Impacts of sales forecasting on the supply chain

Ana Dias
ana.cravo.dias@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

May 2022

Abstract

The supply chain enconpasses all flows of materials and information among the different members of the chain, beginning from purchasing raw materials until the sale of the final product by the consumer. In a *push* environment, demand anticipation is the force for executing chain processes, which asssigns a real importance to the sales forecast. In this context, the question of which impacts the forecast in the supply chain raises, making it the theme of this dissertation. The study was undertaken in *Grupo Boticário* in Portugal, where the areas affected by forecast were verified and solutions were found to reduce the negative effects in the Company and the supply chain in which it takes place.

To accomplish the goals established, it was necessary to understand in detail the processes related to the sales forecasting and how the supply chain of the Company operates. The theoretical knowledge of themes like supply chain management, inventory management and sales forecasting is essential for this work.

With the intention of reducing the impacts, some forecasting methods were applied and, then, assessed through indicators of error. Comparing all models allowed to find which method worked best, depending on the sample selected. Some improvements on the current procedure of sales forecasting were indicated.

No paragraph breaks.

Keywords: Sales forecasting, Supply chain, Forecasting methods, Inventory management

1. Introduction

A supply chain is characterized by all the members involved with the aim of fulfilling a costumer request. It starts form the purchase of the raw materials and ends when the client buys the final product. In between, beyond the physical flow of materials from the producer to the costumer, there is a flow of information and funds, in the other direction, contributing for its dynamic. [1]

As the main purpose of supply chain management is to coordinate the processes associated with the flows identified and to satisfy the costumer demand, trying to minimize the costs associated with every step included in the chain. [5]

The processes within a supply chain can be executed as a response to a costumer order (pull process) or anticipating the order based on a forecast (push process). In a push environment, the demand and sales forecast is the main force for all the activities in a chain. [2]

Is the importance of supply chain and its management, focusing on the role of sales forecasting in a push strategy that the problem discussed in this work arise. The major purpose is to find and analyse the impacts that sales forecasting have on the supply chain and, then, discover ways to maximize the accuracy of the forecast and to minimize the negative effects that can occur within the chain. The study was done in a cosmetics and perfume company environment, Grupo Boticário Portugal, experienced the work of the Department of Planning.

1.1. Grupo Boticário Portugal - Suppy chain and sales forecasting

As a Brazilian company, which expanded to several other geographies, one being Portugal, all products from the company's brands are produced in two factories in Brazil. The team in Brazil receives a monthly order from Portuguese Department of Planning, that it has to be fulfilled and shipped. The distribution is done through maritime transport, with Rotterdam, The Netherlands, as destination. From there, the products are delivered by truck until the wharehouse, located in Montijo, Portugal. This whole process can take up to three months to be completed. If this operation cannot be that long-lasted, the expedition that was done by maritime transport can be replaced by plane, reducing the process to approximately one month. From the wharehouse, the products destined to stores are distributed, where the Department of Logistics decide what the quantity rational, and for the online store and the resales (or *venda direta (VD)*, as it is called in the company), the orders are expedited directly to the costumer or the reseller.

First, it is necessary to understand some concepts used within the company. A year is divided by cycles, periods of different length, based by the themes or products in focus during this time and where distinct types of promotions are arranged and product are launched. The Marketing team is the one that decides these types of information, combining it all in a document called "grade".

Before using the sale history of the products, these values are converted to a average of units sold in each cycle, with the aim of being comparable, because they are in the same calculation basis. For the case of physical and online stores, the unit used is called MDL, and for the case of VD, UPA is applied. These are named fundamental unit of measurement, estimated taking into account the number of days in the cycle and the number of stores in operation, for MDL, and the number of resellers, for UPA.

Every single product from the company's portfolio are listed in a document named "master", where all information related to each product is organized. One of the most important characteristic is the *chave*, a code that refers to a specific product. The information also includes the status (where in the cycle of life the product is); the category which it inserts and more segmented details as the subcategory and the brand; and the price.

Focusing now on the most important task for the present study, performed by the Department of Planning, the sales forecasting is divided in four main groups: baseline, promotions, offers and launching.

The baseline, the regular sales without any type of promotional event, is built through the average of MDL or UPA of the last 3 cycles with no promotion. Sometimes, having a critical sense, if there are values very different from the rest, it is possible to use older cycles. The baseline estimated is applied as the sales forecast of the products that are not included in the "grade" of the cycle in question.

If the product is a part of the "grade", the process is different. The database used in this case is the total historical sales (called Sell Out), where all the information about the cycles, the promotions and it is stated if the product was out-of-stock or was a launching.

To forecast the sales of a promotional products, it is necessary to take into account the type and level of promotion that it is going to received. If that type and level was already on the database for the same or similar product, the found value of MDL or UPA can be used. If not, a rational over the baseline can be used, achieved through the performance analysis of comparable promotions in products of the same characteristics. For offers, the same process is operated, verifying if there was an offer similar and collecting that value.

For the last group, to forecast the launching of a new product is essential to know the moment it is launch and what is the similar product, treated as the bench mark. To this product, the database of its first cycle is checked. It is important to consider also the initial values of the same category and subcategory products.

The forecast done by the methods described is not yet complete, the values obtained are normally estimated by excess. Taking the budget of the company for each month as an objective value, the forecast of a series of product is reduced, maintaining the estimated value for items in the class A+; product in the top 20-30% of the expected revenue; launching, promotional and offered products; and end of life products.

2. Methods

There are two different types of forecasting methods [2]:

- Quantitative methods: Take as an assumption the future repetition of the patterns seen in the past. More suitable for short and medium term forecasting. There are two groups of methods: causal and non causal models.
- Qualitative methods: Subjective methods based on the experienced and intuition of who is making the forecasting and also using analogies and similar situations. More suitable for long term forecasting.

2.1. Causal models

Causal models are statistical techniques of correlation and regression, in order to recognize and establish cause-effect relationships between the chosen variables, through their historic data. The difficulty finding variables that effectively has a causal effect on the variable to be predicted and the length of time it takes to collect all the data. [2]

2.2. Non causal models

Non causal models only use historical series of the variable to forecast by analyzing the data and looking for patterns in behavior that can be projected into the future.

A time series can be defined by different characteristics:

- Level: Base value between which the series stays around;
- Trend: The linear increasing or decreasing behaviour in the time series;

- Seasonality: Repeating patterns or cyclic behaviour that causes the increasing or decreasing patterns for a short time in the times series;
- Noise: Random variations that cannot be explain by any model.

The following methods are used to predict the patterns, the most accurate possible. If not mentioned, the forecast for k periods forward of the instant t ($\hat{Y}_{t+k}$) is represented by the level at t, as demonstrated in the following equation.

$$\hat{Y}_{t+k} = n_t \quad , \text{with}\, k = 1, 2, \ldots \tag{1}$$

2.2.1 Arithmetic moving average

Is a method where the forecast value correspond to the average of the most recent observations. Suitable for time series with no trend or sazonality. The level at the instant t is calculated by the equation 2, in which N correspond to the number of the most recent observations and Y_i the sales observed in the instant i. [2]

$$n_t = \frac{1}{N}\left(Y_t + Y_{t-1} + \ldots + Y_{t-N+1}\right) \tag{2}$$

2.2.2 Weighted moving average

In this method the observations have different weights, according to the proximity to the present value. To the most recent data is assigned more weight. [2]

The final equation for the method is the equation 3, where a_0 represents the weight for the most recent observation and β is a parameter, between 0 and 1.

$$n_t = a_0\left(Y_t + \beta Y_{t-1} + \beta^2 Y_{t-2} + \ldots + \beta^{N-1} Y_{t-N+1}\right) \tag{3}$$

2.2.3 Simple exponential smoothing

Model suitable for time series with stationary trend and no seasonality. With the previous forecast, the method adjusts it to obtain the new forecast, through the equation 4, where α is a smoothing parameter, between 0 and 1. [2]

$$n_t = \alpha Y_t + (1-\alpha) n_{t-1} \tag{4}$$

2.2.4 Holt model

It is a model also suitable for time series with trend and no seasonality. Comparing to the last method, it introduces the trend component b, which is a a smoothing parameter, between 0 and 1, represented by the equation 6. The level is calculated by the equation 5. [3]

$$n_t = \alpha Y_t + (1-\alpha)\left(n_{t-1} + b_{t-1}\right) \tag{5}$$

$$b_t = \beta\left(n_t - n_{t-1}\right) + (1-\beta) b_{t-1} \tag{6}$$

the forecast for k periods forward of the instant t is represented by the expression **??**.

$$\hat{Y}_{t+k} = n_t + k \cdot b_t \tag{7}$$

2.3. Measures of errors

Evaluate the performance of each method studied is fundamental to the choice of the one to use for the data, but also in monitoring its efficiency, during the application.

The measures more commonly used are the mean squared error (MSE) and the mean absolute error (MAE), which equation are 8 and 9, respectively. The m are the number of periods used. [3]

$$MSE = \frac{1}{m}\sum_{t=1}^{m}\left(Y_t - \widehat{Y}_t\right)^2 \tag{8}$$

$$MAE = \frac{1}{m}\sum_{t=1}^{m}\left|Y_t - \widehat{Y}_t\right| \tag{9}$$

The other two measures used are based on perceptual errors, making it possible to compare forecasts of different scales. The measures are the mean absolute percentage error (MAPE) and weighted absolute percentage error (WAPE), which are represented by the equation10 and 11, respectively. The WAPE is different from the others, because it doesn't consider the observations equally important, making the most recent ones more weighted.

$$MAPE = \frac{1}{m}\sum_{t=1}^{m}\frac{|Y_t - \widehat{Y}_t|}{Y_t} \times 100 \tag{10}$$

$$WAPE = \frac{\sum_{t=1}^{m}|Y_t - \widehat{Y}_t|}{\sum_{t=1}^{m} Y_t} \times 100 \tag{11}$$

3. Results

3.1. Sample choice

The first step to take when applying a method is choosing the sample. The products to be analyzed must be distinct so that the sample is representative of the entire company's portfolio.

Using the ABC classification, two class A+ products were chosen, as they are part of the group with the greatest contribution to the overall value of the Company's revenue, and one item from the classes A, B, C and D. Their status is also considered, one of the products is in the end of its life (status of to be discontinued). The sample's information is represented in the table 1.

Table 1: Sample choosen to study the application of the different methods.

Chave	Description	Category	Class ABC	Status
597	Moisturizing cream	Skin care	A+	Active
592	Perfume	Perfumery	A+	Active
4027	Body lotion	Skin care	A	Active
2097	Deodorant	Deodorants	B	Active
4757	Shampoo	Hair care	C	To be discontinued
5209	Eye palette	Make up	D	Active

3.2. Data treatment

After choosing the products for which the study of the forecasting methods will be developed, it is important to gather all the information necessary for their application and making sure that the data are on the same basis, in order to compare them in a reliable way.

The study was carried out only for the elaboration of the baseline, since it is the procedure that requires less attention when it is performed, thus indicating more fragility in the confidence of the results obtained.

In this way, by removing the time series of each product, data from the cycles in which promotional events took place were eliminated. For cycles where the historical values are considered abnormal, comparing with the others, they were removed from the data to use.

The lifetime of each item is different, so the horizon studied may not be the same for all. The cycles are only numbered and not referring to the year and number, as usual.

The data used were taken from the sales history, already represented by the fundamental measurement units. For the items of classes A+ and A, the MDL values were studied and for the others, the UPA values were chosen.

3.3. Methods application

For each product it was applied the four methods previously mentioned, using the equations shown, and after, for a correct evaluation of the results, the different types of indicators of error were calculated. For a resumed presentation of results, only the WAPE will be displayed. When the method requires parameters, they were achieved through the solver tool on Execel, where the desire was to minimize the WAPE and taking into account the limits for these components.

3.3.1 Arithmetic moving average

The application of the arithmetic moving average method led to the results presented in Table 2.

Analyzing the results obtained, it is possible to affirm that the number of observations that result in the minimization of error varies greatly from item to item.

For the most part, 3 observations produce the least error. In case of product with *chave* 592, $N = 5$ is preferable. The case in which a larger N is more desirable is the case with *chave* 4027.

Table 2: Values of WAPE from the application of the arithmetic moving average for different values of N.

Chave	Class	N=3	N=5	N=7
597	A+	32,9%	37,1%	39,3%
592	A+	41,7%	38,9%	43,9%
4027	A	38,2%	36,1%	27,2%
2097	B	30,5%	33,9%	37,7%
4757	C	36,2%	41,9%	41,0%
5209	D	58,4%	61,1%	58,6%

The very high indicator value for the *chave* 5209 are mainly due to the small number of historical data and their randomness. The product 2097 is the one that presents, mostly, the smallest errors, due to the tendency that is expressed.

3.3.2 Weighted moving average

For the method of weighted moving average, the values of WAPE are presented in the Table 3.

Table 3: Values of WAPE from the application of the weighted moving average for different values of N.

Chave	Class	N=2	N=3	N=4	N=5
597	A+	-	30,6%	30,2%	30,4%
592	A+	42,4%	40,0%	-	37,6%
4027	A	44,1%	-	35,2%	-
2097	B	-	26,4%	26,1%	25,8%
4757	C	30,4%	31,3%	30,0%	29,9%
5209	D	62,5%	-	-	-

The results for this method are not considered very satisfactory, due to the difficulty, in certain cases, of reaching the optimized parameters to have the smallest possible error associated.

The products with *chave* 4027 and 5209, whose time series are the shortest, have the fewest possible cases. Analyzing the results, the greater number of observations always produces the most minimized measure.

3.3.3 Simple exponential smoothing

The errors calculated, when the application of the simple exponential smoothing method is done for the sample, are in the Table 4.

Table 4: Values of WAPE from the application of the simple exponential smoothing.

Chave	Class	WAPE
597	A+	28,5%
592	A+	39,0%
4027	A	31,2%
2097	B	26,0%
4757	C	28,8%
5209	D	46,8%

From the results presented, it appears that for the items with *chave* 4027 and 5209, the forecast is almost constant, due to the fact that the optimized parameter α is very small. In these cases, very little weight is given to the most recent observations. A possible explanation must be the reduced number of historical data used.

In the remaining cases, where the application of the method was more successful, it was with the product of class B that the error was the lowest.

3.3.4 Holt model

The last method studied was the Holt model. The results of the error for the application are represented in the Table 5.

Table 5: Values of WAPE from the application of the Holt model.

Chave	Class	WAPE
597	A+	31,5%
592	A+	50,4%
4027	A	46,8%
2097	B	28,4%
4757	C	28,6%
5209	D	46,7%

It is shown again that the item with *chave* 5209 presents an almost constant forescast, where the parameter β has been assigned the value 0.

Regarding the other products, the one that presented the most minimized WAPE was the *chave* 2097.

3.4. Method comparison

Finally, it is interesting to compare, within the study of a product, which method best fits and, in the sample, which method presented the best results.

The arithmetic moving averages method, in general, is the one that presents the worst results, being only for the class A item that this method with $N = 7$ has the lowest WAPE. The Holt model presents, for the products 592 (A^+) and 4072 (A), also unsatisfactory results.

Regarding the method of weighted moving averages, despite its difficulty in finding values for the cases studied, it presents satisfactory values, especially for N greater than 2.

However, the method that unequivocally gives consistently low values of WAPE for all products is the simple exponential smoothing method.

Analyzing item by item, what presents the biggest error in most of the methods is the product of class D, due to the fact that it has a reduced amount of historical data and, those that exist, have a great variability. As it belongs to class D, it contributes little to the global revenue and has very low sales, so small variations in sales greatly affect forecasts.

The item of class B is the one with the lowest WAPE values.

4. Conclusions

The sales forecasting at the Company takes a vital part in the proper functioning of the business, since it influences the supply chain, from one end to the other. The system studied is defined as push, where the flow of materials from the origin to the point of sale is determined based on the estimation of future sales. Variations from the consumer can, therefore, affect the supply chain, in all the areas involved.

In the chain, there are also flows of information between the members. If some kind of distortion occurs, the chain is affected in different ways, according to the bullwhip or whiplash effect. This effect concerns two phenomenas: demand distortion, where the request to the supplier tends to have a greater variation than the sales to the final consumer; and variation amplification, in which the distortion propagates upstream in an amplified way. [4]

The impacts found were divided into three different groups: internally, in the Department; upstream; and downstream.

Analysing the processes practiced by the team, there are some improvements that could be applied to achieve a better forecast. They use mostly Excel, which has its limitations for more complex tasks and, with a very large portfolio of products, there is a high probability of errors occurring. It would, therefore, be important to implement an automated system, making the processes faster and with better performance.

Regarding the baseline, there is some methods more suitable than the one currently implemented. From the results, it is concluded that the simple exponential smoothing model produced the most minimized errors, from the different ones studied. The application requires the optimization of the smoothing parameter each time there is a new historical data, so the method needs continuous update. Using only the solver tool in Excel, it becomes time consuming to update it for all products. Thus, the items with the greatest variations between the forecast and the actual sale must be analyzed, choosing this group for the application of this model. For the rest, the method of arithmetic moving averages should continue to be used, which, according to the results, managed to minimize the errors well.

Another way of improving forecasting processes, both for baseline and for promotional and launched products, would be the development of a database, to be able to compare the growth of MDL/UPA values.

For the baseline, from the historical database, the values from cycles with promotional events or launching or out of stock situations should be removed. This new database would allow the application of the forecasting method in a simpler way, showing only the values that really matter for the process. With all the values gathered, the analysis of seasonality can also be useful, segmenting the values by the characteristics of each product, such as category, subcategory, brand and price. The seasonality value could then be applied to the baseline, according to the cycle to be forecasted and the product's information.

Regarding promotional events and launches, the databases would be composed of historical values from the cycles in which the items had promotional events and in which they were launched, respectively. In the case of promotions, the database would be segmented by promotion typology and discount level and also by their characteristics, where it would be possible to analyze the growth verified in comparing it to the baseline. As for launches, the segmentation of the base would only be according the characteristics of the item and, when necessary, it would be analyzed for the product marked as its benchmark (product similar to the one being launched).

The forecast still undergoes a final adjustment. However, the following effects are not taken into account: cannibalization effect and hangover effect. In order to adjust the values obtained considering these effects, it would be necessary to analyze, in the case of cannibalization, the historical behavior of products that were the target of some promotions or were launched and that compete with others without associated events and, in the case of hangover, what is the variation of sales in the cycles followed by some type of promotion or launch. Using the results of these analyses, it would be possible to regulate the baseline of the items that are the target of the described effects.

Still internally, the sales forecasting process can be compromised if the communication between the teams of the Company is not effective or constant, and if changes in each of one's work are necessary, these should be transmitted in a timely manner, so that there is still time to react.

Considering now the impacts upstream of the chain, if there is any significant change in the frozen forecast compared with the order, the team in Brazil may not have enough stock available and may not be able to supply the product that was guaranteed. In this type of situation, changes in the forecast can mainly affect the installed capacity and the purchase of raw materials in the industry.

Where there may also be constraint, in the scenario where forecast was higher than what was expected, is in the reservation of containers for the transport from Brazil. Due to space limitations, it may not be possible to attend a significant increase in product in a timely manner.

Analyzing the impacts of the forecast downstream of the supply chain, variations in the forecast lead to a risk of supplying the sales channels and, therefore, customer consumption may be compromised. In terms of stock management, the risk of rupture and loss of sales increases or, on the contrary, there is a higher risk of stock accumulation, which leads to possible loss due to obsolescence and increased storage costs.

There is still one more financial impact of the sales forecast. In the situations where more or less product is sold than what was expected, there is always capital at risk and, therefore, it must be verified in any circumstance that arises. The management of orders and stocks is based on the sales forecast. If the forecast, and consequent order, is much higher than reality. the product in stock will increase, increasing the possibility of loss of stock, cost of holding stock and capital will remain still. On the contrary, if the forecast and the order fall short in relation to the real demand, sales will be compromised and there could be a risk of rupture, thus also causing monetary losses.

In conclusion, what is important to take away from this study is that all members of the supply chain are directly linked, and in a push system, even greater importance is given to the role of sales forecasting.

References

[1] S. Chopra and P. Meindl. *Supply chain management : strategy, planning, and operation.* Pearson, 5th edition, 2013.

[2] A. P. Guedes, A. Arantes, and R. Oliveira. *Logística e Gestão da Cadeia de Abastecimento.* Edições Sílabo, Lda., 3 edition, 9 2020.

[3] R. J. Hyndman and G. Athanasopoulos. *Forecasting: principles and practice.* OTexts, 3rd edition, 2021.

[4] H. L. Lee, V. Padmanabhan, and S. Whang. Information distortion in a supply chain: The bullwhip effect. *Management Science*, 43:546–558, 1997.

[5] M. A. Mahmood, L. A. Gemoets, and A. O. Solis. Supply chain management. *Encyclopedia of Information Systems*, 4, 2003.