black) and the forecasts obtained with the selected model (red).
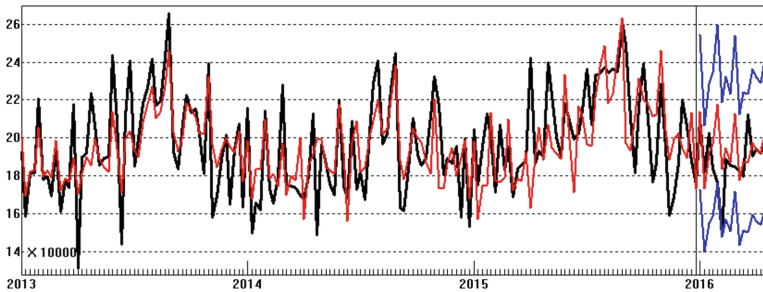

Fig. 7. Actual sales vs Forecasts using Exponential smoothing model. (Color figure online)

The error autocorrelation function for this exponential smoothing model is illustrated in Fig. 8.

Fig. 8. Exponential smoothing error autocorrelation function.

In addition, we evaluated the normality of the errors using Kolmogorov-Smirnov test, which reveals that the errors could be considered normally distributed ($p = .058$). This result is illustrated in Fig. 9 (a and b).
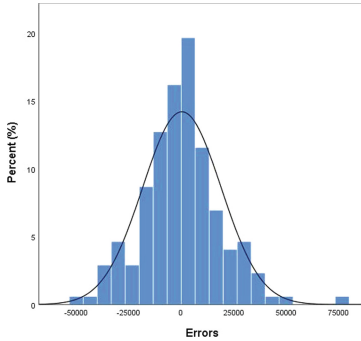
(a) Histogram of ES errors.
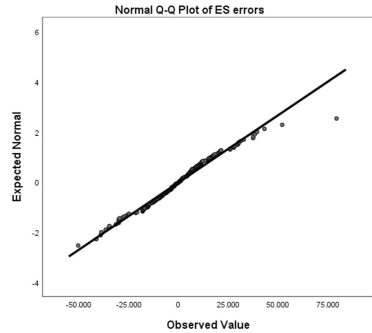
(b) QQ Plot of ES errors.

Fig. 9. Normality assumption for ES errors.

4   ARIMA Models

In 1970 George Box and Gwilym Jenkins popularized ARIMA (Autoregressive Integrated Moving Average) models in their seminal textbook, *Time Series Analysis: Forecasting and Control*. ARIMA models generated a lot of interest in the academic community, due mostly to their theoretical foundations, which proved that, under certain assumptions, the models would produce optimal forecasts.

However, the Box-Jenkins methodology apparently hasn't been widespread adopted among the business community. This was mostly due to the difficult, time-consuming and subjective procedure described by Box and Jenkins to identify the proper form of the model for a given dataset. Furthermore, empirical studies showed that despite the ARIMA model's theoretical superiority over other forecasting methods, in practice the models did not regularly outperform other time series methods.

Generically, a non-seasonal Box-Jenkins model is represented as ARIMA (p, d, q), where p indicates the number of AR terms, d indicates the order of differencing, and q indicates the number of MA terms. A seasonal Box-Jenkins model is symbolized as ARIMA $(p, d, q) * (P, D, Q)$, where the p,d,q indicate the model orders for the short-term components of the model, and P,D,Q indicate the model orders for the seasonal components of the model.

We used Forecast Pro to build the ARIMA models. After transforming data in order to remove trend (using a simple difference) and seasonality (using a seasonal difference) we analysed the Autocorrelation function (ACF) and the Partial Autocorrelation function (PACF), presented in Figs. 10 and 11.
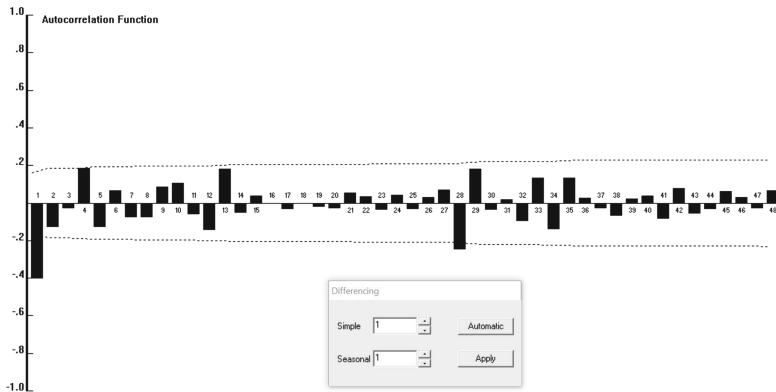

Fig. 10. Autocorrelation Function (ACF) after a simple and a seasonal difference.

This analysis led to an ARIMA$(0, 1, 1) * (0, 1, 1)^{52}$ that minimized within sample $MAD = 9772$ and is summarized in Table 2.

Table 2. ARIMA model selected.

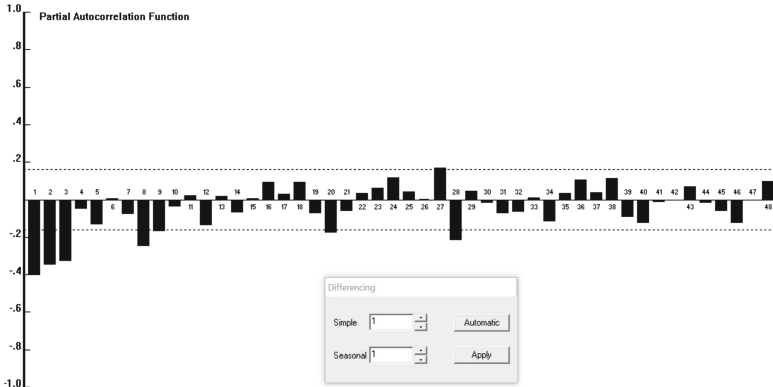| Term	Coefficient	Std. error	t-statistic
$q[1[$	0.8878	0.0447	19.84
$Q[52[$	0.7160	0.0465	15.38

Fig. 11. Partial Autocorrelation Function (PACF) after a simple and a seasonal difference.

Model coefficients are significant and the Error Autocorrelation function (Fig. 12) shows that errors are random. Figure 13 shows the actual sales (black) and the forecasts obtained with the selected model (red).
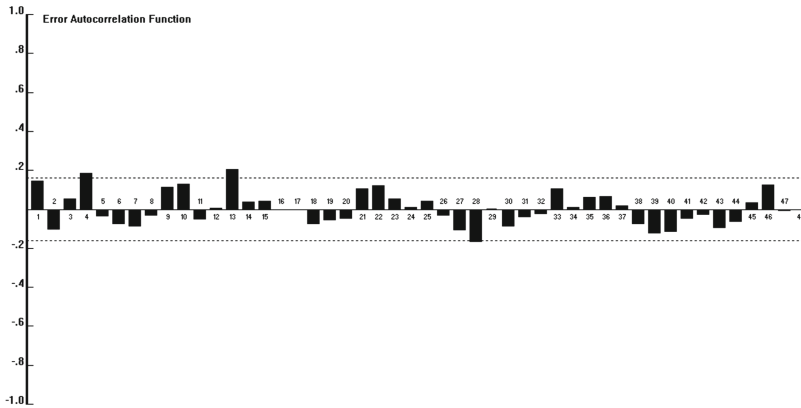

Fig. 12. Error Autocorrelation Function (ACF) for $ARIMA(0, 1, 1) \times (0, 1, 1)^{52}$.

Normality evaluation of errors for ARIMA model, using Kolmogorov-Smirnov test, reveals that the errors could be considered normally distributed ($p = .063$). This result is illustrated in Fig. 14 (a and b).
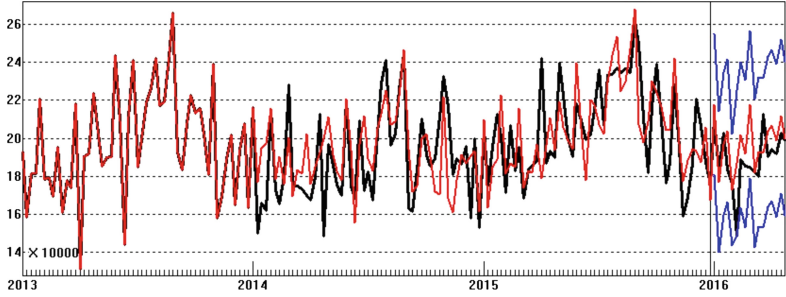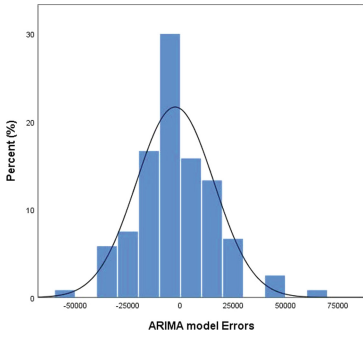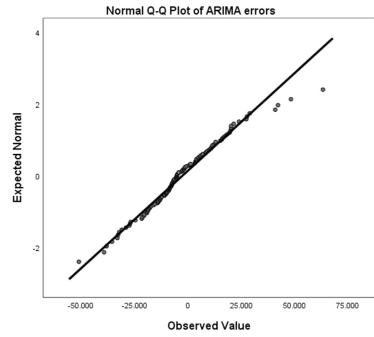
Fig. 13. Actual sales vs Forecasts using $ARIMA(0,1,1) \times (0,1,1)^{52}$. (Color figure online)

(a) Histogram of ARIMA errors.

(b) QQ Plot of ARIMA errors.

Fig. 14. Normality assumption for ARIMA errors.

5   Forecasting Methodology Proposed

5.1   Combining Forecasts

There are many reported situations where a combination of different forecasts can outperform individual forecasts.

The development of research on combining forecasts was largely documented by Clemen [5], who pointed out that "the primary conclusion of this line of research is that forecast accuracy can be substantially improved through the combination of multiple individual forecasts. Furthermore, simple combination methods often work reasonably well relative to more complex combinations."

The value of combining forecasts in an inventory management context has been reported by Chan et al. [4], in a case study where "the results show that combining different forecasts will lead to significant improvements in demand forecasting performance and savings of around 10% in the amount of safety stock that will need to be carried."

5.2   Combining Methods

Although a significant amount of research has been published on combining forecasts, the choice of which method should be used in a particular situation is not always clear. Over the last forty years, many combining methods have been proposed, ranging from the simple average to much more sophisticated approaches to determine optimal combinations.

The simple average is reported to perform well in many published studies. The simple average has the advantage of being uncomplicated and robust, which is a very attractive feature in a context where hundreds or thousands of items have to be periodically forecast.

Another approach, known as outperformance, Bunn [3], gives different weights to individual forecasts. Each weight is made proportional to the probability that the respective forecasting method will have the best individual performance. Each probability may be estimated on the basis of past performance of the alternative forecast methods, and can be revised using a Bayesian analysis. It has the advantage of being robust and having an intuitive meaning. Bates and Granger [2] proposed an approach where weights are calculated in order to minimise the error variance of the combination, assuming that each individual forecast is unbiased. Granger and Ramanathan [9] showed that this method is equivalent to a least squares regression, in which the constant is suppressed and the weights are constrained to sum to one. They proposed a regression method where individual forecasts are used as regressors, and the combining weights are not restricted, arguing that this has the advantage of producing an unbiased combined forecast even when individual forecasts are biased.

However, this has been questioned by a number of researchers. The theoretical work by Diebold [6] implies that using unrestricted least squares regression to combine forecasts will lead to serially correlated errors. His recommendation was to restrict the coefficients to sum to one. The same conclusion was made by de Menezes and Bunn [12] in a study to investigate the issue of serial correlation in combining methods.

5.3   Combining Results

It seems that there is still considerable disagreement concerning the best choice of combining rule. However, the M-Competition, a large scale forecasting competition with 1001 time-series, Makridakis et al. [11] and many other applications, Clemen [5] and Genre et al. [8], gave strong support for the simple average. This strong performance of the simple average may be due, as suggested by some authors, to the instability of "optimal" weights, resulting from changes in the individual forecast methods' performance over time. This may explain why a simple average, although having sub-optimal weights, may produce consistently good results.

We used a simple average to combine the three different sets of forecasts available (the 4-week moving average used by the company, the exponential smoothing and the ARIMA models developed), and compared the performance

Table 3. Model performance comparison.

2016	Observed values	Company forecasts	Exp. smooth forecasts	ARIMA	Combined forecasts
1	210170	195563	213706	217858	209042
2	182101	200109	173385	177080	183525
3	202687	197375	191788	196982	195382
4	183538	194809	196571	203926	198435
5	176603	188812	217957	172769	193179
6	148692	192680	183446	186774	187633
7	188945	194624	194643	201630	196966
8	185514	186232	186930	191736	188299
9	184927	177880	212852	217607	202780
10	182538	174445	177823	181638	177969
11	179680	174939	187624	192807	185123
12	212595	177020	186711	192316	185349
13	190533	185481	197910	203552	195648
14	194113	183165	193957	206776	194633
15	191991	189935	191605	198666	193402
16	202631	191337	204498	211532	202456
17	198675	194230	191943	199927	195367
	MAD	11826	11905	12301	9219
	$MAPE$	6.49%	6.60%	6.77%	5.16%
	$MSE \times 10^8$	2.65	2.92	2.53	1.94

of these combined forecast with forecasts obtained with the model used by the company (described in Sect. 2.2). The combined forecasts reduced the MAD by 22% (from 11 826 to 9219), the MAPE by 21% and the MSE by 27%. These results are shown in Table 3.

6   Conclusions

In an increasingly competitive market where companies are fighting for market shares that are hard to achieve, the pursuit of operational efficiency as a competitive advantage is essential. The volatility of the markets makes planning and replanning an endless task, in order to respond to more demanding customers. As a result, the pursuit of process optimization as a way to ensure better results is a major focus of attention.

This study allowed the formulation of some conclusions about the demand forecasting process in the company under study.

The exponential smoothing and the ARIMA models developed using Forecast Pro software demonstrated their applicability, and the results are considered satisfactory and able to serve as the basis for the company to predict demand.

It has been shown that combining these forecasts significantly improves accuracy. Simple approaches to combining provide adequate results and combining methods can be used with little increase in cost. Therefore, on the basis of simplicity and efficiency, we believe that the simple average of the forecasts produced by the three different models(4-week moving average, exponential smoothing and ARIMA) may provide a good solution for this forecasting problem.

In addition to the model performance against the current methodology applied by the company, other advantages can be pointed out, such as the fact that this is a more robust methodology, able in a timely manner to generate demand forecasts, and thus releasing the managers of this task, so that they can provide more time for the analysis of scenarios and to improve the decisions taken in this area.

References

1. Acar, Y., Gardner, E.S.: Forecasting method selection in a global supply chain. Int. J. Forecast. **28**(4), 842–848 (2012). https://doi.org/10.1016/j.ijforecast.2011
2. Bates, J.M., Granger, C.W.J.: The combination of forecasts. OR **20**(4), 451–468 (1969)
3. Bunn, D.W.: A Bayesian approach to the linear combination of forecasts. Oper. Res. Q. **26**(2), 325–329 (1975). (1970–1977)
4. Chan, C.K., Kingsman, B.G., Wong, H.: The value of combining forecasts in inventory management - a case study in banking. Eur. J. Oper. Res. **117**(2), 199–210 (1999). https://doi.org/10.1016/S0377-2217(98)00277-X
5. Clemen, R.T.: Combining forecasts: a review and annotated bibliography. Int. J. Forecast. **5**(4), 559–583 (1989). https://doi.org/10.1016/0169-2070(89)90012-5
6. Diebold, F.X.: Serial correlation and the combination of forecasts. J. Bus. Econ. Stat. **6**(1), 105–111 (1988). https://doi.org/10.1080/07350015.1988.10509642
7. Gaither, N., Frazier, G.: Administração da Produção e Operações. Thomson Learning (2007)
8. Genre, V., Kenny, G., Meyler, A., Timmermann, A.: Combining expert forecasts: can anything beat the simple average? Int. J. Forecast. **29**(1), 108–121 (2013). https://doi.org/10.1016/j.ijforecast.2012
9. Granger, C.W.J., Ramanathan, R.: Improved methods of combining forecasts. J. Forecast. **3**(2), 197–204 (1984). https://doi.org/10.1002/for.3980030207
10. Holt, C.C., Modigliani, F., Muth, J.F., Simon, H.A., Bonini, C.P., Winters, P.R.: Planning Production, Inventories, and Work Force. Prentice-Hall, Upper Saddle River (1960)
11. Makridakis, S., et al.: The accuracy of extrapolation (time series) methods: results of a forecasting competition. J. Forecast. **1**(2), 111–153 (1982). https://doi.org/10.1002/for.3980010202
12. de Menezes, L.M., Bunn, D.W.: The persistence of specification problems in the distribution of combined forecast errors. Int. J. Forecast. **14**(3), 415–426 (1998)
13. Winters, P.R.: Forecasting sales by exponentially weighted moving averages. Manag. Sci. **6**(3), 324–342 (1960). https://doi.org/10.1287/mnsc.6.3.324
14. Wong, W., Guo, Z.: A hybrid intelligent model for medium-term sales forecasting in fashion retail supply chains using extreme learning machine and harmony search algorithm. Int. J. Prod. Econ. **128**(2), 614–624 (2010)