

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
UNIVERSITY MOHAMED BOUDIAF - M'SILA

FACULTY: MATHEMATICS AND INFORMATICS

DEPARTEMENT : COMPUTER SCIENCE

N° :



DOMAIN : MATHEMATICS AND COMPUTER
SCIENCE

FIELD : COMPUTER SCIENCE

OPTION : IDO

**Dissertation submitted in partial fulfillment of the requirements for
The degree of MASTER**

By: CHAANBI BOUTHAINA

Subject

**forecasting tools for optimizing the supply
chain**

Publicly defended before the jury composed of :

.....

University of M'sila

Chair

Mr. Allaoua Hemmak

University of M'sila

Supervisor

.....

University of M'sila

Examiner

College year : 2019 /2020

Table of Contents

Table of Contents	1
Liste of Figures	4
List Of Abbreviation.....	6
General Introduction	7
Thesis Context.....	7
Chapter 1. General <i>information</i> on supply chains.....	8
1 <i>Introduction</i>	8
2 The Supply Chain	8
3 Supply chain management :	10
3.1. Definition :	10
3.2. The objectives of supply chain management:	11
3.3 Management of requests and stocks:	12
4 . Forecasting in the supply chain:	13
4.1 Demand Forecasting:.....	13
4.1.1 Defintion:	13
4.2 Demand forecasting and the supply chain:	14
4.3 Importance of forecasting :.....	15
5 . Methods of Demand forecasting:	16
6. Qualitative prediction methods:.....	16
6 .1 . The Delphi method:	17
6 .2. executive opinion.....	18
6 . 3 . Sales fores composite	18
7. Quantitative forecasting methods.....	19
7 . 1 Time series models(time series)	19
7 . 2. The components of time series	20
a) Seasonality:	20
b) Trend:	21
7 . 3 Time-Series Transformations	23
7 . 4 . Statistical Methods	25
8 . Machine Learning Medels:.....	28
8.1. Support vector machines (SVM):.....	29
8.2. K-Nearest Neighbor regression (KNN):.....	29
8.3 Decision Tree:	30

8.4	Grouping of k-means:	30
8.5	MLP neural networks:.....	31
8.6	LSTM neural networks:	31
9.1	Hybrid models:.....	33
9.1.1	Box-Jenkins and Artificial Neural Networks:	33
9.2	causal patterns:	33
9.2.1	Simple regression model :	33
9.2.2	multiple regression model:.....	34
10	Evaluation and validation of forecasting models:	34
10.1	Procedure for evaluating training and test sets:	34
10.2	Cross-Validation Time-Series Procedure:	35
10.3	Forecast precision measurements:.....	36
10.3.1	Scale dependent errors:.....	36
10.3.2	Percentage error :.....	36
10.3.3	Errors At Scale :	37
	Conclusion :	38
	Chapter 2: State of Art	39
1	.Introduction :	39
2	Methodology:	39
2.1	Artificial neural networks :	39
2.2	. Rough set theory :	40
2.3	Machine learning:.....	40
2.4	Expert systems:	41
2.5	Genetic algorithms:	42
2.6	Fuzzy logic :.....	43
2.7	Agent-based systems:.....	44
	Chapter 3 . Data Analysis	46
1	Introduction:.....	46
1	.Existing Model and Problem Statement:	46
2	.Functional and Technical Specifications:	47
3	.The Challenge of Choosing the Forecasting Model:.....	47
4	. Preliminary Data Analysis Approach:	48
4.1	Data Description:	48
4.2	Data Exploration:.....	50
4.3	Cross-Correlation Analysis:	51
4.4	Demand Categorization for Inventory Items:	54

General Conclusion :.....	58
Abstract.....	59
Resumé.....	59
ملخص.....	60

Liste of Figures

Figure 1 Supply chain flows	9
Figure 2 Modèle SCOR (supply chain council)	9
Figure 3 house of Stadtler et Kilger (14) SCM	11
Figure 4 Types of requests	12
Figure 5 Demand forecasts and their roles in decision-making in the supply chain	14
Figure 6 Place of forecasting in the global supply chain.....	15
Figure 7 classification of forecasting methods	16
Figure 8 the Delphi method process	17
Figure 9 example of a time series in blue and forecast in orange	19
Figure 10 example of a seasonal time series.....	20
Figure 11 additive and multiplicative seasonality	21
Figure 12 example of the trend in a time series	21
Figure 13 A business cycle in four phases	21
Figure 14 Four examples of time series showing different models	22
Figure 15 multiplicative and additive models	23
Figure 16 multiplicative and additive models	23
Figure 17 An RNN where W is a weight matrix and U is a matrix Consisting of what is called bias and ϕ is an activation function, making the network nonlinear	32
Figure 18 LSTM cell.....	32
Figure 19 Illustration of the division of the training / test dataset.....	34
Figure 20 Illustration of the cross-validation procedure for different forecast horizons-1-.....	35
Figure 21 Illustration of the cross-validation procedure for different forecast horizons-2-.....	35
Figure 22 Preliminary data analysis process	48
Figure 23 Daily sales of the three retailers comprising all the product variants sold by each one of them	52
Figure 24 Cross Correlation analysis of some product variants from inventory S-1088	53
Figure 25 Cross Correlation analysis of some product variants from inventory S-1224	53
Figure 26 Cross Correlation analysis of some product variants from inventory S-767	54
Figure 27 Syntetos Boylan classification decision tree.....	54
Figure 28 Example of a product variant from S-1088 and exhibiting intermittent demand	55
Figure 29 Example of a product variant from S-1224 and exhibiting lumpy demand	56
Figure 30 Example of a product variant from S-767 and exhibiting intermittent demand	56

List Of Tables

<i>Tableau 1 Definitions of supply chain management</i>	<i>10</i>
<i>Tableau 2 Functional specifications</i>	<i>47</i>
<i>Tableau 3 Technical specifications</i>	<i>47</i>
<i>Tableau 4 Structure of Stock&Buy historical sales data.....</i>	<i>49</i>
<i>Tableau 5 Structure of Stock&Buy inventory data.....</i>	<i>50</i>
<i>Tableau 6 Stationarity test using Augmented Dickey-Fuller Test on retailers sales data</i>	<i>51</i>
<i>Tableau 7 Demand categorization for the product variants across stores S-1088, S-1224 and S-767.</i>	<i>55</i>

List Of Abbreviation

ANN	Artificial Neural Networks
ARIMA	Auto Regressive Integrated Moving Average
BNN	Bayesian Neural Network
CART	Classification and Regression Trees
FBPN	Fuzzy Back-Propagation Network
GP	Gaussian Process
GRNN	General Regression Neural Networks
HES	Hyperbolic-Exponential Smoothing
KNN	K-Nearest Neighbors
LES	Linear-Exponential Smoothing
LSTM	Long Short Memory Network
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MASE	Mean Absolute Scaled Error
ML	Machine Learning
MLP	Multi Layer Perceptron
NN	Neural Network
PCA	Principal Component Analysis
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
SB	Syntetos-Boylan Adjusted Croston Method
SBJ	Shale-Boylan-Johnston Adjusted Croston Method
SES	Simple Exponential Smoothing
SKU	Stock Keeping Unit
SVR	Support Vector Regression

General Introduction

Thesis Context

With the availability of data and the increasing capabilities of data processing tools, many businesses are leveraging historical sales and demand data to implement intelligent inventory management systems. Demand forecasting is the process of estimating the consumption of products or services for future time periods. It plays an important role in the field of inventory management and Supply Chain, since it enables production and supply planning and therefore can reduce delivery times and optimize Supply Chain decisions. The results of demand forecasting depend very much on the relevance, quality and accuracy of the forecasts made, which in turn depend on the forecasting model being used.

Demand forecasting is often based on time series data, which are the most common type of historical data. There are several methods that can be used for time series forecasting, the first of which appeared in the 1950s. Since then, it has become a highly sought-after research topic and other methods have emerged which were for the most part statistical methods. The latter provided highly accurate forecasting results and were very useful in practice.

With the emergence of Machine Learning (ML), ML methods have been proposed as alternatives for statistical methods, and have generated such a buzz in the forecasting field. Hence, several recent research works proposed new time-series forecasting methodologies based on Machine Learning.

Within this context, Stock&Buy, a growing online retail platform, wants to build a demand-forecasting system in order to allow for effective inventory control. This system shall use robust forecasting techniques in the field of demand forecasting to handle the diversity of products and their underlying demand patterns, but also to estimate demand for new products for which there is no historical demand data.

Chapter 1. General information on supply chains

1 Introduction

Since the second half of the 20th century, companies have been faced with a more and more competition. This competition can be explained, among other things, by the development of international trade and the market economy. In this context, it is essential for companies to rely on a chain efficient supply necessary for their performance. Among these chains the Supply Chain Management (SCM) is one of the new concepts in the corporate sector of Bangladesh which was practiced from late 90s. Initially the Multinational Companies (MNC) incorporated Supply Chain Management in their structures and later on other privates and local conglomerates embraced the concepts. Since beginning purchase and materials management were the main functions of SCM, but later on SCM took the integrated shape i.e. consists of sourcing, materials management, manufacturing support, and distribution management. Considering the competitive market scenario, SCM becomes the prime functioning area among the companies. SCM deals with direct, indirect, and services from the origin (as input materials) to end customers as final products. A Supply Chain is a network among the supply chain partners such as suppliers, manufacturers, distributors, retailers, transporters etc. who shares information, deliver goods, ensure services and perform other intermediate activities to meet customer demand. Forecasting is far most beginning activities of SCM which initiates the all other actions of SCM. However forecasting plays an important domain in inside as well as outside of the company. Forecasting is the key driving factor in planning and making decisions in SCM as well as enterprise level. Companies are, truly perform professional way, highly depends on true numerical value of forecasting to take major decisions such as capacity building, resource allocation, expansion and forward or backward integration etc.

2 The Supply Chain

The term supply chain was first used in the newspaper "The Independence" in 1905. However, it remains difficult to define the terms "supply chain".

We find in the literature a multitude of definitions which are generally dependent on the problem treated. One of the first generic definitions is proposed by Beamon [1] which defines the supply chain as: "a process in which a variable number of entities (ie, suppliers, manufacturers, distributors, retailers) work together to : acquire raw materials, convert these raw materials into specific finished products and deliver these finished products to retailers. This chain is traditionally characterized by a downstream flow of materials and an upstream flow of information."

This definition presents the supply chain in three aspects that Comelli [2] and Lemoine [3] use to define it: entities, flows and activities. For Christopher [4], entities are interconnected organizations that produce value through different activities. For Beamon, these are suppliers, manufacturers, distributors and retailers. Entities are the physical nodes of the network that perform various tasks such as storage, manufacturing, or distribution in order to meet an objective which is generally customer satisfaction at the lowest cost.

The second aspect is that of flows. [5] propose four in their definition of the supply chain:

- Product and service flows. This flow is generally a downstream flow (from suppliers to the customer). The purpose of this flow is, in most cases, the satisfaction of the customer.

- The financial flow. It is generally reversed to the flow of products and services. He constitutes the compensation of the various entities of the supply chain to wards those that are upstream. This flow is generally optimized locally by each of the independent supply chain organizations although in the case of an intra-organizational supply chain, this flow can be globally optimized.
- The flow of information. It is an essential flow for the proper functioning of the chain supply. According to Comelli and Lemoine it allows coordination between physical and financial flows. It also contains all the necessary data managing the supply chain and carrying out various activities.

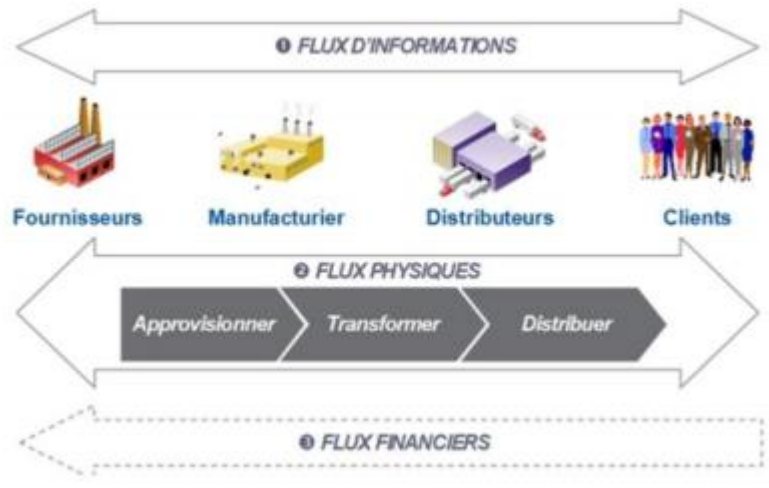


Figure 1 Supply chain flows

The last aspect represents the different activities carried out in the supply chain. They can be varied and are grouped into four categories by the “supply chain council” 1 in the SCOR model. Figure 2 presents this modeling



Figure 2 Modèle SCOR (supply chain council)

The different activities offered are:

- "Source", the procurement process that provides goods and services.
- “Make”, the manufacturing process which allows the transformation of raw material into Finished products.

- “Deliver”, the process of delivering goods and services to the different entities of the chain.
- “Return”, the process of returning items for various reasons (after-sales service).

3 Supply chain management :

Supply Chain Management (SCM) - is a major subject that fits into a wide variety of issues. This is why the notion apparently introduced by Oliver and Webber [6] has been of great interest from the late 1990s to the present day. Indeed, at the end of the previous century, it was difficult to find work in manufacturing, distribution, marketing, customer management or even transportation without seeing the subject of supply chain management associated with it [7; 8]. There are several reasons for the popularity of the concept. One example is globalization, which has made supply chain management very difficult. Also, the need companies to be more and more competitive on different aspects such as the cost of management or delivery times . Thus, a race for performance has been launched between large companies and has aroused general interest in supply chain management because if good chain management doesn't necessarily allow a business to thrive, mismanagement will sink it.

3.1. Definition :

As with supply chains, there are multiple definitions for their management. 5 offer a very interesting study of these definitions. We have retained some definitions presented in Table 1.1.

[9]	SCM requires traditionally separate materials functions to report to an executive responsible for coordinating the entire materials process, and also requires joint relationships with suppliers across multiple tiers. SCM is a concept, “whose primary objective is to integrate and manage the sourcing, flow, and control of materials using a total systems perspective across multiple functions and multiple tiers of suppliers.”
[10]	Supply chain strategy includes: “... two or more firms in a supply chain entering into a long-term agreement; ... the development of trust and commitment to the relationship; ... the integration of logistics activities involving the sharing of demand and sales data; ... the potential for a shift in the locus of control of the logistics process.”
[11]	“The objective of managing the supply chain is to synchronize the requirements of the customer with the flow of materials from suppliers in order to effect a balance between what are often seen as conflicting goals of high customer service, low inventory management, and low unit cost.”
[12]	Differences between supply chain management and classical materials and manufacturing control: “1) The supply chain is viewed as a single process. Responsibility for the various segments in the chain is not fragmented and relegated to functional areas such as manufacturing, purchasing, distribution, and sales. 2) Supply chain management calls for, and in the end depends on, strategic decision making. “Supply” is a shared objective of practically every function in the chain and is of particular strategic significance because of its impact on overall costs and market share. 3) Supply chain management calls for a different perspective on inventories which are used as a balancing mechanism of last, not first, resort. 4) A new approach to systems is required—integration rather than interfacing.”
[13]	Supply chain management is “... an integrative philosophy to manage the total flow of a distribution channel from supplier to the ultimate user.”

Tableau 1 Definitions of supply chain management

Stadler and Kelger [14] discuss the main points that can be found in most definitions with their "home" of supply chain management (Figure 3). This house represents the three main points of supply chain management. The first point, the roof, represents the targets. It can vary, but in general, the overall competitiveness of the chain and customer satisfaction.

The other two are the columns on which the roof rests. The first is integration, which represents the organization and overall structure of the supply chain. The second is coordination between entities in the supply chain. We mention here, for example, the synchronization between orders and delivery. We will now detail the main points of supply chain management

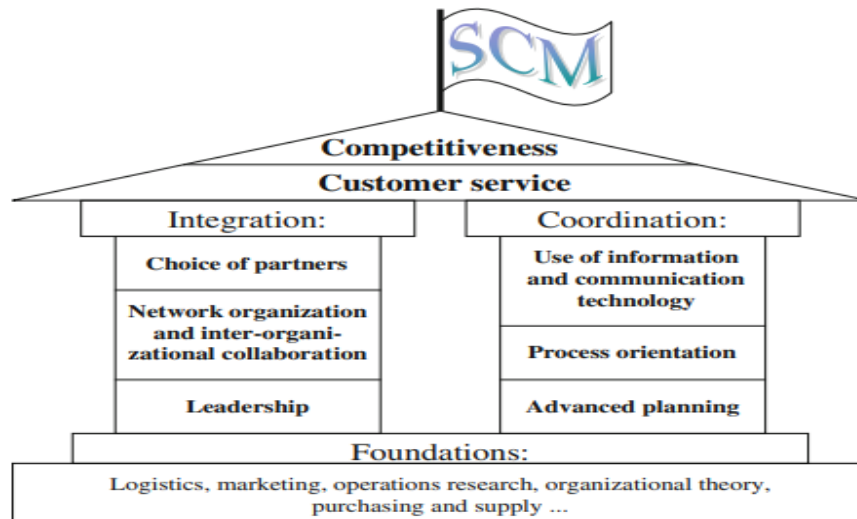


Figure 3 house of Stadler et Kilger (14) SCM

3.2. The objectives of supply chain management:

The main purpose of supply chain management is to improve the competitiveness of the supply chain as a whole. In order to achieve this objective, it is necessary to focus on different more concrete objectives. The most important of these is generally customer satisfaction with the level of service [15; 16]. The level of service is defined as the percentage of requests satisfied on time [17]. A request is considered to be satisfied in time if it is satisfied within the time specified by the contract between the supplier and the customer - in English Service Level Agreement (SLA). Obviously, all businesses can satisfy a maximum level of service. Just invest in efficient transportation and enough inventory to meet all demands imaginable. However, such management would lead companies to bankruptcy because the expenses generated would be clearly greater than the profits.

It is therefore necessary to define objectives transversal to that of satisfaction with the level of service. These goals can be to limit the amount of stock in the supply chain or to limit the number of trips made. In the case of a supply chain with a repair loop, we can try to limit the number of purchases of items to avoid introducing new ones into the chain or even try to spend as little as possible on repairs. . Other goals may be more chain dependent.

supply. So we can for example mention ecological supply chains (green supply chain), one of the main objectives of which is to limit the energy used during all the processes .

Defining objectives is an essential step in building a supply chain if it is to be efficient.

3.3 Management of requests and stocks:

An essential element to establish the chain management method supply is demand. The request corresponds to the quantity of products ordered or planned by the customer. The nature of the demand is highly dependent on the nature of the product linked to it [17]. Products can be divided into two main categories: functional products and innovative products (see Figure 4).

Functional products are the basic products that can be found in any mall. These products generally have a long lifespan and their demand is stable and predictable. On the contrary, innovative products are products . which will quickly go out of fashion and whose demand is largely dependent not only on the quality of the product but also on subjective factors such as fashion effects. Also, it is very difficult to predict the demand for this kind of products.

We will see in section 1.3.4 that this uncertainty is one of the main obstacles to effective management of the supply chain.

Closely linked to demand management, the other axis of supply chain management is inventory management. According to Waters [18], the latter is essential, and should allow a subtle balance to be achieved between production and orders. Indeed, the rates of consumption and production are not always similar, and efficient management of stocks is then necessary to avoid excesses which can be too low a volume of stock to meet the demands or on the contrary a surplus of stock. These risks depend on the type of requests to be answered. In fact, the inventory management techniques that will be used to meet the demands will be significantly different.

Functional	Innovative
Low demand uncertainties	High demand uncertainties
More predictable demand	Difficult to forecast
Stable demand	Variable demand
Long product life	Short selling season
Low inventory cost	high inventory cost
Low profit margins	High profit margins
Low product variety	High product variety
Higher volume per SKU	Low volumes per SKU
Low stockout cost	High stockout cost
Low obsolescence	High obsolescence

Figure 4 Types of requests

Here we introduce the concept of inventory management policy. An inventory management policy results in the choice of a sequence of decisions (purchase, sale, supply, etc.) whose aim is to minimize the cost of the supply chain while maximizing profits.

We divide management policies into two categories: proactive management policies and reactive management policies. Proactive policies produce sequences of decisions made in order to anticipate future demands. On the contrary, the decisions made by reactive policies follow a triggering event (demand, stock level below a threshold, etc.).

A proactive policy can allow a better level of service and / or a less quantity of stock mobilized than a reactive policy. However, it is necessary to have a reliable forecast of future demand. Without it, putting in place a proactive policy is ineffective and will either lead to numerous stock-outs or an accumulation of safety stocks. In the event that information on future requests is unreliable or non-existent, the implementation of a reactive policy will be preferred.

4 . Forecasting in the supply chain:

Forecasting in general is considered an important problem in the management of supply chains, but with research in this area experts believe that it is possible, using appropriate and well-studied methods, to establish good forecasts with a high level of accuracy.

Forecasting is a common statistical task in businesses. It helps make informed decisions about production, transportation and personnel planning, and provides a guide for long-term strategic planning.

Forecasting involves predicting the future as accurately as possible, taking into account all available information, including historical data and knowledge of any future event that may affect the forecast. [19]

Forecasting involves determining what will happen in the future by analyzing what has happened in the past and what is happening now. It is a planning tool that helps business people cope with uncertainty about what may or may not happen. Forecasts are based on past and current data and trend analysis.

According to John G. Glover, "Forecasting is the research procedure aimed at discovering the economic, social and financial influences that govern the activity of companies, in order to predict or estimate current or future trends or forces that may affect company policies or future finances. production and marketing operations. "

The essence of all of the above definitions is that forecasting is a technique for analyzing the economic, social and financial forces that influence business activities in order to predict future events based on past information and present.

Forecasting should be an integral part of management decision-making activities, as it can play an important role in many areas of a business. Modern organizations need short, medium and long term forecasts, depending on the application.

- Short-term forecasts: are necessary for personnel planning, production and transport. As part of the planning process, demand forecasts are also often required.
- Medium-term forecasts: are needed to determine future resource requirements, to buy raw materials, hire staff or buy equipment
- Long-term forecasts: are used in strategic planning. These decisions must take into account market opportunities, environmental factors and internal resources.

4.1 Demand Forecasting:

Demand forecasting plays a very important role in the chain of supply. It is an essential key for planning and setting the objectives of each business. Forecasting demand [20] also provides crucial information for better management of stocks and optimal adaptation of supply to demand. Indeed, more reliable and more precise forecasts favor visibility on demand and thus constitute a guarantor of an efficient management of the supply chain (SCM), allowing to reduce the impact of the inevitable fluctuation of the market and the uncertainty about demand

4.1.1 Defintion:

Demand forecasting is the activity of estimating the quantity of a product or service that consumers will buy over a given period. Forecasting demand involves techniques that include both informal methods, such as informed guesses, and quantitative methods, such as the use of historical sales data or current data from test markets.

Demand forecasting simply refers to the process of forecasting future demand for a company's product. This therefore refers to future customer demand and therefore the products are produced by a company

One finds in the literature a multitude of definitions, one of the first definitions which is proposed by Evan J. Douglas, "The estimation of demand (forecast) can be defined as a process of finding values for demand in future periods. Another definition by Cundiff and still which defines demand forecast as an estimate of sales over a specified future period based on the proposed marketing plan and a set of uncontrollable and competitive particular forces. And in another definition Forecasting consists in predicting future events. Demand forecasting is used to predict future demand for the company's products and services, in order to be able to satisfy customer desires over a specified period using historical data, usually based on demand data. Having a good forecast is crucial for efficient service and for manufacturing operations [21]

4.2 Demand forecasting and the supply chain:

Demand forecasting provides very useful information for all activities in the supply chain (Fig. 5).

Purchasing, marketing, manufacturing, management, budget and logistics decisions are all dependent on the forecast [22], so demand estimation is required in the short, medium or long term. At the operational level, sales analysis is used to make decisions on supply, production, inventory management and organization of vehicle routing activities, etc. In the medium term, demand forecasting is used to plan logistics operations and determine the resources required for production, distribution and transportation. In the long term, the estimate of demand is aimed at making decisions on a strategic horizon, such as the development of new products, the outsourcing or not of a logistics activity, etc. [23]

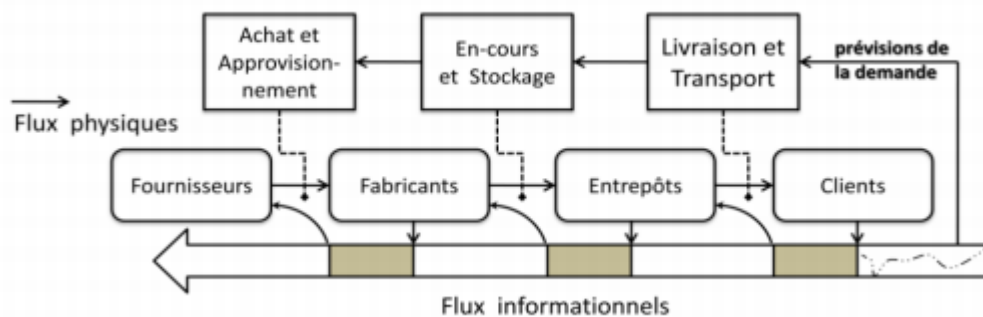


Figure 5 Demand forecasts and their roles in decision-making in the supply chain

To a large extent, supply chain planning begins with forecasting. Matching supply and demand is an important goal for most businesses and is at the heart of operational planning.

It is also very important. Since most production systems cannot respond to consumer demands instantly, an estimate or forecast of future demand is necessary for operational plans to be effective and efficient.

Facilities, processes and work capacity are all defined based on forecasts of future demand.

Capacity planning and facility decisions would be based primarily on the long term, aggregate forecast. However, forecasting is also necessary to plan the appropriate inventory levels.

The forecasts concern most of the functional areas of the company (Fig 6) and constitute the starting point for resource management - Allocation decisions -.

For example, manufacturing must plan production on a daily basis to meet customer orders, while purchasing must know how to align suppliers, deliveries with production schedules. Finance must include forecasting so that appropriate levels of investment can be made in facilities, equipment, inventory and so budgets can be constructed to better manage the business. The marketing function needs to know how to allocate resources for various product groups and marketing campaigns. The forecasts also determine the manpower requirements of the business; the human resources function can make the right hiring and training decisions when demand is expected to increase.

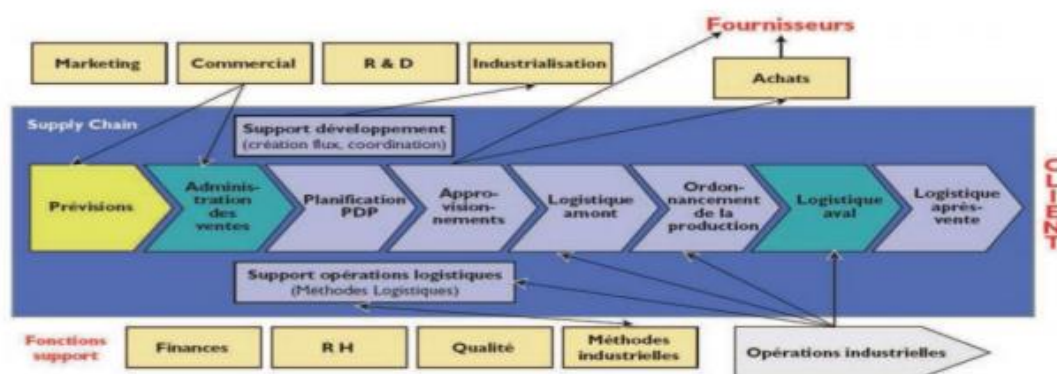


Figure 6 Place of forecasting in the global supply chain

4.3 Importance of forecasting :

As we saw previously the demand forecast and the first phase in the supply chain planning process and it is also the first phase in the supply chain in general, so it plays a very important role: [24] [25][26] [27]

- Forecasting demand helps companies adjust their production schedules so that they can respond effectively to demand.
- It will also help businesses keep inventories low, which could help reduce the costs associated with inventory and the costs of finished products as a result.
- While more production would involve more capital employed, forecasting demand could help the business invest the appropriate amount
- Demand forecasting can be used in pricing decisions. For example, if the forecast indicates a decreasing share of market share, then the company should lower the price to gain more demand.
- Demand forecasting can also be used to assess future capacity needs or to decide whether to enter a new market.

Overall, it is obvious that forecasting demand is essential for a business in terms of production, stocks, capital, prices, etc.

5 . Methods of Demand forecasting:

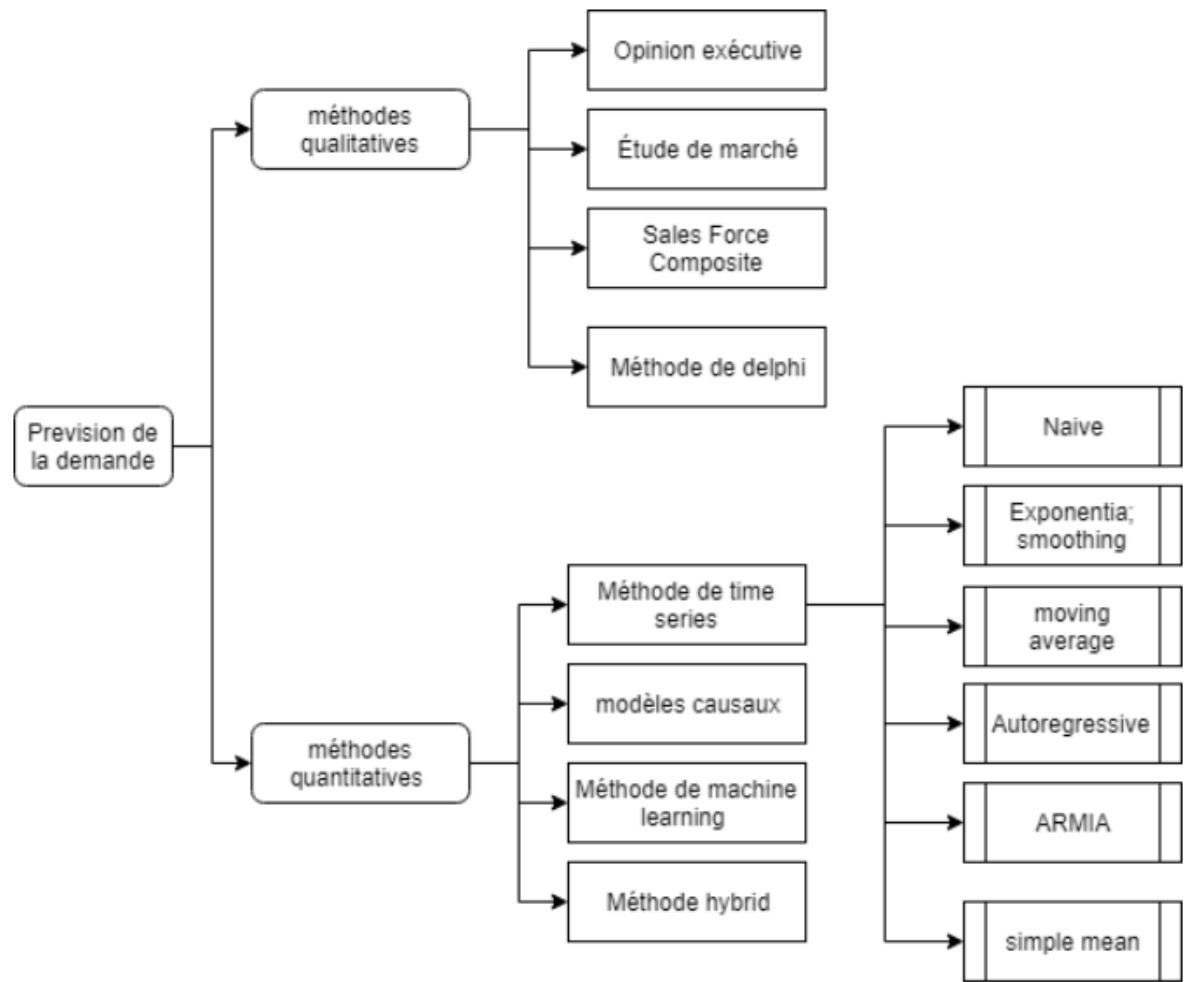


Figure 7 classification of forecasting methods

6. Qualitative prediction methods:

Qualitative forecasting is an estimation method that requires expert judgment rather than numerical analysis. This type of forecast draws on the knowledge of highly experienced employees and consultants to provide insight into future results. This approach is significantly different from quantitative forecasting, where historical data is compiled and analyzed to discern future trends.

Qualitative forecasting is particularly useful in situations where it is suspected that future results are significantly different from the results of previous periods and therefore cannot be predicted by quantitative means. For example, the historical sales trend may indicate that sales will increase further in the next year, which should normally be measured using a trend line analysis; However, industry expert says there will be a shortage of materials from a key supplier, which will force sales down.

Another situation in which qualitative forecasting can be useful is to assimilate large amounts of closely targeted local data in order to discern the trends that a more quantitative analysis might not allow to detect evidence. For example, a construction company needs to know what style of house to build in a given area and relies on a local expert to find out that the area in question is being

abandoned by families younger and replaced by an older, retired society. group. As a result, the builder is building smaller single-storey homes with fewer bedrooms.

This approach also works well when an action plan needs to be derived from inadequate data. In this case, a qualitative analysis will seek to link disparate data in order to build a broader view, sometimes incorporating an intuition for build this view.

Another situation in which qualitative forecasting can add value is when management changes trends derived from history based on expert opinions. In this case, quantitative methods are used to create a preliminary forecast, which is then adjusted by a qualitative examination. In theory, the result should be a forecast derived from the better of the two methods.

The results obtained by the qualitative forecasts can be biased for the following reasons:

- Recency. Experts may tend to focus more on recent historical events by extrapolating future trends.
- Personal worldview. Experts may have built their own vision of how the industry works and tend to reject new influences on this market.

In the following, we briefly describe some of the most popular qualitative forecasting techniques that have led to the acceptance of judgmental forecasting as a forecasting approach.

6.1 . The Delphi method:

The Delphi method is a framework for forecasting process based on the results of several series of questionnaires sent to a panel of experts. Several series of questionnaires are sent to group of experts and anonymous responses are aggregated and shared with the group after each round. Experts are allowed to adjust their responses in subsequent rounds, depending

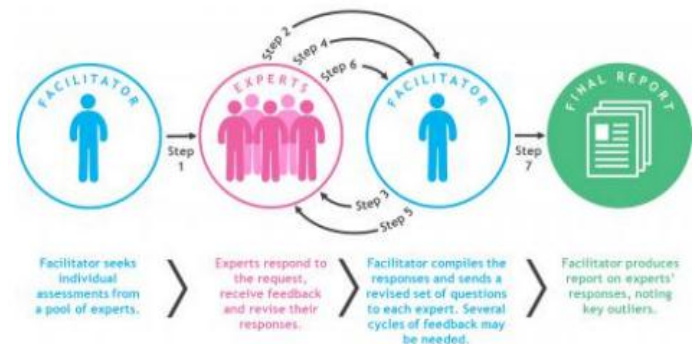


Figure 8 the Delphi method process

On how they interpret the "group response" provided to them. Since several series of questions are asked and the panel is informed of what the group as a whole thinks, the Delphi method seeks to obtain the correct answer by consensus.

The Delphi method was originally designed in the 1950s by Olaf Helmer and Norman Dalkey of the Rand Corporation. The name refers to the Oracle of Delphi, a priestess of a temple of Apollo in ancient Greece known for her prophecies. The Delphi method allows experts to reach a mutual agreement by carrying out a series of circulating questionnaires and by diffusing the corresponding comments to deepen the discussion at each following cycle. The experts' responses change as the series ends on the basis of information provided by other experts participating in the analysis. [28]

How this method works:

The Delphi method generally involves the following steps: [28]

1. A panel of experts is assembled.

2. Forecasting tasks / challenges are defined and distributed to the experts.
3. The experts return the initial forecasts and the justifications. These are compiled and summarized to provide feedback.
4. Comments are provided to the experts, who are now revising their forecasts in the light of these comments. This step can be repeated until a satisfactory level of consensus is reached.
5. Final forecasts are constructed by aggregating expert forecasts.

Each step of the Delphi method has its own challenges. In the following, we provide some suggestions and discussions on each of them.

6 .2. executive opinion

This approach implies that the staff at the top of the management structure generates global forecasts. These forecasts are usually generated during a group meeting, during which leaders provide information from their own area of business. The presence of executives from different functional areas of the company promotes a wide variety of skills and knowledge within the group.

This process has all the advantages and disadvantages of a group meeting that we have already discussed. In this context, it is important to justify and document the forecasting process. This means that managers must be held accountable in order to reduce the biases generated by setting up group meetings.

It may also be possible to apply variations to a Delphi approach in this context; for example, the estimation-conversation-estimation process described above.

6 .3 . Sales fores composite

In this approach, forecasts for each point of sale / branch / store of a company are generated by the sellers and then aggregated. This usually means that sales managers anticipate demand from the store that they manage. Sellers are usually as close as possible to the interactions between customers and products and often develop an intuition about customers' buying intentions. They bring this valuable experience and expertise to the forecast.

However, the fact that sellers generate forecasts violates the fundamental principle of the separation of forecasters and users, which can create biases in many directions. It's common for a seller's performance are evaluated in relation to sales forecasts or previously established expectations. In this case, the seller acting as a forecaster can introduce a selfish bias by generating low forecasts. On the other hand, one can imagine an enthusiastic seller, full of optimism and generating high forecasts.

In addition, a successful salesperson is not necessarily a successful or knowledgeable forecaster. A significant proportion of salespeople will have no formal or limited training in forecasting. Finally, sellers will feel customer dissatisfaction if, for example, the product runs out or is not brought into their store. Such interactions will interfere with their judgment.

7. Quantitative forecasting methods

The second category of forecasting models is quantitative forecasting which, unlike the qualitative model, relies on mathematical (statistical) models to make projections for the future.

It is a statistical technique for predicting the future, which uses numerical measurements and past effects to predict future events. These techniques are based on mathematical models and in nature are mainly objective. They are highly dependent on mathematical calculations.

There are two types of quantitative forecasting methods which are listed below:

1. Time series models (time series).
2. Associative model : are also called occasional models. The model assumes that the variable being forecast is associated with other variables. The forecasts are based on these associations. Linear regression is one of the simplest forms of an associative forecasting model. This regression line predicts the variable dependent on the selected value of the independent variable.

Quantitative forecasting methods are very easy to predict based on the underlying information. The data can be used to automatically forecast without many complications. Anyone can easily predict based on the data available.

One of the main drawbacks of this method is its dependence on data. The entire forecast depends on the data from the underlying model. An error in the available data can lead to erroneous forecasts. These methods can also only be used if the appropriate data are available. This method also does not allow to evaluate the effect of the modifications of the other variables involved.

7 . 1 Time series models(time series)

These models examine past data models and predict the future based on the underlying models obtained from this data. There are many types of time series models, such as the simple and weighted moving average ,seasonal indices, trend projections, simple, exponential smoothing.

And to better understand these models you must first understand what is time series or even time series. And that's what we're going to use for this project.

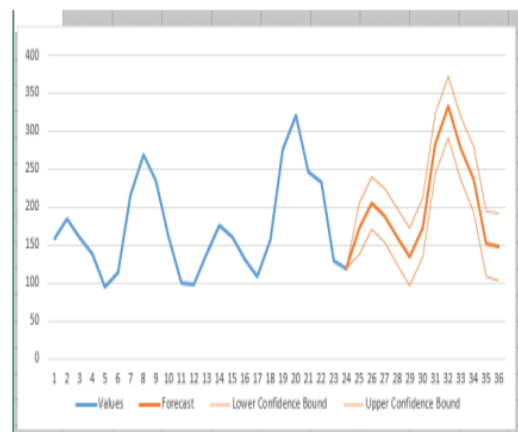


Figure 9 example of a time series in blue and forecast in orange

A timeline is a series of number data points in sequential order. By investing, a time series follows the movement of the selected data points, such as the price of the security, over a specific period, data points They are recorded at regular intervals. There is no minimum or maximum inclusion, which makes it possible to collect data in order to provide the information required by the investor or the analyst examining the activity.

Mathematically, we can formulate a time series as follows: We index the time periods as $1, 2, \dots, T$ so we would have $(y_1, y_2, \dots, y_T)$ or y_t indicating the observation at time t .

The time series is simply a series of notes taken sequentially over time. [28]

Time series analysis involves developing models that capture or describe an observed time series in order to understand the underlying causes. This field of study seeks "Why?" Behind a dataset Time series.

This often includes assumptions about the form of the data and the division of time series into components of the constitution.

The quality of the descriptive model depends on the quality with which it describes all the available data and on the interpretation it provides to better inform the problem area.

"The main objective of time series analysis is to develop mathematical models that provide reasonable descriptions of the sample data."

[29]

Time series prediction uses information about historical values and associated models to predict future activity. This is mainly related to trend analysis, analysis of periodic fluctuations and problems Seasonality. As with all prediction methods, success does not guarantee.

7 . 2. The components of time series

A time series in general is assumed to be affected by four main components, which can be separated from the observed data. These components are:

Trend components, cyclical, seasonal and irregular. A brief description of these four components is given here.

a) Seasonality:

This component exists when a series exposes regularly fluctuations according to the season (for example, each month / quarter / year). Seasonality is always of a fixed and known period. For example, ice cream sales tend to peak during the summer season (summer) and then decrease thereafter.

Seasonality plays a very important role if it is present, because it has a very important impact on the choice of forecasting method to use because there are methods that work well with this aspect of seasonality and there are some who works more therefore an influence on the accuracy of the forecast.

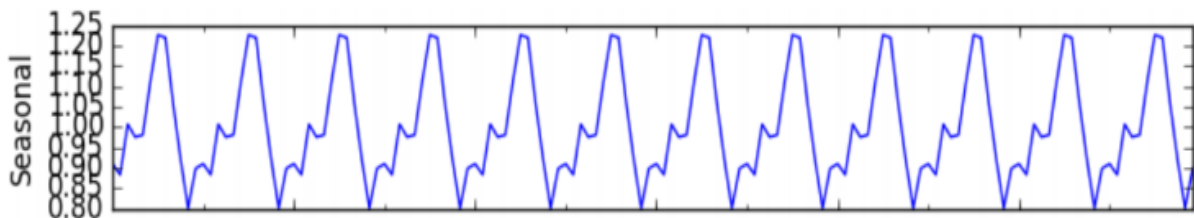


Figure 10 example of a seasonal time series

b) Trend:

The general tendency of a time series to increase, decrease or stagnate over a long period of time is called a secular trend or simply a trend. So you could say that the trend is a long-term movement in a time series. For example, series relating to population growth, the number of houses in a city, etc. show an upward trend, while a downward trend can be observed in the series on mortality rates, epidemics, etc.

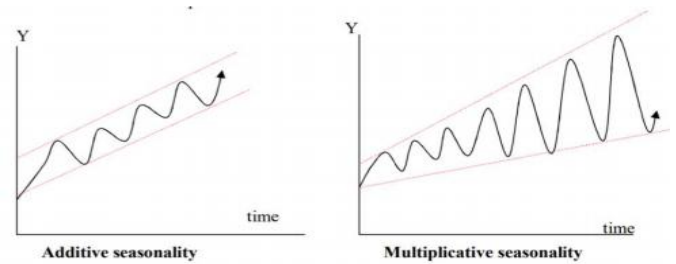


Figure 11 additive and multiplicative seasonality

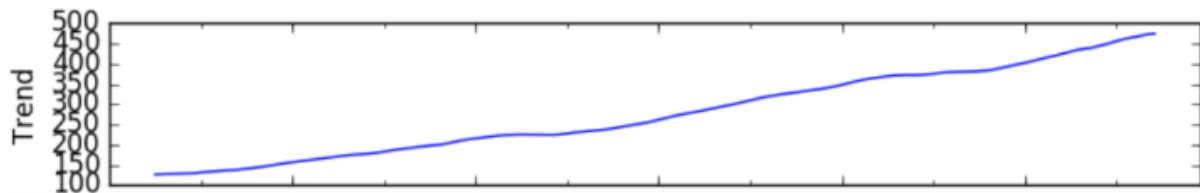


Figure 12 example of the trend in a time series

C) cyclical:

The cyclical variation in a time series describes the medium-term changes in the series, caused by circumstances, which are repeated in cycles. The duration of a cycle spans a longer period of time, usually two years or more. Most economic and financial time series show some sort of cyclical variation. For example, an economic cycle has four phases:

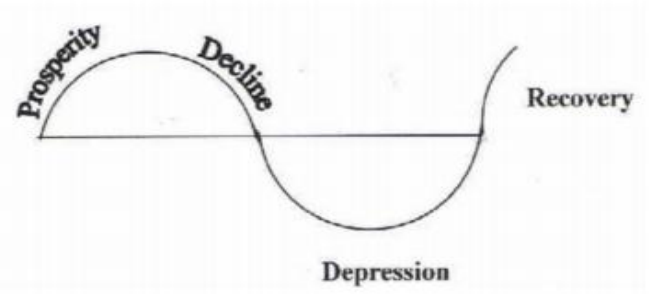


Figure 13 A business cycle in four phases

i) prosperity, ii) decline, iii) depression and iv) recovery.

Schematically, a typical business cycle can be presented as follows:

D) Irregular:

Irregular or random variations in a time series are caused by unpredictable influences, which are not regular and which are not repeated in a particular model either. These variations are caused by incidences such as war, strike, earthquake, flood, revolution, etc. There is no defined statistical technique for measuring random fluctuations in a time series.

This diagram shows the different components of a time series:

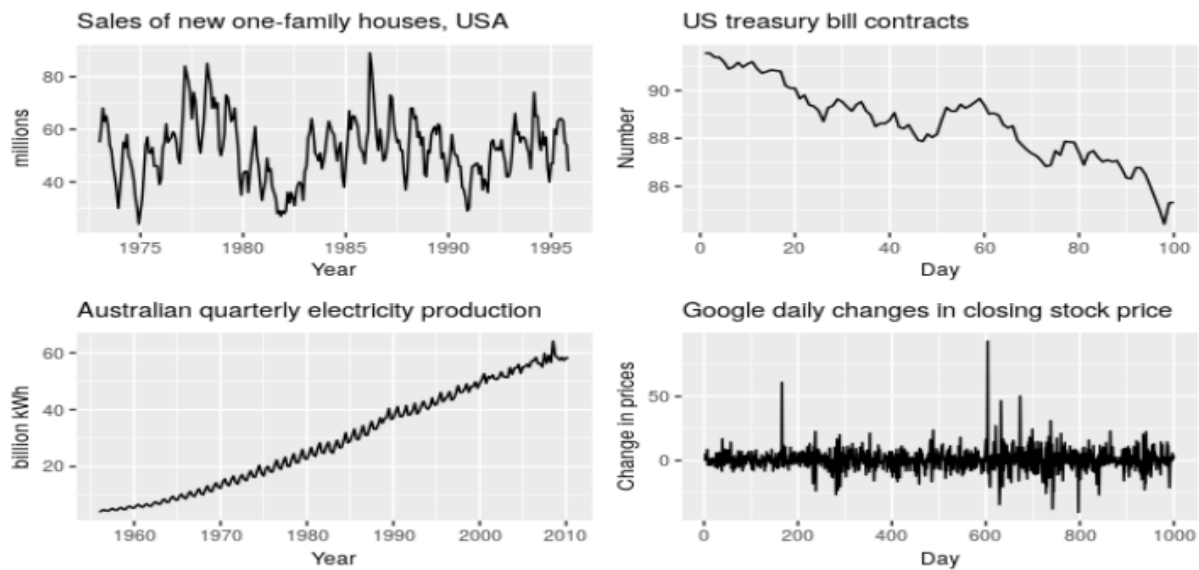


Figure 14 Four examples of time series showing different models

Diagram (a): The monthly home sales (top left) show a strong seasonality during each year, as well as a significant cyclical behavior with a period of around 6 to 10 years. There is no apparent trend in the data for this period.

Figure (b): US Treasury contracts (top right) show Chicago market results for 100 consecutive trading days in 1981. Here there is no seasonality, but a downward trend obvious. Maybe if we had a much longer run, we would see that this downtrend is actually part of a long cycle, but if we only look at it over 100 days, it seems to be a trend.

Figure (c): Australian quarterly electricity production (bottom left) shows a strong upward trend, with strong seasonality. There is no evidence of cyclical behavior here.

Diagram (d): The daily change in the closing price of Google (bottom right) has no trend, seasonality or cyclical behavior. There are random fluctuations which do not seem very predictable, and no strong trend would allow to develop a forecasting model.

Given the effects of these four components, two different types of models are generally used for a time series. Multiplicative and additive models.

$$\text{Multiplicative model: } Y(t) = T(t) \times S(t) \times C(t) \times I(t)$$

$$\text{Additive model: } Y(t) = T(t) + S(t) + C(t) + I(t)$$

With $Y(t)$ observation (time series) and $T(t)$, $S(t)$, $C(t)$, $I(t)$ are respectively the trend, seasonal, cyclic and irregular variation over time.

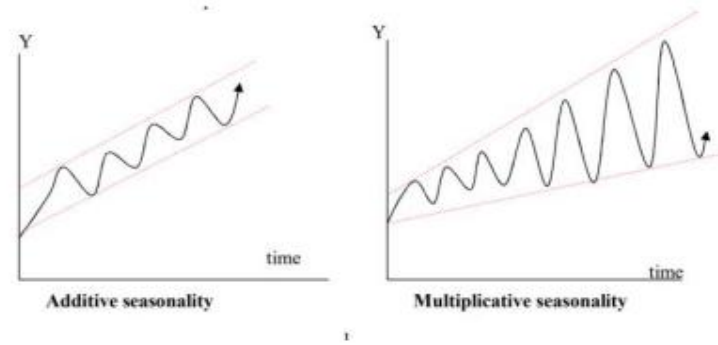


Figure 15 multiplicative and additive models

The multiplicative model is based on the assumption that the four components of a time series are not necessarily independent and can relate to each other; whereas in the additive model, we assume that the four components are independent of each other.

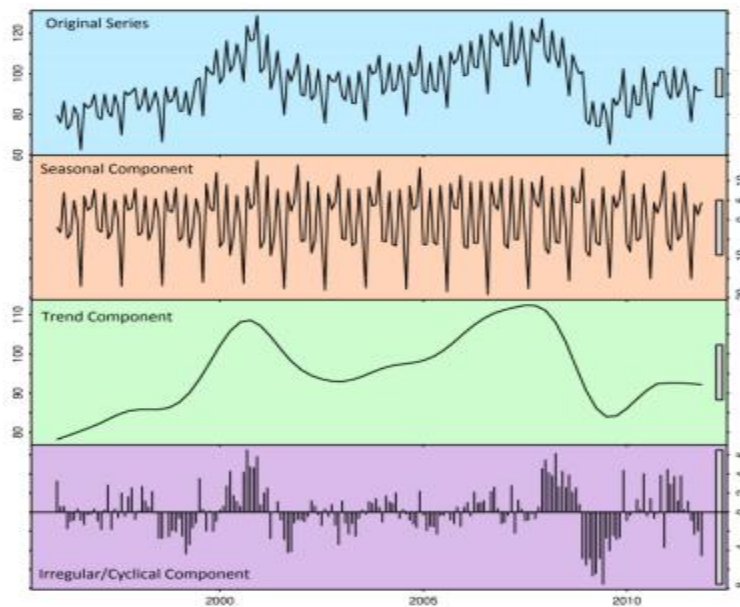


Figure 16 multiplicative and additive models

7.3 Time-Series Transformations

Time series data often requires transformation before forecasting or analysis. Transforming time series is often useful since many forecasting methods are most appropriate for time series that conform to some specific assumptions including the stationarity assumption which is very common in many time series techniques.

Transforming time series can also simplify the forecasting problem by removing trend and seasonal components. If any transform operation is applied to time series data, a similar inverse operation needs to be applied to the prediction. Identifying the best transform method can be very difficult for a given prediction problem as there are no standard rules regarding time series data preprocessing. In what follows, we present the most common transform operations for time series analysis and forecasting.

a) Log Transformation :

The log transformation is a type of power transformation where a logarithmic function is applied to time series values. This transformation is often used in time series forecasting and analysis to stabilize the variance of a series. In other words, forecasts based on $x_t = \log y_t$ can be much better than direct predictions of y_t if x_t has a more stable variance than y_t . It is also useful for forecasting positive time-series that grow exponentially.

b) Box-Cox Transformation:

The Box-Cox Transformation is also a power transform which depends on the parameter λ that aims at stabilizing the variance of non stationary time series. If $\lambda = 0$, we obtain the log transformation. If y_t represent the different observations of a time series, then Box-Cox transformation can be defined as follows :

$$\begin{cases} \log y_t & \text{if } \lambda = 0; \\ \frac{y_t^\lambda - 1}{\lambda} & \text{otherwise.} \end{cases}$$

The forecasting results generally don't depend on the value of λ , however prediction intervals are very sensitive to this parameter [29].

c) Differencing:

The Difference transformation is a very famous time-series transformation that aims at removing trend and seasonality. While power transformations aim at stabilizing the variance of a time series, differencing aims at stabilizing the mean of a time series.

The differenced series can be obtained by computing the differences between consecutive observations in the original data and can be written as :

$$y_t' = y_t - y_{t-1} \quad (1)$$

Equation (1) is called first-order differencing and if applied to a series it will remove linear trend. If the obtained differenced series is not stationary, a second-order differencing may be necessary and it is given by :

$$y_t'' = y_t' - y_{t-1}'$$

In order to remove the seasonal structure of a time series, a seasonal difference is often necessary and it is obtained by computing the difference between an observation and the previous observation from the same season :

$$y_t' = y_t - y_{t-m}$$

Where m is the number of seasons.

d) Standardization and Rescaling :

Standardization is the rescaling of the distribution of values so that the mean of observed values is zero and the standard deviation is one. This can be achieved by subtracting the mean and dividing the result by the standard deviation of the series. Time series data can also be rescaled and the most common approach is rescaling the data from the original range so that all values are within the range of 0 and 1. Scaling the time-series improves the convergence of forecasting methods (particularly ML methods) when applied to the time-series.

7. 4 . Statistical Methods

Statistical methods are the most popular methods in time-series forecasting in general and in demand and sales forecasting in particular. They have a long standing history and a strong mathematical basis. Although they are sometimes described as traditional methods, it has been shown through many empirical studies that they outperform sophisticated Machine Learning methods. In what follows, we describe some of the most popular statistical time-series forecasting methods including the simplest ones which are usually used as benchmarks.

a) Naive Method :

Naive method is a forecasting technique in which forecasts are equal to the last observed value. That is :

$$\hat{y}_{t+h} = y_T$$

If the time series is seasonal, it is better to use another variant of the naive method, that is the seasonal naive method in which forecasts are equal to the last observed value from the same season of the year. That is :

$$\hat{y}_{t+h} = y_{T+h-s(i+1)}$$

Where s is the seasonal period and i is the number of complete years in the forecast period prior to time T + h.

b) the simple average method:

Use the average of all relevant historical observations as a forecast for the next period. So we will have:

$$\hat{y}_{t+1} = 1/t \sum y_t$$

If we have a new observation:

$$\hat{y}_{t+2} = \frac{t\hat{y}_{t+1} + y_{t+1}}{t+1}$$

This method is used for stabilized series and the environment is generally unchanging.

c) The simple moving average method:

Previously, we discussed two simple forecasting methods, one is the naïve method which suggests that the value of the last observation is the new forecast, and the simple average method which assumes that the best predictor of what will happen tomorrow is the average of everything that has happened so far. And here we will detail another method which is a first step to go beyond simple average models, random route models and linear trend models, it is possible to extrapolate non-seasonal profiles and trends to the use of a moving average or a smoothing model. The basic

assumption of the smoothing and smoothing models is that the time series is locally stationary with a slowly varying average. Therefore, we take the moving average (local) to estimate the current value of the average and then use it as a forecast for the near future.

This can be seen as a compromise between the average model and the drift-free random walk model. The same strategy can be used to estimate and extrapolate a local trend. Moving average often called "smoothed" version of the original series because the short term average has the effect of smoothing the reliefs of the original series. By adjusting the degree of smoothing (the width of the moving average), we can hope to find an optimal balance between the performance of the average and random walk models. This is what we call the simple moving average (SMA) model and its equation to predict the value of Y at time $t + 1$ on the database going to time t is as follows:

$$\hat{y}_{t+1} = \frac{y_t + y_{t-1} + \dots + y_{t-m+1}}{m}$$

This average is centered on the period $t - (m + 1) / 2$, which implies that the estimate of the local average will tend to be late compared to the real value of the local average of about $(m + 1) / 2$ periods. So, we say that the average age of the data in the simple moving average is $(m + 1) / 2$ compared to the period for which the forecast is calculated: this is the time during which the forecast will tend to be behind the turning points. In the data. For example, if you average the last 5 values, the forecast will take about 3 late periods to respond to the turning points. If m is very large (comparable to the length of the estimation period), the SMA model is equivalent to the average model. As with any parameter in a forecast model, it is usual to adjust the value of k in order to obtain the best "fit" to the data, that is, the smallest forecast errors on average.

d) exponential smoothing models:

Exponential smoothing was proposed in the late 1950s and has motivated some of the most effective forecasting methods. Predictions produced using exponential smoothing methods are weighted averages of observations weights decrease exponentially as observations age. In other words, the more recent the observation, the higher the associated weight. This framework generates reliable forecasts quickly and for a wide range of time series, which is a considerable advantage and of major importance for industrial applications.

The exponential reading has several variants, the most used are simple exponential smoothing, double exponential reading and triple exponential smoothing.[30][31][32]

d.1. Simple exponential smoothing:

Simple exponential smoothing [31] is best applied to time series that is not. Have a dominant trend and do not exhibit seasonality. The smoothing constant α is used to control the speed at which the updated forecast will be adapted to the local (or average) level of the time series.

Let l_t be the level estimator at the moment t . Since l_{t-1} , once the observation at time t , Y_t , is available, the SES due to Brown (1959, 1963) updates the level estimator via the recurrence equation:

$$l_t = \alpha Y_t + (1 - \alpha) l_{t-1}$$

Where α is a smoothing parameter taking values in the interval $[0, 1]$. According to the equation, the estimator at time t , l_t is a weighted average of the last observation, Y_t , and of the estimator at time $t - 1$, l_{t-1} . The value of α can be used to adjust the sensitivity of the estimator to variations in the level of the series: the higher the value of α , the higher the weight received by Y_t , the more sensitive the estimator is to variations in level. The forecast h step by step made at the moment t by SES is:

$$\hat{y}_{t+H} = l_t, h > 0$$

d.2. double exponential smoothing:

SES does not work well for forecasting time series with a local linear trend.

Holt [31] extended SES to allow forecasting of data with a trend. This method involves a forecasting equation and two smoothing equations (one for the level and one for the trend)

Let b_t be the estimator of the slope at time t . Since l_{t-1} and b_{t-1} , once the observation at time t , Y_t , is available, Holt's method updates the estimators using the recurrence equations

$$\begin{aligned} l_t &= \alpha Y_t + (1 - \alpha) (l_{t-1} + b_{t-1}) \\ b_t &= \beta (l_t - l_{t-1}) + (1 - \beta) b_{t-1} \end{aligned}$$

where α and β are smoothing parameters taking values in the interval $[0, 1]$. The forecast ***h - step - ahead*** made at an instant t by Holt's method is:

$$\hat{y}_{t+H} = l_t + h b_t, h > 0$$

d.3. triple exponential smoothing:

Neither SES nor Holt's local linear method is appropriate for forecasting time series with seasonal changes. To account for seasonal changes, Winters [33] generalized the local linear method of Holt and proposed the so-called Holt-Winters method, which includes three methods.

Winter proposed a new method which consists in updating the equations, one for the level, one for the slope and one for the seasonality. Depending on whether seasonality is combined with the linear trend in an additive or multiplicative way, there are two versions of the Holt-Winters method. [34]

i. Holt-Winters additive method:

The form of the component for the additive method is:

$$\begin{aligned} l_t &= \alpha (Y_t - s_{t-m}) + (1 - \alpha) (l_{t-1} + b_{t-1}) \\ b_t &= \beta (l_t - l_{t-1}) + (1 - \beta) b_{t-1} \\ s_t &= \gamma (Y_t - l_{t-1} - b_{t-1}) + (1 - \gamma) s_{t-m} \end{aligned}$$

ii. Holt-Winters multiplicative method:

The form of the component for the multiplicative method is:

$$\begin{aligned} l_t &= \alpha * (Y_t / s_{t-p}) + (1 - \alpha) (l_{t-1} + b_{t-1}) \\ b_t &= \beta (l_t - l_{t-1}) + (1 - \beta) b_{t-1} \\ s_t &= \gamma Y_t l_t + (1 - \gamma) s_{t-p} \end{aligned}$$

where s_t is the seasonal factor at the moment t , α , β and γ are smoothing parameters taking values in the interval $[0, 1]$ and m is the length of a complete seasonal cycle.

The forecast h - step - ahead made at an instant t by the Holt-winters method is:

$$\hat{y}_{t+H} = lt + hbt + st - m + h, h > 0$$

e) ARIMA method:

An autoregressive integrated moving average, or ARIMA, is a statistical analysis model that uses time series data to better understand the data set or to predict future trends. which was introduced by Box and Jenkins in 1970 [35].

It is a form of regression analysis that measures the strength of a dependent variable relative to other changing variables. The objective of the model is to predict future developments in the securities or financial markets by examining the differences between the values in the series rather than the actual values.

An ARIMA model can be understood by describing each of its components as follows:

e.1. Autoregression (AR):

refers to a model that shows a changing variable that regresses to its own lagged or earlier values. Integrated (I) represents the differentiation of raw observations to allow the time series to become stationary, that is to say that the data values are replaced by the difference between the data values and the previous values.

e.2. Moving average (MA):

incorporates the dependence between an observation and a residual error of a moving average model applied to offset observations.

The ARIMA Model is designated by $ARIMA(p, d, q)$.

Or :

P: the number of offset observations in the model; also known as shift order.

D: the number of times the raw observations are differentiated; also known as the degree of differentiation.

Q: the size of the moving average window; also known as the moving average order.

The equation is given by:

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$$

$\hat{y}_t$: the value.

e_t : the error at time t .

ϕ_t, θ_t : coefficients.

μ : noise.

e.3. The integrated component:

Which represents the amount of differentiation to be made on the stationary appointment series.

8 . Machine Learning Medels:

Machine learning is machine learning or statistical learning: a field of study of artificial intelligence which is based on statistical approaches for give computers the possibility to “learn” from data, that is to say to improve their performance in solving tasks without being explicitly programmed

for each one. More broadly, it concerns design, analysis, development and the implementation of these methods.

Applied to the Supply Chain, Machine Learning makes it possible to project and analyze very finely the demand of its customers in order to anticipate future demands.

Several ML algorithms can be used to predict time series, the most notable being:

- Support for vector machines (SVM)
- K-nearest neighbor regression (KNN).
- Decision trees
- Grouping of K-means
- MLP neural network
- LSTM neural network

8.1. Support vector machines (SVM):

Support vector machines (SVM) are a popular tool.

Machine learning method for classification, regression and other learning tasks. The basic principle of SVM is to give a set of points which must be classified into two classes, find the separating hyper plane which maximizes the margin between the two classes. This will ensure a better classification of invisible points, that is to say a better generalization. In SVR, our objective is to find a function $f(x)$ which has at most a deviation from the targets actually obtained from the targets y_i for all the training data. Forecasting is a systematic effort to anticipate future events or conditions. Forecasting involves: precipitation forecast, stock market forecast, price forecast, temperature forecast, bank cash forecast, etc. The forecast is usually made using various statistical analyzes. [36]

8.2. K-Nearest Neighbor regression (KNN):

KNN is considered to be one of the simplest classification techniques. In K-Nearest The vector closest to K based on an appropriate distance vector is determined to classify the input characteristic vector X. The attribution of the vector X to a specific class is then performed by the majority of the closest neighbors to K at which they belong to. The nearest neighbor K algorithm is based on the distance and the voting function in the nearest K neighbors, the metric used is the Euclidean distance. KNN is a supervised classification technique which gives good performance results for an optimal value of K. Like any other supervised learning algorithm

KNN has two phases, namely the training phase and the test phase. In the learning phase, the classifier receives data values to train it. The training data is associated with labels representing their class. During the test phase, the KNN classifier receives unlabeled data points and the algorithm generates a list of the K closest data values. The class of unlabeled data points is determined by these values. KNN algorithm includes steps following: [37]

Step 1: Determine an appropriate distance metric

Step 2: In this phase, the predictor is trained for data sets whose classes are known to us. Store all training data sets X in pairs according to the selected characteristics $X = \{(x_i, y_i), i = 1, \dots, n\}$, where x_i is a learning model in the learning data set, y_i is the class assigned to it and n represents the number of learning models.

Step 3: In this phase, the predictor receives data sets whose class is not known to us. For the identification of the class of these datasets, the distance between the newly introduced entity and the already stored entities is calculated.

Step 4: The nearest K neighbors are chosen and invited to vote for the new dataset class.

8.3 Decision Tree:

CART was introduced by Breiman et al. (1984) [38]. It is a non-parametric learning technique for decision trees which gives classification or regression trees, depending on the nature of the variable (categorical or numeric). Decision trees are generated by a group of rules based on variables in the modeling dataset (Steven, 2014). The rules based on the values of the variables are chosen to develop the best distribution in order to differentiate the observations according to the dependent variable. Once a rule is selected and divides a node in half, each "child" node is also subjected to the same process (that is, it is a recursive procedure of splitting a node). The division stops when CART no longer detects any gain or when certain predefined stopping rules are observed. Otherwise, the data is split as much as possible and the tree is then pruned. Each branch of the tree ends with a terminal node. Each observation falls into one and exactly one terminal node, and each terminal node is uniquely defined by a set of rules. The algorithm produces binary trees, each internal node has exactly two outgoing edges, while the two algorithms ID3, C4.5 give decision trees with variable branches per node. The Hunt-based algorithm is due to the fact that it is used for regression analysis using regression trees. The regression analysis function is used to predict a dependent variable (result) from a set of predictor variables over a given period. The decision tree CART is a recursive binary partitioning procedure mastering the attributes of continuous and nominal processing, both targets and predictors. In CART, trees are cultivated using the Gini Index attribute selection for construction and the split procedure, to a maximum size without the use of the surfacing rule, then pruned (essentially divided by split) at the root by pruning cost complexity. Pruning of cost complexity is used to eliminate little branches decision tree to improve accuracy. The CART mechanism is intended to produce not one, but a sequence of nested pruned trees, all of which are candidates for optimal trees.

$$\text{Gini index: } 1 - \sum_{i=1}^n (p_i)^2$$

Pseudocode for the construction of CART trees by exhaustive research:

- I. Start at the root node.
- II. For each X , find the set S which minimizes the sum of impurities from the nodes in the two child nodes and choose the division $\{X * \in S\}$ Which gives the global minimum X and S .
- III. If a stop criterion is reached, quit. Otherwise, apply step 2 to each child node.

8.4 Grouping of k-means:

The best known partitioning and the most used methods are k-means. The k-means algorithm takes the input parameter k and partitions a set of n objects into k clusters so that the intra-cluster similarity is high, but the similarity between the clusters is low. Cluster similarity is measured in terms of the average value of the objects in a cluster, which can be thought of as the center of gravity or the center of gravity of the cluster. [39] First, it randomly selects k objects, each of which initially represents an average or a cluster center. For each of the remaining objects, an object is assigned to the cluster for which is the most similar, based on the distance between the object and the cluster mean. It then calculates the new average for each cluster. This process iterates up to the criterion

The function converges.

The K-Means grouping is one of the most famous clustering algorithms applied in different types of fields such as biology and zoology, medicine and Psychiatry, sociology and criminology ... etc.

8.5 MLP neural networks:

Multilayer neural networks (or perceptrons) define a class of functions whose interest is to be able to approach any continuous function with compact support (see density theorem). No other type of neural network will be studied and thereafter, any neural network will be considered as multilayer (therefore not the Kohonen networks). [40]

The model for a neural network (MLP) is defined recursively as follows:

$$x_{l+1} = \begin{cases} g(\theta_l x_l), & \text{si } l = 0, 1, \dots, L-1 \\ f(\theta_L x_L), & \text{si } l = L \end{cases}$$

where L is the number of hidden layers. We define the partitioned weight matrix

$\theta_l = (bl \mid Wl)$ As being the matrix comprising the weight matrix $Wl \in R_{nl+1} \times nl$ Between the neurons and the bias vector $bl \in R_{nl+1}$. We take the following notations $n_0, n_1, \dots, n_L, n_{L+1}$ to signify the respective number of neurons in each of the $L+2$ layers of the network going from the input layer to the output layer.

Thus, x_0 is seen as an explanatory vector of variables for the network, while x_{L+1} represents the vector or scalar of variable (s) explained by it. The vectors $x_1, \dots, x_L$ contain the neurons for each of the hidden layers of the network.

We choose the activation function g in order to make the connections between the different hidden layers of a neural network. We will choose an activation function that can be distinct for leaving the network, denoted by f . This last transforms the neurons of the last layer into output for the network. For example, for a binary classification task, we will have the choice to use f equal to the sigmoid or soft max function depending on whether we use one or two Neurons respectively for the output layer. In the case where we have a classification task where the number of categories is $K \geq 2$, the softmax function is often used in practice since it allows to associate a respective probability to each of the different classes.

8.6 LSTM neural networks:

To understand what the LSTM model is, short term memory, and how works, ANNs and RNNs need to be briefly described. [40]

A. ANN:

The Anns are vaguely inspired by the neural connections of our brain, trying to mimic the neural pathways and their behavior. ANN, robust and self-adapting, generally provides satisfactory solutions to non-linear problems difficult to implicitly implement, as well as the resolution of tasks such as speech recognition, natural language processing and forecasting. The network has so-called layers between input and output, where increasing the number of layers also increases the complexity of the network. By supplying the network with known data, the network is formed by selecting the weight relative to the desired output of the input. The weights are scalar adjusted by the network to reduce the error between the, by the Concluded network, the desired output and the actual output using the error

gradient with for the weights. These types of networks are not suitable for sequential data, however. As the network does not have the memory of the previous time steps, it cannot analyze the sequential data. Therefore, a memory of some sort is desired.

B. RNN:

RNN, recurrent neural networks, linking several layers of networks, where information from previous time steps, in addition to the output, is carried over to future time steps. As the input and parameters of each layer are processed, the outputs of the previous layers are taken into account, which gives the network some form of memory.

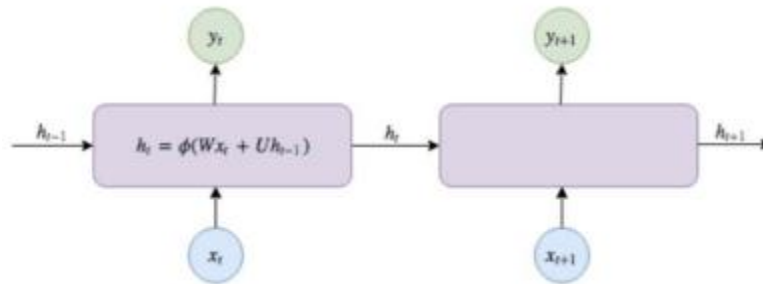


Figure 17 An RNN where W is a weight matrix and U is a matrix Consisting of what is called bias and ϕ is an activation function, making the network nonlinear

As a time dimension must be taken into account for RNNs, the gradient can become relatively complex, which means that the information from previous time steps begins to disappear or to be amplified in a way significant. These phenomena are called a gradient that is disappearing or exploding. Because of these phenomena, it can be difficult for the network to remember the information in The previous time steps are correct if they are far out of order. LSTM is a solution to these problems.

C. LSTM

LSTM networks implement a door-to-door cell to store information, similar to computer memory. Unlike the aforementioned networks, the LSTM cells also learn when to authorize reading and writing of information from previous time steps. The LSTM model therefore solves the problem of the disappearance or explosion of the gradient and allows the network to memorize information correctly. Far back in the sequence. The following figure shows the internal components of the LSTM cell. A cell manages a data time step and transmits the selected information to the next cell.

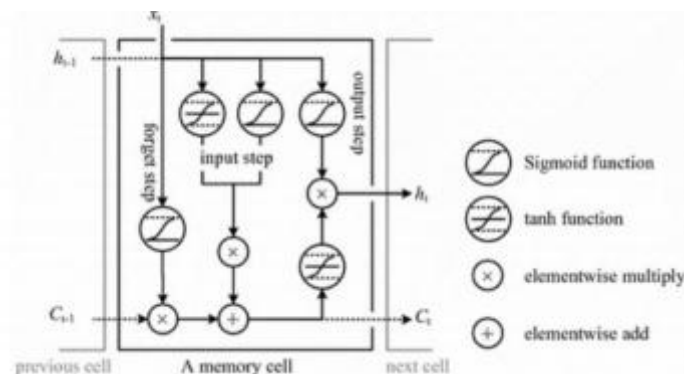


Figure 18 LSTM cell

X_t : Enter at time step t

C_t : additional dependencies recalled from previous time steps which are added to the ordinary output

h_t : output of time step t

9.1 Hybrid models:

Another approach to forecasting sales or demand is hybridization.

Hybrid methods combine the features of multiple forecasting methods to take advantage of the benefits of each. The combination may involve different methods of ML, different statistical methods or both methods of ML and statistical methods.

Examples: Box-Jenkins and Artificial Neural Networks.

9.1.1 Box-Jenkins and Artificial Neural Networks:

This hybrid method uses in combination the static Box-Jenkins (BJ) method and the neural network (ANN), a case of using this method appeared in Kuwait for forecasting water consumption.

ANN, together with the BJ approach, has been shown to model monthly water consumption in Kuwait. The BJ approach was used for the first time to predict missing values for monthly water consumption due to of Iraq's invasion of Kuwait. Once the monthly unregistered water consumption was predicted, the BJ approach was then used to discover the delayed variables or the appropriate input nodes in the input layer of neurons. This approach offers a superior and reliable alternative to traditional methods for choosing the appropriate number of delays or variables shifted by a time series. It was found that when the variables of the input layer in ANN are chosen based on the BJ approach rather than traditional methods, the average relative error of the training and test datasets is reduced by 24%.

9.2 causal patterns:

A causal method is an estimation method based on the assumption that the future value of a variable is a mathematical function of the values of other variables. Used when sufficient historical data are available and the relationship (correlation) between the dependent variable to be predicted and the associated independent variable (s) is well known. Also called a causal forecasting model.

The most common quantitative causal forecasting method is regression analysis

There are two types of regression models:

9.2.1 Simple regression model :

In this model, a single independent variable predicts the dependent variable. For example, the sale of soft drink on a given day depends on the temperature of that day.

The simple linear model is expressed using the following equation:

$$Y = a + bX + \epsilon$$

Or :

Y- A dependent variable

X - An independent variable

a- intercept
b- slope
 ε - residual (error)

9.2.2 multiple regression model:

Multiple linear regression analysis is essentially similar to the simple linear model, except that several independent variables are used in the model. The mathematical representation of multiple linear regression is as follows:

$$Y = a + bX_1 + cX_2 + dX_3 + \varepsilon$$

Or :

Y- A dependent variable

X_1, X_2, X_3 -An independent variable

a- intercept

b, c, d - slopes

ε - residual (error)

10 . Evaluation and validation of forecasting models:

10 . 1 Procedure for evaluating training and test sets:

It is important to assess the accuracy of forecasts using authentic forecasts. Therefore, the size of the residuals is not a reliable indication of the likely magnitude of the true forecast errors. Prediction accuracy can only be determined by considering a model's performance on new data that was not used when fitting the model.

When choosing models, it is common to separate the available data into two parts, the training and test data, the training data being used to estimate the parameters of a forecasting method and the test data for assess its accuracy. Since test data is not used to determine the forecast, it should provide a reliable indication of how likely the model is to predict new data.



Figure 19 Illustration of the division of the training / test dataset

The size of the test set generally corresponds to around 20% of the total sample, although this value depends on the length of the sample and the duration of the forecast. The test set should ideally be at least as wide as the horizon maximum forecast required.

The following points should be noted.

- A model which corresponds well to the training data will not necessarily be well planned.
- A perfect fit can always be obtained using a model with enough parameters.
- Over-fitting a model to the data is as serious as not identifying a systematic model in the data.

Certain references describe the test set as “reserve set” because this data is “kept” from the data used for the adjustment. Other references call the learning game "intra-sample data" and the test define "data outside sample". We prefer to use "training data" and "test data" in this work.

10.2 Cross-Validation Time-Series Procedure:

In this procedure [41], there is a series of test sets, each consisting of a single observation. The corresponding learning set includes only the observations prior to the observation constituting the test set. Thus, no future observations can be used in the construction of the forecast. The following diagram illustrates the training and test series, where the blue observations are the training sets and the red observations are the testing sets.

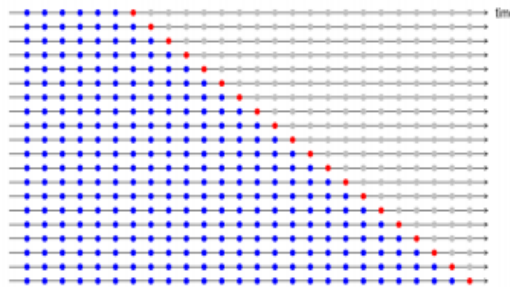


Figure 20 Illustration of the cross-validation procedure for different forecast horizons-1-

Forecast accuracy is calculated by averaging over test sets. This procedure is sometimes called "evaluation on a sliding forecast origin" because the "origin" on which the forecast is based is carried over time.

With time series forecasts, one-step forecasts may not be as relevant as multi-step forecasting. In this case, the cross-validation procedure based on a sliding forecast origin can be modified to allow the use of errors in several stages. Suppose we are interested models that produce good 4-step forecasts. Next, the corresponding diagram is presented below.

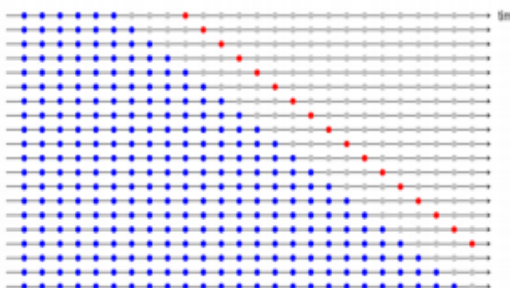


Figure 21 Illustration of the cross-validation procedure for different forecast horizons-2-

10 . 3 Forecast precision measurements:

Forecast error is a measure of the accuracy of our forecast over a given period. It is calculated as the actual demand minus the forecast, It can be written as:

$$e_t = Y_t - \hat{Y}_t$$

With Y_t the real value and $\hat{Y}_t$ is the forecast at the moment t

We can measure forecast accuracy by summarizing forecast errors in different ways.

10 . 3 . 1 Scale dependent errors:

Forecast errors are on the same scale as the data. Accuracy measures based only on e_t therefore depend on the scale and cannot be used to make comparisons between series involving different units.

The two most commonly used scale dependent measures are based on absolute or squared errors:

$$MAE = \frac{1}{N} \sum_{t=1}^N |e_t|$$

$$MSE = \frac{1}{N} \sum_{t=1}^N e_t^2$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^N |e_t^2|}$$

Where N is the number of observations in the test sample.

When comparing forecasting methods applied to a single time series or multiple time series with the same units, MAE is popular because it is easy to understand and calculate. A forecasting method that minimizes the MAE will lead forecasts of the median, while minimizing the $RMSE$ will lead to forecasts of the average. Therefore, $RMSE$ is also widely used, although it is more difficult to interpret.

10 . 3 . 2. Percentage error :

The percentage of error is given by:

$$pt = 100 / Y_t$$

Percentage errors have the advantage of being devoid of units and are therefore frequently used to compare forecast performance between data sets. The most commonly used measure is:

$$MAPE = 100 \times \frac{1}{N} \sum_{t=1}^N \left| \frac{e_t}{Y_t} \right|$$

Measurements based on percentage errors have the drawback of being infinite or indefinite if $y_t = 0$ for all t in the period considered and of having extreme values if any y_t is close to zero. Another often overlooked problem with the error percentages is that they assume that the unit of measure has a significant zero. For example, a percentage error does not make sense when measuring the accuracy of temperature forecasts on the Fahrenheit or Celsius scales, because the temperature has an arbitrary zero.

They also have the disadvantage of imposing a heavier penalty on negative errors than on positive errors. This observation led to the use of the so-called "symmetrical" MAPE proposed by Armstrong [42] which was used in the M3 forecast competition. It is defined by:

$$sMAPE = \frac{100}{N} \times \sum_{t=1}^N \frac{|\hat{Y}_t - Y_t|}{(|\hat{Y}_t| + |Y_t|)/2}$$

However, if Y_t is close to zero, $\hat{Y}_t$ is also likely to be close to zero. Thus, the measurement always involves dividing by a number close to zero, making the calculation unstable. In addition, the value of $sMAPE$ can be negative and therefore does not constitute a measure of "absolute percentage errors".

Hyndman & Koehler [43] recommend not using $sMAPE$. It is included here only because it is widely used, although we do not use it in this work.

10.3.3 Errors At Scale :

Hyndman and Koehler first proposed scale errors in 2006 [43]. The precision measurements of the staggered errors are based on the learning MAE using a simple forecasting method such as a naive method. The best known form of scaled error is the Average Absolute Scaled Error (MASE) and it is given by:

$$MASE = \frac{\sum_{t=1}^N |Y_t - \hat{Y}_t|}{\frac{N}{N-1} \sum_{t=2}^N |Y_t - Y_{t-1}|}$$

The MASE is estimated similarly for a seasonal time series, but instead of using a naive method, it uses a seasonal naive method to resize the error. Like the percentage of errors, MASE is very useful for comparing the accuracy of forecasts generated by time series of different scales. However, unlike MAPE, the measured mean absolute error (MASE) is symmetrical since it also penalizes errors in large forecasts and small forecasts. In addition, when the actual values are zero or close to zero, MASE does not have extreme values. With regard to interpretability, MASE is easily interpretable: a MASE value less than 1 means that the forecast method provides better forecasts than the naive average forecasts calculated on the training set. Conversely, if it is greater than 1, it means that the forecasting method is worse than a naive forecasting method.

Conclusion :

Nowadays, world is a global village and every organization in the world can have access to any customer segment without doing any investment in other country through internet technology and e-business concept. In this tough competition, the only organization can survive which meet up the ever changing demands of the customers and deliver the right product, at right place, in the right quantity, at the right time and in right price. This cycle could not be completed without inter and intra organizational coordination, customer and supplier relationship management, transportation & distribution management and rapid flow of information, which means a strong integration among all functions of the supply chain is required.

The modern information and communication technology provides software solutions (SCM), web based internet facilities (e-business, e-procurement, e-commerce, etc.) and centralized databases which help out in efficient communication and coordination within as well as outside the organization. Supply Chain management software solution provide analysis of demand forecasting, production planning, inventory management and distribution which reduces uncertainty with less process and inventory holding costs. In this competitive market, the organizations have very limited margins and ICT investments are the best way to get efficient return on investment in shape of competitive advantage.

In this chapter, we have presented Demand forecasting it is generally based on time series forecasting. Designing efficient and robust algorithms for accurate demand forecasting is one of the most important challenges in the field of time series analysis and forecasting. Designing such robust demand-forecasting models cannot be achieved without understanding the core elements of demand forecasting conceptual and theoretical framework, which was described in this chapter. Although, this chapter is not meant to be a detailed overview of demand forecasting conceptual framework, it aimed at elucidating the main concepts related to the topic of this thesis, including the state of the art of demand forecasting, which we discuss in the next chapter.

Chapter 2: State of Art

1 .Introduction :

In an era of greater demand uncertainty, higher supply risk, and increasing competitive intensity, supply chain (SC) excellence often hinges on the organization's ability to integrate and orchestrate the entire spectrum of end-to-end processes of acquiring materials or components, converting them into finished goods, and delivering them to customers. Since such ability can be enhanced by increased visibility across the end-to-end SC processes, many leading-edge organizations have attempted to enrich their information sources and share real-time information with SC partners.

Thus, SC management (SCM) is becoming more information intensive and its focus has been directed toward the substitution of assets (e.g., inventory, warehouses, transportation equipment) with information. Recognizing the increasing significance of information to SC success, SC professionals have explored various ways to better manage information and leverage it to make better business decisions. One of those ways may include artificial intelligence (AI) that has been in existence for decades, but has not been fully utilized in the area of SCM.

In general, AI is referred to as the use of computers for reasoning, recognizing patterns, learning or understanding certain behaviors from experience, acquiring and retaining knowledge, and developing various forms of inference to solve problems in decision-making situations where optimal or exact solutions are either too expensive or difficult to produce . Put simply, the main objectives of AI are to understand the phenomenon of human intelligence and to design computer systems that can mimic human behavioral patterns and create knowledge relevant to problem-solving. Thus, AI should have the ability to learn and comprehend new concepts, learn from experience (“on-their-own”), perform reasoning, draw conclusions, impute meaning, and interpret symbols in context. Due to such ability, AI has been successfully applied in areas such as game playing, semantic modeling, human performance modeling, robotics, machine learning, data mining, neural networks, genetic algorithms (GAs), and expert systems.

2 Methodology:

AI is known for its ability to think like humans, act like humans, think rationally, and act rationally (Russell and Norvig 1995). Thus, with respect to these distinctive features, AI can be further classified into a number of sub-fields: artificial neural networks (ANN) and rough set theory (“*thinking humanly*”); machine learning, expert systems, and GAs (“*acting humanly*”); fuzzy logic (“*thinking rationally*”); and agent-based systems (“*acting rationally*”). These sub-fields are discussed below.

2.1 Artificial neural networks :

The theory of an ANN was predicated on the way that the living organ's brain cells, namely neurons,

function. Using the interconnected network of computer memories, ANN can learn from experience, distinguish features, recognize patterns, cluster objects, and process ambiguous or abstract information. To elaborate, an ANN is composed of a number of nodes which correspond to biological neurons. Those nodes are connected to each other by links. Each link has a numeric weight assigned to it. The links and their weights are the primary means of the long-term memory storage. The network processes information in such way that the output of one neuron is an input to another neuron linked to it. The weights are responsible for the strengthening or weakening of the information passed via the link. The links are placed and the values of weights are set in a process called learning. ANN can be taught to respond to various data patterns according to our wishes or to learn hidden interrelationships among the data. Once the network is initialized, ANN can be modified to improve its performance using an inductive learning algorithm and be trained in either supervised or unsupervised environments (McCulloch and Pitts 1943, Russell and Norvig 1995).

ANN has been proved to be useful for semantic modeling due to its ability to learn to pronounce English vocabularies. In the logistics field, ANN can be useful for maneuvering autonomous vehicles using its image processing technique. Indeed, Pomerleau (1993) utilized ANN to steer a land vehicle along a single lane on a highway by mimicking the performance of a human driver. Although the application of ANN to auto-piloting land vehicles is still limited to a certain type of road condition and traffic environment, it showed promise in autonomous vehicle navigation. ANN has also been applied to a traditional lot-sizing problem (Gaafar and Choueiki 2000).

In a broader context, ANN was successfully utilized to develop hierarchical SC planning that determined the time/capacity needed for setups, estimated optimal lot-size between successive SC processes, and linked inventory and scheduling decisions at the lower level to demand and production planning decisions at the higher level (Rohde 2004). As such, ANN is designed to reflect the interconnectivity and interdependence of SC planning processes better than traditional operational research (OR) techniques that were primarily intended for solving less-integrated sub-problems (e.g., inventory or production or transportation planning) of SC planning.

2.2 . Rough set theory :

Rough set theory was introduced by Pawlak (1982) as a way of synthesizing the approximation of concepts from the acquired data using a data table comprising of one or more classification attributes. These attributes include equivalence classes, indiscernibility relations, set approximation, and rough membership. These classification attributes are necessary to implement mechanisms similar to those that humans use for object classification and recognition. Also, based on the description of common features of indiscernible objects, they can be useful for developing decision rules (Pawlak 1984, 1989, 1997). As such, rough set theory can be employed to classify decision criteria and then develop decision rules relevant to SCM. For example, Li et al. (2007) used rough set theory to select the most desirable supplier among a pool of qualified suppliers with respect to multiple but conflicting supplier selection criteria. Zheng and Lai (2008) also utilized rough set theory to develop multiple criteria decision rules for measuring dynamic SC performance.

2.3 Machine learning:

Machine learning, coined by Samuel (1995), was designed to provide computers with the ability to learn without being explicitly programmed. In other words, machine learning

investigates ways in which the computer can acquire knowledge directly from data and thus learn to solve problems (Ratner 2000). Depending on the method of learning tasks, machine learning can be sub classified into a number of categories: concept learning that is designed to correctly recognize or construct concepts relevant to future decision-making processes following an inductive learning process; decision tree learning that aims to classify all the objects by testing their values for certain properties and then constructing a decision tree; perception learning that aims to acquire useful knowledge, reduce the error, and solve decision problems using a single layer of the network called a “perceptron”; fort Bayesian learning that trains the computer to learn representations of probabilistic functions; and reinforcement learning that trains the computer to perform at high levels by giving constant feedback in a form of rewards (see e.g., Luger 2002 for various machine learning processes).

Regardless of differences in learning tasks, machine learning techniques often attempt to mimic nature based upon the knowledge and experience that the human race has amassed over the eons of its existence. Some of the machine learning techniques were motivated by the neurological studies of the human brain function, some by the processes dictating human evolution, some by the mathematical theory of human knowledge acquisition and reasoning and some by the sociological theory behind human collaborative behavior. In particular, machine learning can be a useful tool for understanding the motivation behind collaborative behavior among SC partners for sharing critical information and improving ways of strengthening the partnership among SC partners through the organizational learning process. For instance, Charbonneau et al. (2008) recently used machine learning to forecast the distorted demand information at the end of a SC, namely the bullwhip effect, if that demand information was not shared among the SC partners due to lack of collaboration.

2.4 Expert systems:

Expert systems represent computer programs capable of emulating human cognitive skills such as problem-solving, visual perception and language understanding, and are capable of performing reasoning about a problem domain complex enough for a considerable amount of human expertise (Jackson 1999). Expert systems are comprised of four components: knowledge base, inference engine, justifier/scheduler, and user interface. To elaborate, the knowledge base is the repository of the rules, facts, and knowledge acquired from the human expert. The inference engine is a cluster of problem-solving programs (the “brain” of the expert system) that coordinate the searching, reasoning, and inference based on the rules of the knowledge base. The justifier explains how and why an expert arrives at a solution, while the scheduler is set up to coordinate and control the sequencing rules. The user interface facilitates communication and interaction between the system and its user through a series of user queries (Awad 1996).

Since an expert system operates at a level and in terms and concepts with which the user can feel affinity, the expert system is much more easily understood by the practitioner and thus more applicable to practical SC problems (Basden 1984). In particular, the expert system is known to increase productivity in managing logistics (Eom and Karathanos 1996). For instance, Allen (1986) successfully solved multiple echelon inventory control problems in the US Air Force Logistics Center using more than 400 rules and heuristics created by a number of experts and verified the effectiveness and efficiency of the proposed expert system in terms of improved inventory accuracy. Sullivan and Fordyce (1990) also developed a real time, transaction-based expert system that aimed to schedule, monitor, and

control the logistics flow of the IBM's semiconductor facility near Burlington, Vermont. They reported that the use of the expert system increased IBM's product output by 35% and saved approximately \$10 million of capital expenditure.

The application potential of the expert system is limitless in the SC field, as evidenced by its successful application to air traffic control, spatial mapping, airline yield management, and vehicle repair and maintenance scheduling (Findler 1987, Jefferies and Yeap 2008). In addition, an expert system can be more useful than a single more sophisticated forecasting method (or a mixture of such methods) for demand forecasting at every stage of the SC, in terms of forecasting accuracy, computational speed, user understanding, and cost effectiveness (DeLurgio 1998). Other intriguing applications of the expert system to the SC discipline include: concurrent product design and planning (Zha et al. 1999); gas pipeline operations (Uraikul et al. 2000); supplier evaluation (Kwong et al. 2002); evaluation and selection of third-party logistics (3PLs) providers (Yan et al. 2003); formulation of a logistics strategy (Chow et al. 2005); and production planning and control (Lawrynowicz 2007).

2.5 Genetic algorithms:

As a branch of evolution programs, a GA imitates the principles of natural evolution and derives a set of rules from natural selection processes that create organisms that most fit the surrounding environment. GAs have often been used to solve combinatorial optimization problems for which it is possible to construct a function that can estimate a fitness of a given representative (solution) to a given environment (problem). The GA encodes possible solutions to the problem in numerical strings called chromosomes. By iterative application of genetic operators (crossover, mutation, and selection) to a whole population of such chromosomes, the GA produces solutions that are not necessarily optimal, but quite satisfactory in terms of the fitness to the optimization problem.

In general, a GA is referred to as a stochastic AI technique that utilizes a solution search process that mimics natural evolutionary phenomena: genetic inheritance and Darwinian struggle for survival (Holland 1975, Michalewicz 1999). The GA typically comprises five components (Michalewicz 1999, Gen and Cheng 2000):

- (1) A genetic representation of potential solutions to the problem.
- (2) A way to create a population (an initial set of potential solutions).
- (3) An evaluation function measuring the fitness of solutions to see whether they will survive.
- (4) Genetic operators that alter the genetic composition of offspring. These operators include reproduction, crossover, and mutation. Reproduction is a process in which individuals (solutions) are copied through the selection of individuals that are the most fit. Crossover combines the features of two parent chromosomes (potential solutions) to form two similar offspring by exchanging corresponding attributes of the parents. Mutation randomly alters one or more features of a selected solution to introduce extra variability.
- (5) Parameter values that determine population size (how many individuals should be in the population); crossover rate (the probability that the individual will crossover); and mutation rate (the probability that a certain gene will mutate).

Although the GA aims to produce global optimal solutions efficiently, its population size cannot be infinite. Thus, its final solution may be biased due to the finite sampling of potential solutions. Also, the performance of a GA may depend on the specific rates of crossover and mutation.

As a result, a GA may suffer from premature convergence. Furthermore, the size of population does not usually improve the performance of a GA with respect to the speed of finding solutions (see, e.g., Holland 1975 and Goldberg 1989 for distinctive features of a GA). Nevertheless, GAs have been applied successfully to a variety of challenging SC network design problems.

These problems include: vehicle routing and scheduling (Malmberg 1996, Potvin et al. 1996, Chen et al. 1998, Park 2001); minimum spanning tree (Zhou and Gen 1998, 1999); delivery and pickup (Jung and Haghani 2000); bus network optimization (Bielli et al. 2002); and location– allocation problems (Hosage and Goodchild 1986, Jaramillo et al. 2002, Zhou et al. 2002, 2003, Min et al. 2005). In addition, a GA was employed to solve well-known logistics and purchasing problems involving facility layout (Tam and Chan 1998, Balamurugan et al. 2006); pallet loading (Fontanili et al. 2000); inventory control (Disney et al. 2000, Haq and Kannan 2006); container loading (Gehring and Bortfeldt 1997, Bortfeldt and Gehring 2001); material handling (Wu and Appleton 2002), delivery reliability assurance (Antony et al. 2006); freight consolidation (Min et al. 2006a, b); supplier selection (Rao 2007); and express courier services (Ko et al. 2007).

2.6 Fuzzy logic :

Fuzzy logic can be a powerful tool for building knowledge bases for particular domains and acquiring knowledge from the experts. In a broad sense, fuzzy logic uses expert opinions as an input to specify “good” and “bad” areas of each variable and then determines the likelihood of “goodness” and “badness” levels after comparing the input variable with the expert opinion (Tanaka 1997). Fuzzy logic is an extension of Boolean logic that was designed to conceptualize partial truth – somewhere between definitely true and definitely false. Typically, fuzzy logic consists of five basic components: linguistic variables, linguistic values, fuzzy sets, membership functions, and fuzzy IF-THEN rules. Thus, fuzzy logic is in contrast to crisp logic that is predicated on clear distinction between objects (or values). In other words, fuzzy logic can handle ambiguity, imprecision, and uncertainty of objects. For instance, fuzzy logic may help us answer questions of how cold the temperature is, how heavy a person is, or how expensive the product price is, without setting a clear-cut boundary.

In fuzzy logic, since the membership of an object in a fuzzy set takes a value between 0 and 1, the transition from membership to non-membership in the set is gradual. This gradual transition allows for the mathematical expression of objects with varying conditions and states. For example, a given temperature (object) can be expressed as cold to the degree of 0.1 and warm to the degree of 0.8; a car can belong to the set of a cheap car to the degree of 0.2 and an expensive car to the degree of 0.6. Thus, fuzzy logic can be useful for developing a set of rules for SC decision environments where subjective performance criteria have to be employed. For instance, in the airport location decision, a transportation planner may not know exactly how convenient the location of the airport is to the nearby shippers and carriers. Fuzzy logic has also been applied to solve the well-known traveling salesman problem (TSP) in a SC setting (Michalewicz and Fogel 2000). Other applications of the fuzzy logic to SCM include: supplier performance evaluation (Lau et al. 2002); inventory cost control (Wang and Shu 2005); measurement of the bullwhip effect (Balan et al. 2007);

agro-industry SC planning (Yandra et al. 2007); supplier selection (Carrera and Mayorga 2008); and order fulfillment (Amer et al. 2008).

2.7 Agent-based systems:

An agent-based system is one of the distributed problem-solving techniques that divides a decision problem into sub-problems and solves those sub-problems using independent entities called agents. Each agent can use different methodology, knowledge and recourses to process given tasks.

According to Reis (1999), an agent refers to an autonomous entity that can take certain actions to accomplish a set of goals and can compete and cooperate with other agents while pursuing its individual goals. An agent is characterized by its ability to exploit significant amounts of domain knowledge, overcome erroneous input, use symbols and abstraction, learn from the decision environment, operate in real time and communicate with others in natural language (Newell 1989).

Exploiting such characteristics, an agent-based system has often been employed to handle various SC issues including shop floor control (Van Dyke Parunak 1998, Shen et al. 2000, Usher 2003, Wang and Shen 2003); logistics planning (Satapathy et al. 1998); air traffic control (Iordanova 2003); aggregate demand planning and forecasting (Yu et al. 2002, Liang and Huang 2006); joint production planning (Lima et al. 2006); new product development (Liang and Huang 2002); order monitoring (Chen and Wei 2007); business-to-business negotiation (Ito and Saleh 2000, Lenar and Sobecki 2007); bidding evaluation (Kang and Han 2002); outsourcing relationship management (Logan 2000); customer relationship management (CRM) (Baxter et al. 2003); SC relationship management (Ghiassi and Spera 2003); SC performance assessment (Gjerdrum et al. 2001); SC coordination (Swaminathan 1998, Fox et al. 2000, Nissen 2001, Sadeh et al. 2001, Ono et al. 2003, Chan and Chan 2004, Lou et al. 2004, Xue et al. 2005); SC collaboration under uncertainty (Kwon et al. 2007); information exchange among SC partners (Garcia-Flores et al. 2000, Turowski 2002); information tracking across the SC (Zimmermann et al. 2001); material handling (Ito and Mousavi Jahan Abadi 2002); retail merchandise purchasing (Park and Park 2003); e-logistics (Santos et al. 2003); strategic e-procurement (Cheung et al. 2004); e-supply chains (Singh et al. 2005); traffic incident management (Tarver and Fae 2007); and the procurement of maintenance, repair, and operating (MRO) supplies (Nissen and Sengupta 2006).

A byproduct of agent-based methodologies that may be useful for solving complex SC problems is ant colony optimization. This algorithm mimics the social behavior of ants, who often find the shortest path to their food sources and nests using their innate capability to follow pheromone trails (hormone deposited by ants reflecting their collective memory) released by other ants.

To elaborate, an ant colony optimization algorithm is a meta-heuristic inspired by knowledge sharing behaviors of ants in solving combinatorial problems pertaining to different realms. The ant colony optimization algorithm is known to stabilize the solution with a reasonable amount of computational time without detriment to the solution accuracy, by exploiting the positive feedback provided by armies of ants working as multiple agents (Dorigo et al. 1996). Due to its past success in handling NP-hard problems, the ant colony optimization algorithm has been employed to solve TSP, vehicle routing problems, sequential ordering problems, and process plan selection problems within the SC framework (Gambardella and Dorigo 2000, Tiwari et al. 2006).

Conclusion:

Demand forecasting is vital for effective inventory management and control. Over the past years, many demand forecasting methods have been proposed. We provided in this chapter a cohesive overview of the state of the art related to demand forecasting. Hence, we reviewed the most prominent demand forecasting methods, which are essentially based on time-series forecasting. Demand forecasting is largely dominated by the univariate traditional statistical methods. Other methods have emerged as potential competitors of the traditional forecasting methods, most of them are based on Machine Learning. Combination and hybrid methods have also been shown to result in, at least, as accurate forecasts as other statistical or Machine-Learning-based methods. In the absence of a universal approach and metrics to assess the superiority of forecasting methods upon others, comparative empirical studies in the field have remained the best approach to compare the accuracy of forecasting methods.

Chapter 3 . Data Analysis

1 Introduction:

Forecasting demand for different products exhibiting varying demand patterns is the most important challenge facing Stock&Buy. In order to tackle these challenges, we discuss in the following chapters the solution that we propose to address the problem. This chapter

discusses the first phase of this project, which aims at analysing the existing solution but also at analysing the available demand data that can be used as part of building the solution.

The preliminary data analysis that we discuss in this chapter consists of exploring the different demand patterns that the products exhibit in order to appropriately choose the adequate forecasting solution later on. It also consists of studying any underlying cross correlations within the available data that may impact the demand. Before we dive in into the preliminary data analysis, we will briefly remind the functional and technical specifications of our project.

1 .Existing Model and Problem Statement:

Stock&Buy with its growing on-line business which has been attracting hundreds of customers all over the world, wants to improve the existing demand forecasting solution which is based on a simple linear regression and has the following limitations :

- The solution uses a very basic linear regression model that doesn't meet customers expectations in terms of forecasts accuracy. Linear Regression is an ML algorithm based on supervised learning. It predicts a dependent variable value y based on a given independent variable x by finding a linear relationship between x (input) and y (output).
- The existing model doesn't take into account cyclic changes such as seasonality.
- The existing model doesn't handle intermittency that slow-moving products, which account for a large proportion of products, exhibit.
- Estimation of demand for new products which don't have historical sales and demand data is not possible with the existing model.

These limitations resulted in a poor forecasting accuracy and required the proposal and implementation of a new forecasting module which can achieve reasonable weekly and monthly forecasting accuracy for different inventory items. The demand forecasting model shall allow retailers to accurately estimate inventory stocking requirements necessary to meet weekly and monthly service levels. Moreover, since a large proportion of products are slow-moving inventory items, the model shall be able to predict obsolescence risk for such products.

2 .Functional and Technical Specifications:

We describe in Table 1 and Table 2 the functional and technical specifications respectively of the forecasting tool that we were required to build. For each specification, a priority was attributed based on the needs expressed by our client.

We remind that the main goal of our project was to come up with a forecasting method accurately forecasts demand for different products. The way we would implement this method was an open-ended specification, hence the type of the solution (standalone or web API) was not explicitly defined by the client.

ID	Specifications	Priority
01	The forecasting tool shall allow for weekly and monthly demand forecasting for a given product variant.	MUST
02	The forecasting tool shall handle intermittent demand exhibited by slow moving products and, in case of product obsolescence, it shall predict obsolescence.	MUST
03	The forecasting tool shall forecast demand for a newly introduced product with no or few historical sales data.	SHOULD

Tableau 2 Functional specifications

ID	Specifications	Priority
01	The forecasting tool back-end shall be easily integrated into Stock&Buy environment (Cloud native micro-services).	MUST
02	The forecasting models shall use a recommended programming language in the field of Data Science such as R or Python.	MUST
03	The forecasting tool shall have a web UI to display forecasting models' results.	SHOULD

Tableau 3 Technical specifications

3 .The Challenge of Choosing the Forecasting Model:

In the literature of demand-forecasting, there are plenty of quantitative forecasting models, which can be described as statistical, Machine-Learning based or hybrid models. The diversity of the existing forecasting models which claim their superiority and accuracy in the field of demand-forecasting, represents a significant challenge in terms of choosing the most appropriate model for Stock&Buy demand-forecasting solution.

The initial project's charter delivered by Stock&Buy, expressed the desired forecasting model by our client, that is a Machine-Learning-based solution. Since ML has been extensively used in different prediction applications in the industry for the last few years, and the fact that it is based on Artificial Intelligence, opting for an ML solution would have been a promising and appealing solution for the retailers using the platform.

These findings were later discussed with our client who was open to expanding the project's forecasting model scope, to include any promising forecasting approach that achieves excellent forecasting accuracy.

We later decided to investigate several ML, statistical and hybrid forecasting models in order to design the most appropriate forecasting approach for Stock&Buy.

4 . Preliminary Data Analysis Approach:

Before studying the models that we can include in our solution, we first conducted a preliminary analysis, which aims at:

- Visualizing sales and demand historical data.
- Studying the main characteristics of demand and sales time-series such as stationarity to determine whether a preprocessing pipeline is required to make the data stationary so that we can use it with the statistical methods that are based on stationarity assumptions.
- Studying the cross correlation between demand and other time-dependent variables (in our case, the product's price) to determine if we should also investigate multivariate time-series models that are best suited to cross-correlated time-series.
- Categorizing demand for all products sold by each retailer to see if slow-moving products make up a large proportion of the products that are generally sold on the platform, which will require us to investigate intermittent-demand forecasting methods that are best suited for intermittent demand forecasting.

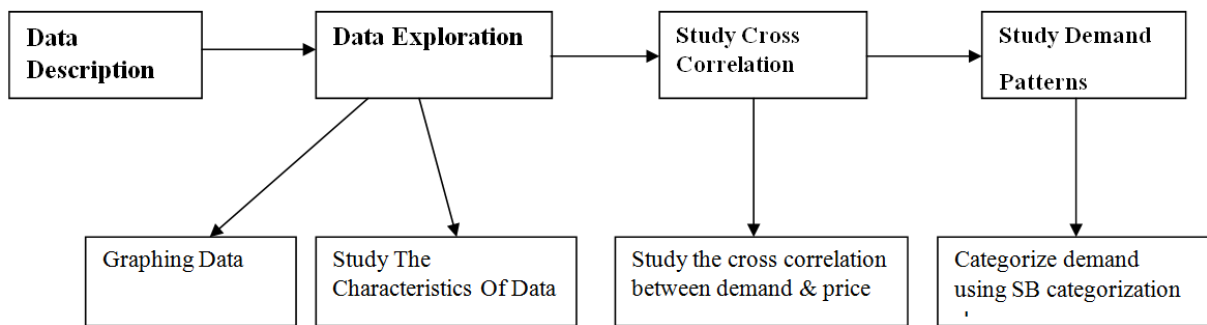


Figure 22 Preliminary data analysis process

At this stage, we used both sales data at the store level which represents daily sales for all product variants of each retailer, and sales data at the product variant level, which represents daily demand and sales data for a particular product variant. Figure 1 showcases the preliminary data analysis process that we went through.

4.1 Data Description:

We describe in this section the data that our Stock&Buy client provided us to carry out this project. Note that, these data only represent a proportion of the wealth of data available at the client side, since it covers retail data of three different retailers and not

all retailers. However, it is a common practice in Data Mining to validate the models using a representative proportion of the available data. Hence, it has been decided with the client to implement during this project, a proof of concept prototype that would be tested and validated using the data that we will describe hereafter. The solution that we will implement could then be generalized to other retailers. The choice of retailers to include in this project was based on their willingness to provide confidential retail data.

a . Provided Data:

We were provided with inventory and sales data of three different Stock&Buy customers, who have been using the e-commerce platform for more than three years, and who accepted to share with us their confidential sales and inventory data since they were interested in the forecasting project to enhance their business.

The first retailer (that we denote by S-1088 1, for confidentiality matters) , which represents the store ID) sell different furniture and accessories from designers, jewelers, artists and ceramicists. The second retailer (that we denote by S-1224) sell paper crafting products. Finally, the third retailer (that we denote by S-767) sell different types of footwear (shoes, trainers, etc.).

Sales and inventory data were provided in CSV format. Each sales CSV file corresponds to a retailer sales transactions data from the end of 2016 through September 2018, where each data record represents a time-stamped sale transaction (e.g. { OrderItemId: 127632, ProductVariantsId: 667975, Quantity: 4, Price: 4.99, CreationTime: 2016-12-28 22:10:51.000}). Table 3 shows the structure of the sales data file of each retailer that we have been provided with:

Attribute	Type	Description
OrderItemId	Long Integer	Unique sale-transaction identifier.
ProductVariantsId	Long Integer	Represents one version of a product with several options. Forecasts shall be generated for every product variant ID.
Quantity	Integer	Number of units that the customer purchased.
Price	Double	Purchase price of the product variant.
CreationTime	Date (y-m-d H:m:s)	A time-stamp indicating the time of the sale transaction.

Tableau 4 Structure of Stock&Buy historical sales data

In addition, we were provided with inventory data which contains product-related data such as the current quantity in stock, entry date, wholesale price, weight, product options, etc. Table 4 shows the structure of the inventory data file for each retailer. Each data record (line) in this file corresponds to a unique product variant (e.g. { ProducVariantsId: 550174, Quantity: 0, AllocatedInventory: 0, Cost: 5.43, MovingAverageCost: 5.60, PurchasePrice: 3.99, WholeSalePrice: 0, RetailPrice: 8.99, Sku: STOOBAM, ProductId: 200254, EntryDate: 2016-11-05 18:28:22.353 , OptionValues: "[\"Small\", \"Medium\"]", Barcode: NULL, IsBundle: 0, HasLots: 0, ImageId: NULL, Wight: 0, InventoryId: 1088}).

Attribute	Type	Description
ProducVariantsId	Long Integer	Represents a unique version of a product with several options. Forecasts shall be generated for every product variant ID.
Quantity	Double	The current quantity available in stock.
AllocatedInventory	Integer	It refers to the quantity of product existing in a warehouse, which is already allocated for specific uses, such as for a requisition, customer order, etc.
Cost	Double	The cost of the product variant.
MovingAverageCost	Double	Moving average cost of the product used for calculating

		Ending Inventory cost.
PurchasePrice	Double	Purchase price of the product variant.
WholeSalePrice	Double	The cost of a good sold by a wholesaler.
RetailPrice	Double	Purchase price of the product variant.
Sku	String	Stock Keeping Unit.
ProductId	Long Integer	ID of a product with several options. Each product variant belongs to a product family represented by a ProductId.
EntryDate	Date (y-m-d H:m:s)	Entry of the product in the warehouse.
OptionValues	String	Product options related to its color, size, etc.
Barcode	Long Integer	Product barcode.
ReferenceId	Long Integer	Product reference.
BundleName	String	The name of the product kit (if the product is sold as a kit).
IsBundle	Boolean	Indicates if the product is sold as a kit of many products from different categories.
HasLots	Boolean	Indicates if the product has lots.
ImageId	String	Indicates the ID of the image that is associated with the product (image uploaded by the retailer).
Wight	Double	Product weight.
InventoryId	Long Integer	ID of the inventory.

Tableau 5 Structure of Stock&Buy inventory data

b. Features Selection:

Determining potential predictors of demand that could be derived from this data was a collaborative process with our Stock&Buy client and based on a preliminary analysis that we conducted. After that, we decided to limit the forecasting model to the demand variable (quantity/number of purchased units) and the price variable, to forecast demand for the product variants which have sufficient historical sales data. When it comes to forecasting demand for product variants which have little to no historical sales data (new products), it has been decided to include other relevant inventory attributes (RetailPrice, Cost, PurchasePrice, WholeSalePrice, ProductId, IsBundle, HasLots, ImageId 2, Weight, OptionValues 3, etc.).

4.2 Data Exploration:

As part of the preliminary data exploration, we performed a stationarity test on sales data (which contains daily sales records of all product variants) of each retailer. The method we used is the Augmented Dickey-Fuller test (ADF). ADF tests the null hypothesis that a unit root is present in time series sample. ADF statistic is a negative number and the more negative it is the stronger the rejection of the hypothesis that there is a unit root.

There are two hypothesis in ADF tests:

- Null Hypothesis (H0): If accepted, it suggests the time series has a unit root, which means it is non-stationary.
- Alternate Hypothesis (H1): The null hypothesis is rejected; it suggests the time series does not have a unit root, meaning it is stationary.

The acceptance or rejection of the hypothesis H0 is based on the p-value 4 :

- If $p\text{-value} > \alpha$: Accept H0, the data has a unit root and is non-stationary.
- If $p\text{-value} \leq \alpha$: Reject H0. the data does not have a unit root and is stationary.

Such as α is the critical value (generally 0.01, 0.05 or 0.1). Table 4.5 shows the results of Dicker-Fuller test applied to our three retailers data (denoted by S-1088, S-1224, S-767).

The general regression equation which incorporates a constant and a linear trend is used and the t-statistic for a first order autoregressive coefficient equals one is computed.

	S-1088	S-1224	S-767
Dickey-Fuller	-18.665	-19.369	-22.254
p-value	0.01	0.01	0.01
A	0.05	0.05	0.05
Interpretation	Sales time series is trend-stationary	Sales time series is trend-stationary	Sales time series is trend-stationary

Tableau 6 Stationarity test using Augmented Dickey-Fuller Test on retailers sales data

The ADF test demonstrated that the three sales time-series are trend-stationary. Figure 2 shows the daily sales time-series which we used to perform the ADF test. The depicted sales data covers the period from 2016 through September 2018.

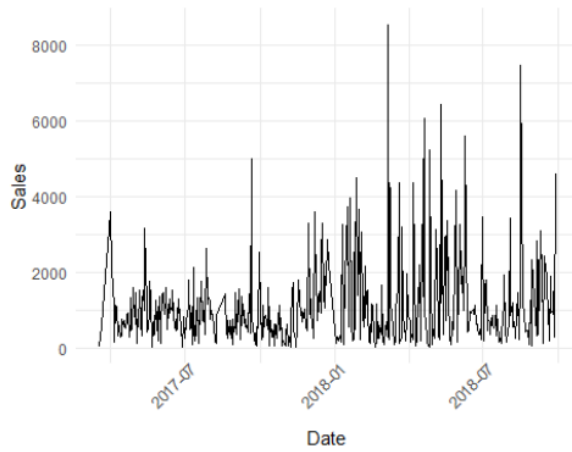
We also performed the ADF test on several demand time-series of different product variants (individual products time-series) sold by the three retailers and found different results. This is due to some product variants that are trend-stationary and some others that have a trend component.

4.3 Cross-Correlation Analysis:

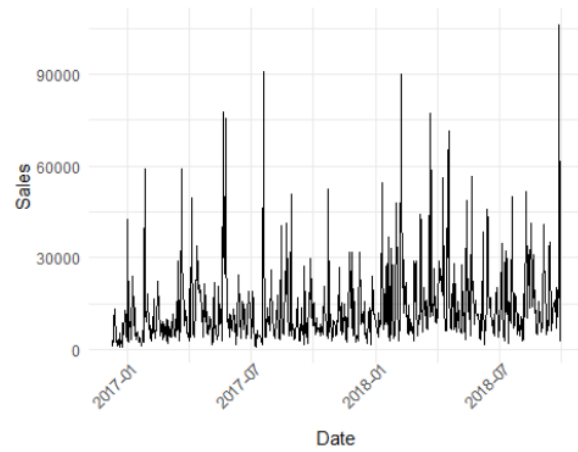
Cross correlation is the estimate of the degree to which two series are correlated. If we consider two series X and Y, the cross correlation between the two series means that the series Y may be related to past lags of the series X. The cross correlation r at delay d is defined as:

$$r = \frac{\sum_{i=1}^N ((x_i - m_x) \times (y_{i-d} - m_y))}{\sqrt{\sum_{i=1}^N (x_i - m_x)^2} \sqrt{\sum_{i=1}^N (y_{i-d} - m_y)^2}}$$

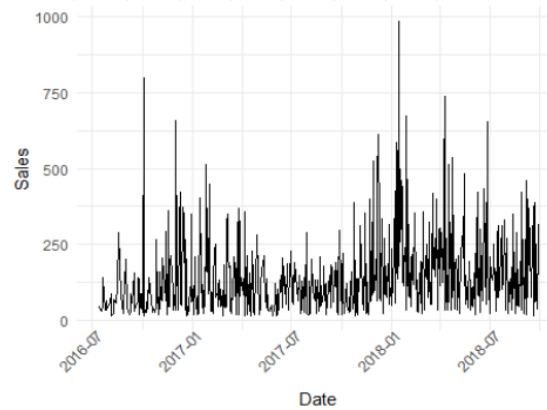
Where m_x and m_y are the means of the corresponding series.



(a) S-1088 daily sales



(b) S-1224 daily sales



(c) S-767 daily sales

Figure 23 Daily sales of the three retailers comprising all the product variants sold by each one of them

We were provided with data containing two time-dependant variables, the price and demand for all the product variants which were sold from 2016 through 2018. It is therefore interesting to study if there exists a correlation between the price and the demand. If a strong correlation exists between the two time-dependant variables, considering a multivariate time-series forecasting model such as the VAR model, may lead to more accurate forecasts than considering a univariate time-series forecasting model, which only takes into account the evolution of demand across time.

Since the retailers sell thousands of products, we only conducted the cross correlation analysis for some product variants to gain insights into the impact of price on demand, and whether this aspect shall be taken into account when building the solution. Figures 4.3, 4.4 and 4.5 show the results that we obtained by studying the cross correlation function of the demand and the price of some product variants sold by each of the three retailers respectively.

The results from the cross correlation analysis show that there is a strong correlation between the price and the demand, especially at lag 5 0 which corresponds for most of the product variants, to the highest correlation coefficient and represents the best fit between the two series.

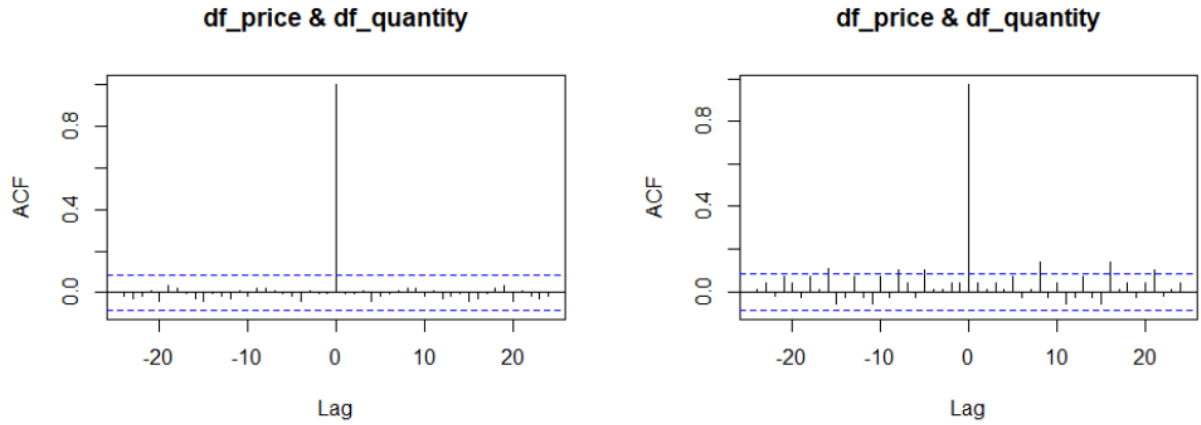


Figure 24 Cross Correlation analysis of some product variants from inventory S-1088

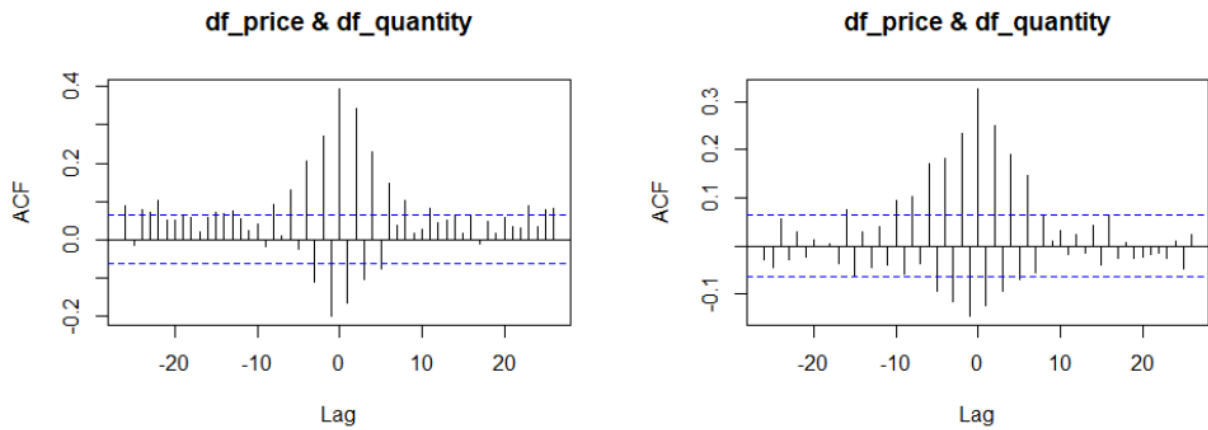


Figure 25 Cross Correlation analysis of some product variants from inventory S-1224

Moreover, we notice that for S-1224, there are significant positive and negative cross correlation coefficients (for example at lags -4, -3, -2, 0, 1, 2, 3, 4). This indicates that the price and demand can either move in the same direction (increase or steadiness in one will result in an increase or steadiness in the other) or move in two opposite directions (decrease in one will result in a increase in the other). However, S-1088 and S-767 time series, don't exhibit such patterns: the most significant cross correlation coefficient is positive and is obtained at lag 0. This means that the price and demand are strongly positively related. The explanation for that is the fact that the prices that these retailers use for their products are for the most part steady across time.

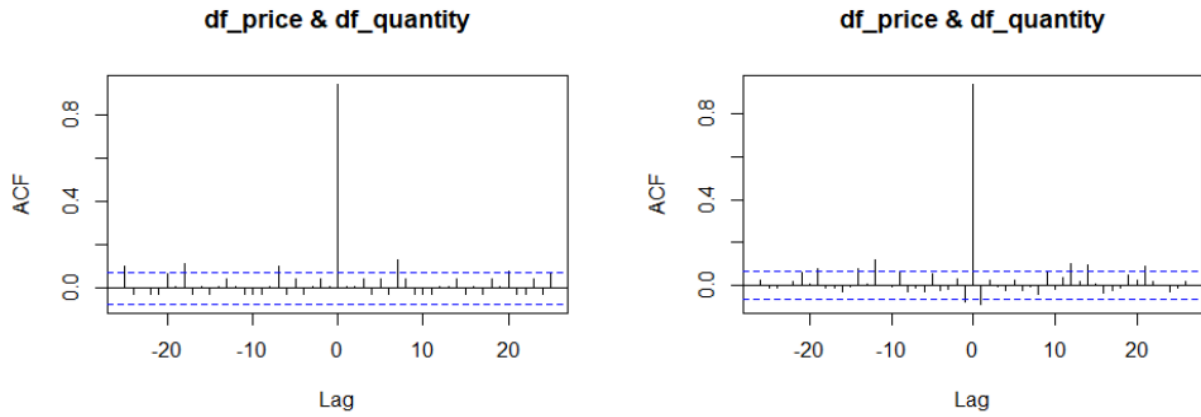


Figure 26 Cross Correlation analysis of some product variants from inventory S-767

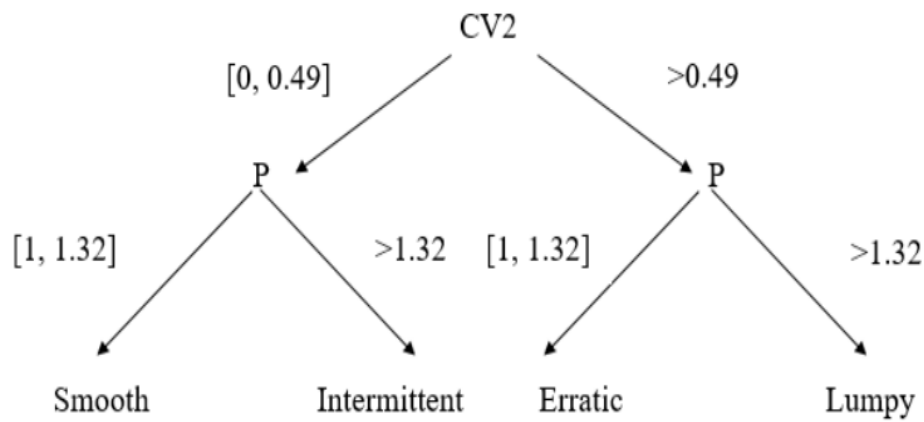


Figure 27 Syntetos Boylan classification decision tree

4.4 Demand Categorization for Inventory Items:

In this section, we will discuss the categorization of demand of the different inventory items that the three retailers who accepted to share their inventory and sales data sell.

In order to categorize demand, we used Syntetos and Boylan Categorization Scheme which is based on two main parameters : inter-demand interval (p) and squared coefficient of variation ($CV2$).

Figure 6 showcases SB decision tree.

This classification results in four quadrants each representing a particular demand pattern:

- If $CV2 > 0.49$ and $p > 1.32$ then the SKU item exhibits a lumpy demand which is characterized by a high demand variation and longer inter-demand interval. It is often considered as the most difficult demand pattern to forecast.
- If $0 \leq CV2 \leq 0.49$ and $p > 1.32$ the SKU item exhibits an intermittent demand which is characterized by a low demand variation but longer inter-demand interval.
- If $CV2 > 0.49$ and $1 \leq p \leq 1.32$ then the demand is erratic (characterized by a high demand variation but a shorter inter-demand interval).
- If $0 \leq CV2 \leq 0.49$ and $1 \leq p < 1.32$ then the demand is smooth (characterized by low demand variation and short inter-demand interval).

Table 6 summarizes the findings of the demand-categorization analysis that we conducted to determine patterns underlying demand based on SB classification. The number of product variants under each demand pattern are given in Table 4.6. Demand categorization step is crucial for choosing the appropriate forecasting method for such demand patterns. We conclude from Table 4.6 the following:

- The three retailers sold 1420, 2135, 489 product variants respectively (e.g. blue adidas, African carpets, etc.) from 2016 through September 2018, with different demand-patterns.
- The most dominant demand pattern for the majority of products was either lumpy demand or intermittent demand.
- Intermittent demand (characterized by weak randomness and long inter-demand intervals) is more prevalent within S-767 and S-1088.
- Lumpy demand (characterized by strong randomness) is more prevalent within S-1224.

	S-1088	S-1224	S-767
Total	1420	2135	489
Lumpy demand	48	1491	19
Intermittent demand	808	540	342
Erratic demand	1	16	0
Smooth demand	9	16	1
Other	554	72	127

Tableau 7 Demand categorization for the product variants across stores S-1088, S-1224 and S-767

Most product variants that the three retailers sell exhibit lumpy and intermittent demand, which indicates that the vast majority of these products are slow-moving products. Figure 7, 8, 9 show some examples of different product variants sold by the three retailers and exhibiting different demand patterns.

The intermittency of demand (that comes in the form of lumpy demand and intermittent demand) for the vast majority of product variants requires the selection of the appropriate forecasting models for these demand patterns in order to achieve the desirable forecasting accuracy.

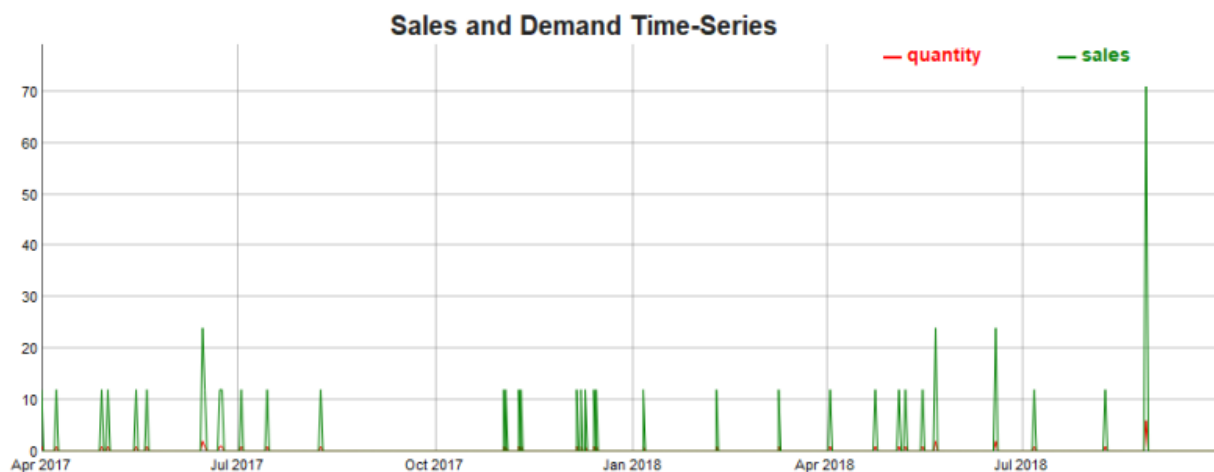


Figure 28 Example of a product variant from S-1088 and exhibiting intermittent demand

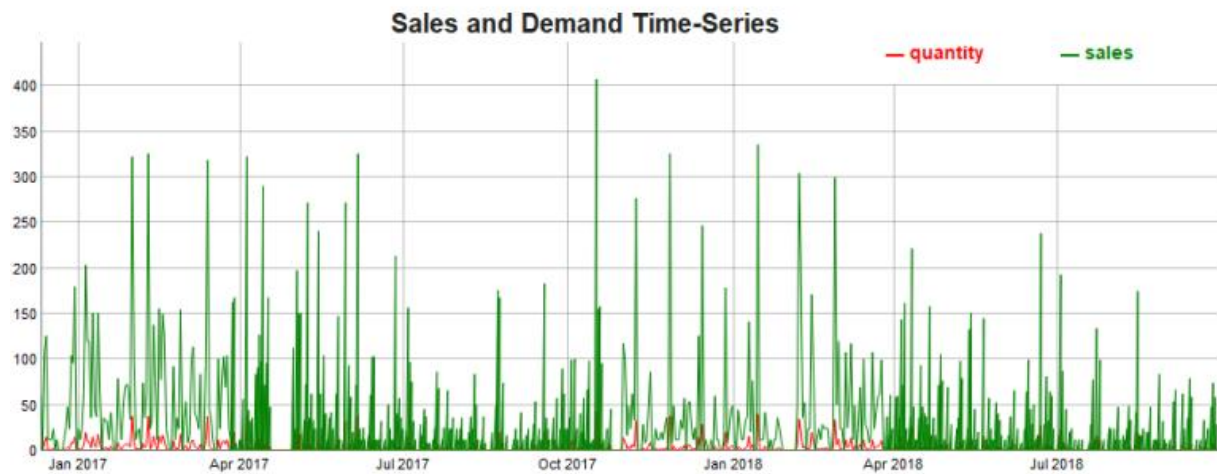


Figure 29 Example of a product variant from S-1224 and exhibiting lumpy demand

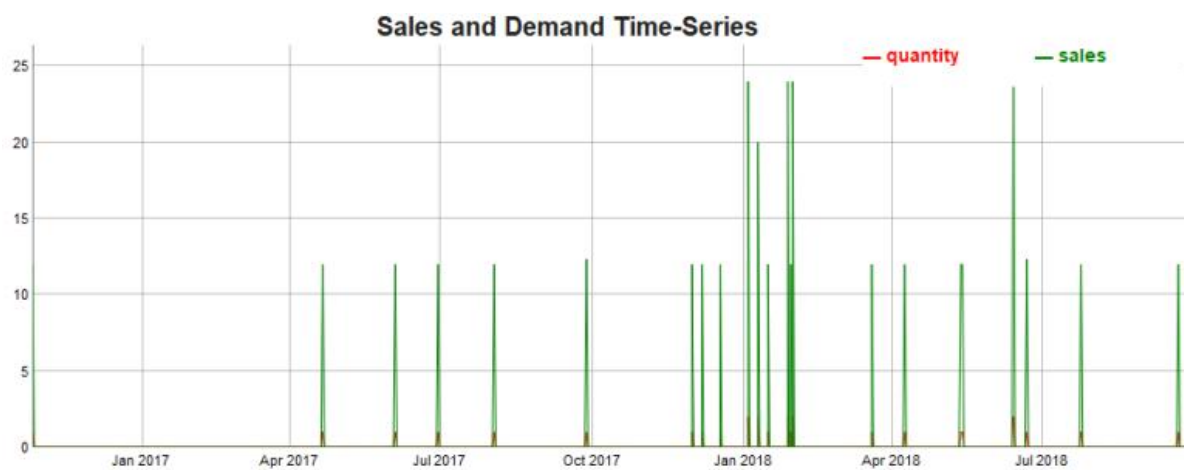


Figure 30 Example of a product variant from S-767 and exhibiting intermittent demand

Conclusion:

We discussed in this chapter a phase of our project which consists of analysing the problem through a preliminary data analysis. The latter mainly involved analysing cross correlation underlying demand time-series, and studying the different demand patterns exhibited by the product variants sold across three different stores of three retailers. We remind that these data only represent a proportion of the wealth of data available at Stock&Buy and doesn't cover all the retailers selling through the platform. It has been decided with the client to use these data to test and validate the final solution. Moreover, since retail business is very homogeneous, we expect retail data from other stores to be very similar to the ones we analysed.

This preliminary analysis showed that price and demand are strongly correlated. It also showed that many product variants sold by the different retailers exhibit intermittent and lumpy demand as defined by the Syntetos-Boylan categorization scheme. This lumpiness and intermittency need to be properly handled by the forecasting solution to obtain accurate forecasts.

General Conclusion :

Through this thesis entitled "forecasting tool for optimizing the supply chain", we reported on our end-of-study project. The success of any project depends on several parameters including the methodology. This is why we started our project with a bibliographic study of demand forecasting and the techniques applied to carry out this analysis (demand forecast). This study was guided at the same time by a needs analysis.

The second part is about the practical side of our job of creating a forecasting tool. We recall that the objective of our project is to provide a forecasting solution in the field of industry and more specifically the field of production. We presented throughout this thesis the work that we conducted with Stock&Buy, a Norwegian inventory management platform, servicing hundreds of customers all over the world. Stock&Buy main challenge is demand forecasting for thousands of products exhibiting different demand patterns. Indeed, the existing solution, which consists of a simple linear regression, doesn't provide customers with accurate demand forecasts for the products they sell. Another interesting research track was also to predict demand for products which don't have historical sales and demand data.

To address these issues, it was important to dive into the State of the Art related to demand forecasting, a field that has witnessed huge progress since 1950s with many forecasting methodologies being proposed to address the new challenges that the business environment faces today. It was also important to understand the nature of the data that we were provided with to propose a method that handles the patterns underlying these data. We remind that for this study, we were provided with sales and inventory data of three different retailers located in different countries who accepted to share with us their inventory and sales data from 2016 through September 2018. We summarize in what follows the work that we conducted during this project:

Preliminary analysis The preliminary data exploration and analysis was a crucial step since it allowed us to gain deeper insights into the data we were provided with. At this stage, we studied the cross correlation that may exist within sales time-series. More specifically, we studied the cross-correlation between the price and demand to determine if these two series were correlation. Studying cross correlation is often times useful since it would eventually guide us toward using a multivariate time-series model instead of a univariate model that only takes into account demand. At this stage, we also studied the different demand patterns that the products sold across the three stores exhibit. To categorize demand patterns, we used the Syntetos-Boylan classification scheme which categorizes demand either as smooth, erratic, lumpy or intermittent, based on the intermittent demand interval (p) and the squared demand variation coefficient (CV). This classification showed that an important proportion of the products sold across the stores exhibit intermittent and/or lumpy demand, which means that a lot of products can be described as slow-moving products.

Working with time-series data requires pre-processing the series to obtain accurate results.

Abstract

Supply chain planning is a critical activity in the supply chain management strategy. Having smart work tools to develop concrete plans is a necessity in today's business world. Demand forecasting plays an important role in many areas, including supply chain management and inventory management. In fact, inventory control and efficient inventory management are often based on forecasts that make it possible to estimate the demand, for weeks or even months in advance, and thus order the necessary amount.

We present our project with a growing online retail platform, Stock&Buy, servicing hundreds of customers all over the world, as an example of how retailers can use the wealth of data to optimize supply chain and inventory decisions through demand forecasting. One of the main challenges of Stock&Buy is estimating demand for thousands of products exhibiting different underlying demand patterns (seasonality, intermittency, etc.)

In practice, many quantitative and qualitative methods are used to forecast demand and minimize the forecasting error. Within this context, and in order to tackle Stock&Buy challenges, we provide a review of the theoretical framework and the state of the art related to demand forecasting, essentially time-series forecasting. We also propose a demand-forecasting tool to estimate demand for different product variants, which are, for the most part, associated with intermittent and lumpy demand patterns, whereby only a handful of distinct product units are sold every day.

Keywords

Supply Chain, Demand Forecasting, Time Series, Machine Learning, Statistical Methods

Resumé

La planification de la chaîne d'approvisionnement est une activité essentielle de la stratégie de gestion de la chaîne d'approvisionnement. Disposer d'outils de travail intelligents pour élaborer des plans concrets est une nécessité dans le monde des affaires d'aujourd'hui. La prévision de la demande joue un rôle important dans de nombreux domaines, y compris la chaîne d'approvisionnement, gestion et gestion des stocks. En effet, le contrôle des stocks et la gestion efficace des stocks reposent souvent sur des prévisions qui permettent d'estimer la demande, sur des semaines voire des mois à l'avance, et donc de commander la quantité nécessaire.

Nous présentons notre projet avec une plateforme de vente au détail en ligne en pleine croissance, Stock & Buy, desservant des centaines de clients partout dans le monde, comme un exemple de la manière dont les détaillants peuvent utiliser la richesse des données pour optimiser la chaîne d'approvisionnement et les décisions d'inventaire grâce à la prévision de la demande. L'un des principaux défis de Stock & Buy est d'estimer la demande pour des milliers de produits présentant différents modèles de demande sous-jacents (saisonnalité, intermittence, etc.)

En pratique, de nombreuses méthodes quantitatives et qualitatives sont utilisées pour prévoir la demande et minimiser l'erreur de prévision. Dans ce contexte, et afin de relever les défis Stock & Buy, nous passons en revue le cadre théorique et l'état de l'art lié à la prévision de la demande, essentiellement la prévision de séries chronologiques. Nous proposons également un outil de prévision de la demande pour estimer la demande pour différentes variantes de produits, qui sont, pour la plupart, associées à des modèles de demande intermittents et irréguliers, dans lesquels seule une poignée d'unités de produits distinctes est vendue chaque jour.

Mots clés

Chaîne d'approvisionnement, prévision de la demande, séries chronologiques, apprentissage automatique, méthodes statistiques .

يعد تخطيط سلسلة التوريد نشاطاً أساسياً لإستراتيجية إدارة سلسلة التوريد. يعد امتلاك أدوات عمل ذكية لوضع خطط ملموسة أمراً ضرورياً في عالم الأعمال اليوم يلعب التنبؤ بالطلب دوراً مهماً في العديد من المجالات بما في ذلك سلسلة التوريد وإدارة المخزون. في الواقع ، غالباً ما تعتمد مراقبة المخزون والإدارة الفعالة للمخزون على التنبؤات التي تجعل من الممكن تقدير الطلب قبل أسابيع أو حتى أشهر ، وبالتالي طلب الكمية اللازمة

، تخدم مئات العملاء في جميع Stock & Buy نقدم مشروعنا من خلال منصة بيع بالتجزئة متنامية عبر الإنترنت ، أنحاء العالم ، كمثال على كيفية استخدام تجار التجزئة لثروة البيانات لتحسين السلسلة. قرارات العرض والمخزون من في تقدير الطلب على آلاف المنتجات ذات Stock & Buy خلال التنبؤ بالطلب. يتمثل أحد التحديات الرئيسية في (أنماط الطلب الأساسية المختلفة) الموسمية ، والنقطع ، وما إلى ذلك

في الممارسة العملية ، يتم استخدام العديد من الأساليب الكمية والنوعية للتنبؤ بالطلب وتقليل أخطاء التنبؤ. في هذا السياق ، ومن أجل مواجهة تحديات الأسهم والشراء ، نقوم بمراجعة الإطار النظري وأحدث التطورات المتعلقة بالتنبؤ بالطلب ، وبشكل أساسي التنبؤ بالسلاسل الزمنية. نقدم أيضاً أداة التنبؤ بالطلب لتقدير الطلب على متغيرات المنتجات المختلفة ، والتي يرتبط معظمها بأنماط طلب متقطعة وغير منتظمة ، حيث يتم بيع عدد قليل فقط من وحدات المنتج المنفصلة. كل يوم

الكلمات الدالة

سلسلة التوريد ، التنبؤ بالطلب ، السلاسل الزمنية ، التعلم الآلي ، الأساليب الإحصائية

Bibliography

- [1] . Beamon, M . *Designing the green supply chain. Logistics information management*. 12. 1999. pp. 332-342.
- [2]. Comelli, M . *Modélisation, optimisation et simulation pour la planification tactique des chaînes*. 2008. pp. 21-22.
- [3] .Lemoine, D. *Generic models and resolution methods for tactical planning in mono-site and*. Université Blaise Pascal - Clermont-Ferrand II. 2008. pp. 21- 22.
- [4] .Christopher, M . *Logistics and supply chain management*. [éd.] Pearson UK. 2012.
- [5] . *Defining supply chain management*. J. T. M entzer, W. DeWitt, J. S. Keebler, S. M in, N. W. Nix, C. D. Smith, et Z. G. Zacharia. 2001, Journal of Business logistics, Vol. 22, pp. 1-25.
- [6] . R. K. Oliver et M . D. Webber. *Supply-chain management : logistics catches up with strategy*. 1982. pp. 42-47. Vol. 5.
- [7] . Ross, D. F. *Competing through supply chain management*. Springer Science & Business Media. 1998.
- [8] . *Purchasing and supply management : trends and changes throughout the 1990s*. R. J. Trent et R. M . M onczka. 3, 1998, International Journal of Purchasing and Materials Management, Vol. 34, pp. 2-11.
- [9] *Integrating the supply chain*. Stevens, G. C. 8, 3 1989, International Journal of Physical Distribution & materials management , Vol. 19.
- [10] . *Supply chain management : more than a new name*. M . C. Cooper, D. M . Lambert, et J. D. Pagh. 1, 1997, The international journal of logistics management, Vol. 8, pp. 1-14.
- [11]. Simchi-Levi, D. *Designing and managing the supply chain*. McGraw-Hill College. 2005.
- [12] . Christopher, M . *Logistics and supply chain management*. [éd.] Pearson UK. 2012.
- [13] . *From a literature review to a conceptual framework for sustainable*. Müller, S. Seuring et M . 15, 2008, Journal of cleaner production, Vol. 16, pp. 1699- 1710.
- [14] . H. Stadler et C. Kilger. *Supply chain management and advanced planning*. Berlin : s.n., 2000.
- [15] . *International supply chain management*. Houlihan, J. B. 1, 1985, International Journal of Physical, Vol. 15, pp. 22-38.
- [16] . *Using inventory for competitive advantage through supply chain*. T. C. Jones et D. W. Riley. 5, 1985, International Journal of Physical Distribution & Materials Management, Vol. 15, pp. 16-26.
- [17] . Fisher, M . L. *What is the right supply chain for your product ?* 1997.
- [18] . *Vendor-managed inventory in the retail supply chain*. M . Waller, M . E. Johnson, et T. Davis. 199, Journal of business logistics, Vol. 20, pp. 183-204.
- [19]. Rob J. Hyndman, George Athanasopoulos. *Forecasting: Principles and Practice*. [éd.] Otexts. 2013.
- [20]. DESPAGNE, Wilfried. *Construction, analyse et implémentation d'un modèle de prévision. Déploiement sous forme d'un système de prévision chez un opérateur européen du transport et de la logistique*. Université de Bretagne Sud. 2010. Mémoire de fin d'études.
- [21] . Jay Heizer, Barry Render, Chuck Munson. *Operations management*. 10th. 2014.
- [22] . *Forecasting Systems for Production and Inventory Control*. Fildes , Charles beard. 5, December 1992, International Journal of Operations & Production Management , Vol. 12, pp. 4-17.
- [23] . *The impact of information quality on supply chain performance*. Helena FORSLUND et Patrik JONSSON. 2007, International Journal of Operations & Production Management.
- [24] . Mark A. Moon, John T. Mentzer, Carlo D. Smith, and Michael S. Garver. *Seven Keys to Better Forecasting*. 1998. pp. 44-52.
- [25] . *The Impact of team-based forecasting*. Kenneth B. Kahn and John T. Mentzer. summer 1994, Journal of business forecasting, pp. 18-21.
- [26] . *Principles of electricity demand forecasting. Part 1: Methodologies*. AlAlawi, S. M ., & Islam, S. M . 3, 1996, Power engineering Journal , Vol. 10, pp. 139-143.
- [27] . Nitisha. Demand Forecasting: Concept, Significance, Objectives and Factors. <http://www.economicsdiscussion.net/demand-forecasting/demand-forecastingconcept-significance-objectives-and-factors/3557..> [En ligne]

- [28] . George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, Greta M. Ljung. *Time Series Analysis: Forecasting and Control*. 5th. 2015.
- [29] . Shumway, Robert H. *Time series analysis and its applications*. 4th. 2000.
- [30]. AN ANALYSIS OF THE LIMITING METHOD. J. Brown V. R. Cane. 1959, British Journal of Statistical Psychology banner.
- [31]. Holt, C.C. *Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages*. ONR Memorandum. 1957. Vol. 52.
- [32]. *Forecasting Sales by Exponentially Weighted Moving Averages*. Winters, Peter R. 3, April 1960, Journal, Vol. 6, pp. 324-342.
- [33]. *Modeling and forecasting vehicular traffic flow as a seasonal ARIMA*. Williams, B. M. 6, 2003, Journal of Transportation Engineering, Vol. 129, pp. 664-672.
- [34]. *Using Holt Winter's Multiplicative Model to Forecast Assisted*. Odame Owired, Emmanuel, Atinuke Adebajji. 9, 2014, Journal of Biology, Agriculture and Healthcare, Vol. 4, pp. 83-88.
- [35]. Box, G. and Jenkins, G. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco. : s.n., 1970.
- [36]. *Support Vector Regression to Forecast the Demand and*. V. Anandhi and R. Manicka Chezian. 3, June 2013, International Journal of Future Computer and Communication, Vol. 2, pp. 266-269.
- [37]. *A Prediction Approach for Demand Analysis of Energy*. Fazli Wahid, DoHyeun Kim. 2, 2016, International Journal of Smart Home, Vol. 10, pp. 97-108.
- [38]. *Electric Power Load Forecast Using Decision Tree Algorithms*. 4, Hambali, M., Akinyemi, A., Oladunjoye, J. & Yusuf, N : s.n., December 2016, Computing, Information Systems, Development Informatics & Allied Research Journal, Vol. 7, pp. 29-42.
- [39]. *K-Means Clustering Algorithm Applications in Data Mining and Pattern Recognition*. Huda Hamdan Ali, Lubna Emad Kadhum. 8, August 2017, International Journal of Science and Research (IJSR), Vol. 6, pp. 1577-1584.
- [40]. AJLA ELM ASDOTTER, CARL NYSTRÖM ER. *A comparative study between LSTM and ARIMA for sales forecasting in retail*. SCHOOL OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE, KTH ROYAL INSTITUTE OF TECHNOLOGY. 2018.
- [41]. Hyndman, R.J., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., Yasmine, F. *Forecasting functions for time series and linear models. R package version*.
- [42]. Armstrong, J. Scott. *Principles of Forecasting: A Handbook for Researchers and Practitioners*. 2001. p. 384.
- [43]. *Another look at measures of forecast accuracy*. Rob J Hyndman, Anne B Koehler. 4, October–December 2006, International Journal of Forecasting, Vol. 22, pp. 679-688.
- [44]. Min, Hokey (2010) 'Artificial intelligence in supply chain management: theory and applications', International Journal of Logistics Research and Applications, 13: 1, 13 — 39, First published on: 24 March 2009 (iFirst)
To link to this Article: DOI: 10.1080/13675560902736537
URL: <http://dx.doi.org/10.1080/13675560902736537>
- [45] Kostenko, A. V. & Hyndman, R. J. (2006). A note on the categorization of demand patterns.
- [46] Kourentzes, N. (2013). Intermittent demand forecasts with neural networks. International Journal of Production Economics, 198-206, 2013. Available: 10.1016/j.ijpe.2013.01.009.
- [47] Kourentzes, N., Petropoulos, F. (2019). tsintermittent: Intermittent Time Series Forecasting. R package.
- [48] <http://faculty.smu.edu/tfomby/eco6375/BJ%20Notes/ADF%20Notes.pdf>

