Implementation of demand forecasting models for fuel oil
- The case of the Companhia Logística de Combustíveis, S.A. -

Rosália Maria Gairifo Manuel Dias

1DEG, IST, Universidade Técnica de Lisboa, Lisboa, Portugal

Abstract

The management of the oil supply chain has raised impact in the world-wide economy nowadays. Industries, services and oil dependent individuals live under pressure with the capricious fluctuations of the oil prices. Consequently, with the decrease of the petroleum products consumption, the petroleum industries, customers and suppliers try therefore to develop tools to forecast the future demand, in order to reduce the risk and adjacent damage to the existing variability. In this thesis they are approached some methods that companies currently apply in demand forecasting. Additionally, it is developed a case study analysis based on the implementation and development of forecasting models adapted to the demand of the seven products that CLC - Companhia Logística de Combustíveis, S.A. – distributes. The analysis addressed in this work will enable CLC’s decision makers to estimate the consumption of their customers in order to accomplish more effectively and efficiently the petroleum products planning the company sells.

Keywords: fuel oil supply chain, demand forecasting, exponential smoothing

1. Introduction

The demand forecast is a key element in the decision making of management activities, having as main goal to predict the factors that influence their performance and outcomes.

The implementation of the various forecasting methods is frequently done in areas like stock management, production planning and process control. The companies commitment, that actually play in the market, to reduce the uncertainty dependency of their decisions and to obtain a better risk management is translated in increasing investments in appropriate methods of demand forecasting.

Companhia Logística de Combustíveis (CLC) is a company dedicated to storage and distribution of oil products. The products are made available at CLC’s distribution center by a multi-product pipeline. One of the company’s main activity in its operational level is to establish accurate levels of inventory that satisfy customer demands [GalpEnergia, BP, e Repsol]. Note that the ideal situation for CLC has been to store only the quantities needed to satisfy them. CLC has a single supplier, the owner of Sines refinery, GalpEnergia: CLC’s planning has a major weight on the refinery planning since it represents around 40% of the total production of the refinery. This planning is done on a monthly basis, based on calendar months, and is made available to refinery planners by CLC to the managers responsible for the planning in Galpenergia. In order to develop the pipeline schedule, the company needs to receive (always in the middle of previous month) from all its customers a monthly forecast of your daily needs of each product. After this task, the CLC begins to develop the schedule of the pipeline, in parallel with the inventory management plan of all the storage tanks. These plans ensure that there is enough product level to satisfy the customer at any time, according to the quality requirements and the capacity available to receive all products from the pipeline.

Currently, the CLC uses the Microsoft Excel® and the experience of its schedulers to carry out a good monthly plan. For this purpose, they were developed two main working files that constitute the entire planning and scheduling system of CLC: Pump Order and Tank Moving.1,2

The main objective of doing this paper relates to the fact that it is essential for CLC to develop methods that allow the prediction of future consumption, so as to avoid any problems of stockout or product obsolescence (due to chemical properties). That happens because all their planning process is
totally dependent on the values sent by their customers. The present work aims at optimizing the product planning and scheduling that CLC store, through the study of appropriate forecasting methods\(^3,4,5,6\) for each product managed by CLC. The forecasting models based on exponential smoothing methods\(^7\) will be processed using Microsoft Excel\(^8\), and the results are then compared with real output data from the distribution center and the data sent by the customers themselves.

The rest of the paper is structured as follows. Section 2 presents the problem definition. In section 3 it is made a review of the state-of-art in the field of forecasting models based in exponential smoothing and its role on operations management. Section 4 shows the implementation and analysis of the three models implemented. Section 5 analyses the results and refers the main advantages and disadvantages of the forecasting models used. Section 6 refers some relevant work to do in the future developments. Some conclusions are outlined in section 7.

**2. Problem Definition**

As referred in section 1, CLC’s planning method is based solely on forecasts that customers send with their needs for the next month. Currently, with fluctuations in oil prices and consumption in the fuel market, the company feels the need to develop and implement a method that allows a comparison between what customers forecast and the real product outputs (product quantity that CLC customers real needs). This work has as objective to study all demand forecasting methods and apply them to each product distributed by CLC, to enable the company store only the required quantities and make the right pipeline pumping scheduling of every products.

The demand forecasting methods implementation is based on the analysis and understanding of the historical data provided by CLC, regarding forecast and real outputs. To identify and understand patterns in the data it is important to perform a graphical and numerical analysis\(^9,10\). The numerical analysis will be used to determine the mean values of quantities and the spread of data (deviations from the mean value). After analyzing the data of each product, the proper forecasting methods are applied based on the specific characteristics identified so as to to obtain more accurate forecasts. The methods used in this case study are referred in section 4.

**3. State-of-the-art of forecasting models based in Exponential Smoothing for fuel oil**

The application presented in this paper is a pioneer work in the portuguese academic community. These methodologies are frequently used in the distribution companies, and they are very important to help the Operational Department to planning their inventory. From the research within literature, no similar case studies on the oil business were found using exponential smoothing based case studies.

All publications cited in this case study describe methodologies for the implementation of demand forecasting methods for any kind of product. The principal methodologies used by companies are models based on the exponential smoothing\(^4,11,12\). For example, this type of methods are majorly used for the calculation of company sales\(^5\), or to forecast the inputs/outputs in the supply chain\(^4,12\). This type of model is able to calculate estimates for small time horizons, with little historical data, not requiring any hardware with large computing power and small know-how of users. Microsoft Excel\(^6\) tool has functions that help and support the implementation and required calculations of demand forecasting, being this tool chosen to use in this case study\(^6,9\).

Nowadays, it is clear that there is not a single best approach for all problem types and researchers are still drawing the map that will answer which model to use as a function of the characteristics of the real problem.

**4. Implementation and analysis**

**4.1. Analysis**

For the graphical analysis it is required to build seasonal and time plots to identify all the patterns in the data. Four types of data patterns can be distinguished\(^12\): horizontal (when the data values fluctuate around a constant mean); seasonal (when the data is influenced by seasonal factors); cyclical (when the data exhibit rises and falls that are not of a fixed period); and trend (when there is a long-term increase or decrease in the data).

The numerical analysis recurs to the most common descriptive statistics such as the mean, the standard deviation, and the variance, which are used to develop numerical summaries. To measure the data spread it is used the mean of the absolute deviation (MAD) and the mean squared deviation (MSD).
Among the seven products distributed by CLC they have been identified 3 different patterns and, consequently, applied 3 different exponential smoothing models: in the Diesel data set was identified a horizontal pattern and applied the method of single exponential smoothing; in Unleaded Gasoline 95 and 98 was identified a downward trend pattern and applied the Holt Linear Model; in Butane and Propane the pattern identified was the seasonal pattern, and the model applied was the Holt-Winters trend and seasonality method; in Farming Diesel and Jet-A1 was not identified any pattern, but out of curiosity we applied the method of single exponential smoothing to observe the behavior of the model in this kind of product.

4.2. Models description

4.2.1. Single Exponential Smoothing

The method of single exponential forecasting\textsuperscript{12,13} takes the forecast for the previous period and adjusts it using the forecast error. Thus, the forecast for the next period is given by:

$$F_{t+1} = F_t + \alpha (Y_t - F_t) \quad (1)$$

where $\alpha$ is a constant between 0 and 1, $F_t$ is the most recent forecast, $Y_t$ is the most recent observation and $F_{t+1}$ is the forecast that we want calculate.

It can be seen that the new forecast is simply the old forecast plus an adjustment for the error that occurred in the last forecast. When $\alpha$ has a value close to 1, the new forecast will include a substantial adjustment for the error in the previous forecast. Conversely, when $\alpha$ is close to 0, the new forecast will include reduced adjustments.

Another way of writing (1) is:

$$F_{t+1} = \alpha Y_t + (1 - \alpha) F_t \quad (2).$$

4.2.2. Holt Linear method

Holt (1957) extended single exponential smoothing to linear exponential smoothing to allow forecasting of data with trends. The forecast for Holt linear method\textsuperscript{12,14} is determined by using two smoothing constants, $\alpha$ e $\beta$ (with values between 0 and 1), and tree equations:

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1}) \quad (3)$$
$$b_t = \beta (L_t - L_{t-1}) + (1 - \beta)b_{t-1} \quad (4)$$
$$F_{t+m} = L_t + b_t m \quad (5)$$

Here $L_t$ denotes an estimate of the level of the series at time $t$ and $b_t$ denotes an estimate of the slope of the series at time $t$.

Equation (3) adjusts $L_t$ directly for the trend of the previous period, $b_{t-1}$, by adding it to the last smoothed value, $L_{t-1}$. Equation (4) then updates the trend, which is expressed as the difference between the last two smoothed values. This is an appropriate procedure because if there is a trend in the data, the new values should be higher or lower than the previous ones. Finally, equation (5) is used to forecast ahead. The trend, $b_t$, is multiplied by the number of periods ahead to be forecast, $m$, and added to the base value, $L_t$.

4.2.3. Holt-Winters trend and seasonality method

Holt method was extended by Winters (1960) to capture seasonality directly. The Holt-Winters method\textsuperscript{12,14} is based on three smoothing equations – one for the level, one for trend, and one for seasonality. In fact there are two different Holt-Winters methods, depending on whether seasonality is modeled in an additive or multiplicative way, but in this work only the multiplicative way is used.

Multiplicative seasonality

The basic equations for Holt-Winters multiplicative method are as follows:

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + b_{t-1}) \quad (6)$$
$$b_t = \beta (L_t - L_{t-1}) + (1 - \beta)b_{t-1} \quad (7)$$
$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s} \quad (8)$$
$$F_{t+m} = (L_t + b_t m)S_{t-s+m} \quad (9)$$

where $s$ is the length of seasonality (e.g., number of months in a year), $L_t$ represents the level of the series, $b_t$ denotes the trend, $S_t$ is the seasonal component, and $F_{t+m}$ is the forecast for $m$ periods ahead.

4.2.4. Initialization

The reason why initial values for the exponential smoothing methods are needed is that the methods are recursive equations, and so they need to be initialized somehow. In this work the method used to calculate the initial values is the Least Squared Estimates\textsuperscript{15}. In single exponential smoothing $F_t$ can be found by averaging the past values. For methods that assume a trend, a straight line can be fitted to the first few values of the slope and intercept used for initial values.

4.2.5. Measuring forecast accuracy

Usually, accuracy is treated as the overriding criterion for selecting a forecasting method. In many instances, the word
“accuracy” refers to “goodness of fit”, which in turn refers to how well the forecasting model is able to reproduce the data that are already known. They exist a variety of measures for forecasting accuracy. If there are observations and forecasts for \( n \) time periods, then there will be \( n \) error terms, and the following standard statistical measures can be defined:

\[
\text{Mean Error} \quad ME = \frac{1}{n} \sum_{t=1}^{n} e_t \\
\text{Mean Absolute Error} \quad MAE = \frac{1}{n} \sum_{t=1}^{n} |e_t| \\
\text{Mean Squared Error} \quad MSE = \frac{1}{n} \sum_{t=1}^{n} e_t^2 \\
\text{Squared Root of Mean Squared Error} \quad RMSE = \sqrt{MSE} = \left( \frac{1}{n} \sum_{t=1}^{n} e_t^2 \right)^{1/2}
\]

The most important, and that really influences the decision of the best method is the Mean Squared Error. How much lesser it will be the value of MSE, the better is the accuracy of the method applied.

4.3. Implementation by product

4.3.1 Diesel

The diesel values range from a mean value over the time interval. The application of single exponential method proves to be the most effective method for calculating the forecasts of diesel consumption. At the same time, it was applied the Holt linear method to prove which method has the best accuracy. The choice of this comparison method had to do with fact that in the year 2009 the output of diesel decreased compared to previous years. The comparison showed that single exponential method is better than Holt linear method to predict diesel demand, because your mean squared error is lower. The results of applying this model are shown in Figure 1.

![Figure 1. Diesel results - Forecast values vs. Real outputs vs. needs of clients (m³)](image)

4.3.2. Unleaded Gasoline 95

In the study of Unleaded Gasoline 95 two methods were implemented: the Holt linear method and the single exponential smoothing method. From their comparison it was concluded, as expected, that the method that better predict the Unleaded Gasoline 95 consumption is the Holt linear method, since this product presents a pattern of decreasing trend. The forecast values were compared with the values of the real outputs and with the values provided for CLC by the customers, and can be seen in Figure 2.

![Figure 2. Unleaded Gasoline 95 results - Forecast values vs. Real outputs vs. needs of clients (m³)](image)

4.3.3. Unleaded Gasoline 98

Similarly to Unleaded Gasoline 95, also in Unleaded Gasoline 98 they were implemented the same two methods: Holt’s linear method and single exponential method, due to the fact that this product also presents a decrease pattern, obtaining exactly the same conclusion. As in the previous product, the MSE obtained from the linear method was lower than the MSE obtained from the single exponential smoothing, concluding that the first method has the best performing in the calculation of forecasts of Unleaded Gasoline 98. The forecast results are shows in Figure 3.

![Figure 3. Unleaded Gasoline 98 results - Forecast values vs. Real outputs vs. needs of clients (m³)](image)
4.3.4. Butane

The analysis of the available data from the monthly outputs of Butane enabled the detection of a seasonality pattern and, consequently, the implementation of the Holt-Winters method. However, it is still necessary to select which of the Holt-Winters methods is best suited to this product, applying the two methods to Butane and comparing the values of Mean Square Errors. This comparison considered an $s$ period of 3 months, concluding that the method that best predicts the Butane demand is the multiplicative method. Then, it was necessary to compare different values of $s$, for the seasonal period presented by Butane. However, there isn’t a stable period, ranging from 3 in 3 months, 4 in 4 months and from 5 in 5 months. After calculating their forecasts and adjacent errors, it was concluded that the seasonal period would be 4 months; it is the one with a smaller MSE value, producing forecasts closer to actual output of this fuel. The results obtained with the application of this method can be seen in Figure 4.

4.3.5. Propane

As in the case of Butane, in Propane’s forecasting was essential to compare the two models of Holt-Winters, in order to choose the most accurate. The multiplicative model continues to have a lower mean square error, after applying this method to this gas. In the analysis of data from the outputs of Propane they were identified 3 distinct periods in which the values vary seasonally, ranging from 2 in 2 months, 3 in 3 or 4 in 4 months. The period that best predict the propane consumption, by analysis of the errors squared average of the 3 periods, is 2 months. The results are shown in Figure 5.

4.3.6. Farming Diesel

In farming diesel, despite the reduced and unreliable available historical data, it was applied the method of single exponential smoothing to observe the behavior of the model in this kind of product, presented in Figure 6.

4.3.7. Jet-A1

In the graphical analysis of the product Jet-A1 it was not possible to identify a pattern of specific data, observing only the peaks of higher consumption. Comparing the results obtained by the two methods, the single exponential smoothing is the best method to predict the monthly output of Jet-A1, resulting in more accurate forecasts. The results obtained with the application of this method can be seen in Figure 7.
5. Results analysis – advantages and disadvantages

Currently, most companies use forecasting methods to help in their decision making. However, the forecasting methods are simple tools to be used as support but are not always 100% reliable. The main reason relies on the usage of past data that may not correspond to reality today or even the quality of the information used.

The main disadvantage of using these methods, despite their advanced level of development, is that the outcomes may not match the actual demand. To work around this limitation, it is crucial in the case of CLC, to use the data provided by the models to compare them with data from customers and, with CLC schedulers directly involved in the planning and scheduling of the transportation and storage tasks, to shape and adapt the quantities of each product to be made available.

Another of the limitations of the models developed in this study is that they do not consider economic factors\(^\text{16}\) regarding the oil market, since world economy is still heavily dependent on oil derivatives.\(^\text{17,18}\) For individual situations, such as the accident that took place in Sines refinery in January 2009 and the interruption of the supply of butane in September 2008, prediction may be impossible. In these situations, any model used for forecasting demand, although well-designed and adapted to the product, would not be able to do that.

However, the implemented models, except for Jet-A1 and Propane, delivered optimized forecasts when compared to the forecasts given by the customers, being very useful for CLC, to support their operational planning. Additionally, the methods based on the methodologies of exponential smoothing use, in each new application, reduced amounts of data, since in the majority of the cases to foresee the consumption of the following period, it is only necessary the values of the consumption of the previous period and the corresponding forecast. In this way, this method requires reduced computational effort, being also possible to develop easy-to-use interfaces, based on the simplicity of the developed forecasting models presented in this work.

6. Future work

The present work can be seen as kick-off work regarding the application done in the area of oil and oil derivatives forecasting. This industry consists of a complex area where the amount of factors that can influence the demand of a product may be hard to specify, unknown or with a broad frontier. Even for known factors, the quantification is a difficult task, since the application of forecasting methods in this area are scarce, being this study a pioneering approach. It is indispensable to have follow-up studies so as to improve the presented approaches, namely including some of the factors that can influence the results, in order to minimize the errors and to improve the forecasts accomplished by the algorithms.

Another aspect to be improved in future works is the development of an algorithm that optimizes the smoothing constants so as to supply better values of mean quadratic error, and therefore forecast values even more accurate. In this study, we tested values of smoothing constants ranging between 0.1 and 0.9, but the ideal would be to find more precise values to obtain predictions even closer to the actual values.

After the accomplishment of this work, would also be very interesting to develop an user interface, so that CLC’s employees could use this tool for planning and scheduling the amounts of product to manage, even without deep knowledge on forecasting techniques.

7. Conclusions

This paper describes a planning and scheduling inventory problem, involving short-term decisions, in particular for calculating the quantities to be stored for seven products that CLC is responsible for distributing in Portugal. The oil sector is a business area that has an important role in the economy of our country, due to the high level of oil dependency still verified.

This work gave a greater emphasis on forecasting methods that best suited to the products involved and to the Operational Planning of a real world company. Within these methods, exponential smoothing based methods can be found, which are easy to use and produce, in general, highly accurate data in the presence of historical data over a short time horizon periods, up to one month.
The exponential smoothing models studied in this work are models that provide forecasts without significant errors, for short forecast horizons, meeting the objectives of CLC in developing forecasting tools.

The main purpose of this study consisted of the implementation of a customized model for forecasting demand for each oil derivative, in order to optimize the current schedule of transported quantities from the refinery, and stored in the distribution center. On a second level, this work compared the results obtained with the available values of actual outputs, but also with the values of the forecasts given by CLC customers, and which serve as the basis for the the current schedule development procedure. The objectives of the various models used were to calculate the optimal amount (the most similar to the real values) to minimize the forecast errors. These methods and outcomes can be used by CLC to contrast the amounts requested by their customers, which in most products remain inappropriate to what they really load at CLC’s facilities.

References