



**UNIVERSITAT ROVIRA I VIRGILI (URV) Y UNIVERSITAT OBERTA DE CATALUNYA (UOC)**

**MASTER IN COMPUTATIONAL AND MATHEMATICAL ENGINEERING**

**FINAL MASTER PROJECT (FMP)**

**AREA: MODELLING AND SIMULATION**

# **Analytical Methods for Logistic Tenders**

**Time Series Winters' Method Forecast with RStudio**

Author: Cristina Álvarez Benítez

Tutor: Alberto García Villoria

15<sup>th</sup> September 2019

Dr./Dra. (name) Alberto García Villoria, certifies that the student Cristina Álvarez Benítez has elaborated the work under his/her direction and he/she authorizes the presentation of this memory for its evaluation.

Director's signature:

A handwritten signature in blue ink, consisting of a stylized 'A' followed by a long, sweeping horizontal line that curves upwards at the end.



This work is subject to a licence of Recognition-  
NonCommercial- NoDerivs 3.0 Creative Commons

## Alternative licences

(choose any of the following and substitute the one of the previous page)

To) Creative Commons:



This work is subject to a licence of Attribution-NonCommercial-NoDerivs 3.0 of Creative Commons



This work is subject to a licence of Attribution-NonCommercial-ShareAlike 3.0 of Creative Commons



This work is subject to a licence of Attribution-NonCommercial 3.0 of Creative Commons



This work is subject to a licence of Attribution- NoDerivs 3.0 of Creative Commons



This work is subject to a licence of Attribution-ShareAlike 3.0 of Creative Commons



This work is subject to a licence of Attribution 3.0 of Creative Commons

## **B) GNU Free Documentation License (GNU FDL)**

Copyright © YEAR YOUR-NAME.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no **\*\*Back-Cover Texts**.

To copy of the license is included in the section **\*\*entitled "GNU Free Documentation License"**.

## **C) Copyright**

© (The author/to)

Reserved all the rights. It is forbidden the total or partial reproduction of this work by any half or procedure, comprised the impression, the reprography, the microfilm, the computer treatment or any another system, as well as the distribution of copies by means of rent and loan, without the permission written of the author or of the limits that authorise the Law of Copyright.

# FINAL PROJECT SHEET

<b>Title:</b>	Analytical Methods for Logistic Tenders, Time Series Winters' Method Forecast with RStudio
<b>Author:</b>	Cristina Álvarez Benítez
<b>Tutor/a:</b>	Alberto García Villoria
<b>Date</b> (mm/yyyy):	09/2019
<b>Program:</b>	Master in Computational and Mathematical Engineering
<b>Area:</b>	Modelling and Simulation
<b>Language:</b>	English
<b>Key words:</b>	Tender Management, Forecast, Time Series Analysis

## Quote

"Prediction is very difficult, especially if it's about the future."

- Nils Bohr, Nobel laureate in Physics

# Acknowledgments

I would like to acknowledge the following people and institutions for their collaboration on this project, since without them it would not have been possible to realize it.

First of all, Alberto García Villoria has been the ideal Final Master Project supervisor. His insightful remarks aided the writing of this thesis. On the most technical side, Dr. Angel A. Juan from the IN3 investigation group ISCO, provided the needed sage advice and also the motivation and encouragement to fulfil this academic stage.

I would also like to acknowledge Sachin Nataraj, colleague in Tender Management Air & Ocean Europe department, closely involved in all the projects. Also to the rest of the people inside the logistic company that have contributed with knowledge and resources.

Finally, a mention to Xènia Casassas for the introduction to the topic and Lluna Sada for the collaboration in the paper based on the project.





# Abstract

The concept of reverse auction or tender is gaining importance in the logistics and transportation industry. The number of services contracted through this process is increasing, and many provider companies need to provide competitive prices to gain them.

A first part of the approach details different projects to improve organizationally the tender process. A second part of the approach presents the complete structure of the *Smart Tenders Management with Artificial Intelligence* project, which is divided in four stages; (1) Automation of Transport Cost Analysis, (2) Predicting Cost Evolution and Establishing Long-term Tariffs, (3) Modelling Success Probabilities in Tenders and (4) Optimization of the Tenders Portfolio.

The main goal of this project focuses on the second stage of the *Smart Tenders Management with Artificial Intelligence* project. That is, the development of a methodology with few parameters (alpha, beta and gamma), which are automatically adjusted, based on time series and more concretely Winters' Method. It allows to calculate *low-risk* forecasted intervals, giving the decision-maker information regarding which is the most adequate long-term tariff to offer in the tender.

For the experimentation with RStudio, ocean freight rates from Rotterdam to Shanghai during 7 years are used, from 2012 to 2018. Finally, after reading the data, fitting and testing the model, the calculation for the *low-risk* interval for the data used is obtained.

**Keywords:** Final Master Project, Logistic Service Provider, Tender Management, Forecast, Time Series Analysis, Winters' Method

# Index

1. Introduction.....	10
1.1 Context and justification of the Work.....	10
1.2 Aims of the Work.....	11
1.2.1 Global objective .....	11
1.2.2 Partial objectives .....	11
1.3 Approach and method followed .....	13
1.4 Planning of the Work .....	15
1.5 Brief summary of products obtained.....	16
1.6 Brief description of the others chapters of the memory.....	16
2. Literature Review .....	18
2.1 Introduction .....	18
2.2 Literature by project stages .....	19
2.2 Summary of literature review.....	22
3. Formal description of the problem .....	25
3.1 Logistics and logistic providers .....	25
3.1.1 3PL and 4PL vs. freight broker .....	26
3.1.2 Logistic company structure .....	27
3.2 Spot quotation vs reverse auction .....	28
3.2.1 Advantages of reverse auctions.....	29
3.2.2 Disadvantages of reverse auctions .....	29
3.2.3 Challenges.....	29
3.3 Forecasting ocean freight rates .....	30
4. Our approach.....	32
4.1 Implementation Plan .....	32
4.1.1 First steps .....	32
4.1.1.1 Standard Operating Procedure.....	32
4.1.1.2 Distribution Process .....	33
4.1.1.3 Tender Management Process Flow .....	35
4.1.2 Rating automation .....	36
4.1.3 Reporting and analysis .....	37
4.1.3.1 Sugar CRM reporting and analysis .....	37
4.1.3.2 QlikSense tender analysis.....	38
4.1.4 Scoring system .....	39
4.1.5 Database creation for RFIs.....	40
4.2 Smart Tenders Management with Artificial Intelligence .....	41
4.2.1 Project overview.....	41
4.2.2 Predicting Cost Evolution and Establishing Long-term Tariffs using Time Series.....	42
4.2.2.1 Goals .....	42
4.2.2.2 Use cases .....	43
4.2.2.3 Process .....	43
4.2.2.4 Data and components .....	44
4.2.3 Winters' Method .....	45
4.2.3.1 Technique justification.....	45
4.2.3.2 Winters' method definition .....	48
4.2.3.3 Method equations .....	49
4.2.3.4 Method evaluation, error measures .....	50
4.2.4 Modelling Success Probabilities in Tenders.....	51
4.2.4.1 Introduction .....	51
4.2.4.2 Data .....	52
4.2.4.3 Results .....	53
5. Computational experiments and analysis of the results.....	55
5.1 Introduction .....	55
5.2 Reading Time Series Data.....	56
5.3 Plotting Time Series.....	57
5.4 Decomposing Time Series.....	60
5.4.1 Decomposing non-seasonal data .....	60
5.4.2 Decomposing seasonal data.....	60

5.4.3 Seasonally adjusting .....	65
5.5 Forecast using Exponential Smoothing .....	66
5.5.1 Simple Exponential Smoothing .....	66
5.5.2 Holt's Exponential Smoothing .....	67
5.5.3 Holt-Winters' Exponential Smoothing .....	67
5.5.3.1 Verification of the application of the Winters' method.....	68
5.5.3.2 Model fit.....	70
5.5.3.3 Forecast .....	72
5.5.3.4 Confidence intervals.....	74
5.5.3.5 Finding the best fit of the model .....	77
5.5.3.6 Ljung-Box test.....	81
5.6 Establishing low-risk intervals for long-term tariffs for tenders .....	84
6. Conclusions.....	87
7. Self-assessment .....	89
8. Future work .....	91
9. Glossary .....	92
10. Bibliography .....	93
11. Annexes.....	96

## List of figures and tables

Figure 1.1. A scheme of a general RFQ process .....	10
Figure 1.2. Historical data (circles), model (squares), and forecasted values (triangles) ....	12
Figure 3.1. A scheme of a general RFQ process .....	28
Figure 4.1. Distribution Process Flow for RFQ/RFI .....	34
Figure 4.2. Tender Management Process Flow .....	36
Figure 4.3. Sugar CRM, Freight Tender Management Module, Basic tab .....	37
Figure 4.4. QlikSense application for tender analysis, participation and status .....	38
Figure 4.5. QlikSense application for tender analysis, clients and volumes .....	39
Figure 4.6. Tender Scoring System from Road Division, general template and example ....	40
Figure 4.7. RFI database in Wiki page of Air & Ocean Freight Tender Management .....	41
Figure 4.8. Structure of the STeM-AI project .....	42
Figure 4.9. Exponential smoothing methods graphics .....	47
Figure 4.10. Example of Winters' method with additive and multiplicative seasonality .....	48
Figure 4.11. Relation between the margin applied and the success of tenders .....	53
Figure 5.1. Structure of the STeM-AI project, stage 2 .....	55
Figure 5.2. Ocean freight costs real data .....	58
Figure 5.3. Ocean freight costs artificially generated, excel .....	59
Figure 5.4. Decomposition of multiplicative time series .....	64
Figure 5.5. Ocean freight time series seasonally adjusted .....	66
Figure 5.6. Ocean freight forecast using Holt Winters' Method 2018-2019 .....	68
Figure 5.7. Ocean freight Holt-Winters' fitting .....	71
Figure 5.8. Ocean freight forecast 2019-2020 using Holt Winters' Method .....	73
Figure 5.9. Winters' Method forecast using 7 periods .....	75
Figure 5.10. Winters' Method forecast using 14 periods .....	76
Figure 5.11. Winters' Method forecast using 21 periods .....	76
Figure 5.12. Winters' Method forecast using 28 periods .....	77
Figure 5.13. Forecast graph (alpha = 0.4, beta = 0.1, gamma = 0.3) .....	78
Figure 5.14. Forecast graph (alpha = 0.9, beta = 0.1, gamma = 0.3) .....	79
Figure 5.15. Forecast graph (alpha = 0.4, beta = 0.9, gamma = 0.3) .....	79
Figure 5.16. Forecast graph (alpha = 0.4, beta = 0.1, gamma = 0.9) .....	80
Figure 5.17. Forecast graph (alpha = 0.9, beta = 0.1, gamma = 0.9) .....	80
Figure 5.18. Forecast graph RStudio optimized parameters .....	81
Figure 5.19. Ljung-Box test, ocean freight forecast residuals .....	82
Figure 5.20. Time plot of the forecast errors .....	83
Figure 5.21. Histogram of the forecast errors .....	84
Figure 5.22. Ocean Freight forecast with the corresponding low-risk interval .....	86
Table 1.1. Chronological planning of the Work .....	15
Table 4.1. Exponential smoothing methods definition and prediction .....	69
Table 4.2. Logistic regression model to estimate the probability of success .....	47
Table 5.1. Forecast and confidence intervals 2018 .....	54
Table 5.2. Forecast and confidence intervals 2019 .....	73
Table 5.3. Confidence intervals of the forecasts for different sizes of data sets .....	74
Table 5.4. Measure errors for different parameter combinations .....	78



# 1. Introduction

The Final Master Project on ocean freight rates forecasting is introduced in Section 1. In this first section the work will be contextualized and justified and the objectives, methodology and planning detailed. Finally, the results and other sections on the project will be briefly summarized.

## 1.1 Context and justification of the Work

In recent years the world of industry has grown exponentially and, in turn, the need to transport their products. In addition, we live in a trend of globalization that generates wider markets and therefore with more competitors. In the last decades the selection of suppliers has been developed through a contest in which the transport companies compete providing their best offers. These contests receive the name of Tender or RFQ (Request For Quotation) which, as its name indicates, is a price request for a specific package of transport services.

Although there is no single way to carry out this process because each client requests according to their own needs, there are a set of common steps to follow (Figure 1.1):

- The requester sends the set of services for which the price is asked.
- The supplier answers the request with a price proposal.
- The client provides a feedback (position with respect to other companies).
  - If the requester gives another opportunity to improve prices, we return to the second point.
  - If not, in case of reaching an agreement, the contract is signed. Otherwise, the proposal is rejected.

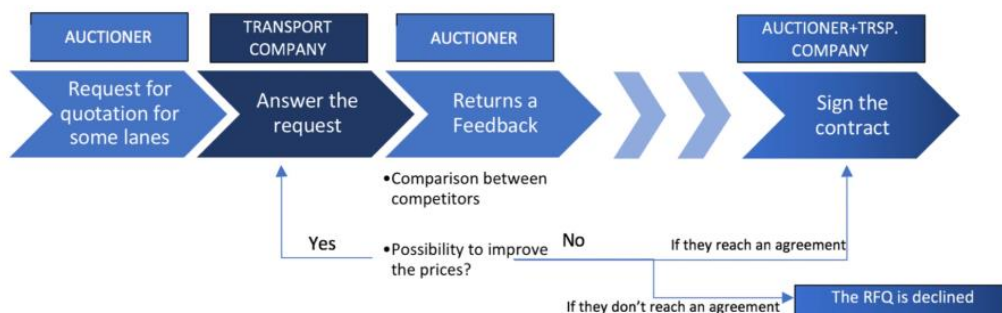


Figure 1.1. A scheme of a general RFQ process

The changing dynamics of the business world are making logistics and transportation companies to modify their strategies to achieve maximum results. The position inside a logistics and transportation company in charge of the process is a tender coordinator. This person takes care of the whole structured process starting from an invitation to tender to designing of an effective tender response. The different tasks of a tender coordinator will be further detailed in following sections of the master thesis. These tasks are linked to different challenges which justify the project.

Between the different tasks and projects to improve the tender process and organization, the focus of the Final Master Project will be put on the ocean freight rates forecasting. The transport costs extracted from the tools in a logistic company are only valid for short-term periods; hence they cannot be used to estimate the evolution of costs throughout longer periods; e.g., one year or even more as requested in tenders. Without reliable cost estimation on the long run, it is not possible to provide *risk-aware* fare tariffs for tender calls.

## 1.2 Aims of the Work

### 1.2.1 Global objective

To generate 'risk-aware' fare tariffs that can be valid for one year or more, and thus can be employed in tender calls.

### 1.2.2 Partial objectives

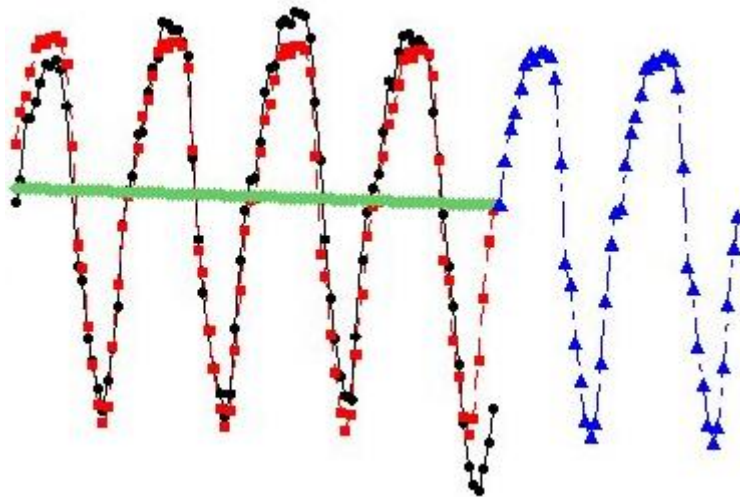
In addition to the main objective exposed, there are previous and posterior steps. They can be translated into sub-objectives which must be reached, defining therefore an efficient methodology to give a solution to the raised problem. Below, the list of partial objectives is presented:

- **Study of the current bibliography related to the theme.** This will be the first formal approach to the subject and the starting point of the way. Knowing other investigations and clarifying some ideas will allow a better definition and cognition of the theme to finally choose a perspective to work on it.
- **Meeting professionals in the logistic industry and align with researchers.** The analysis of the literature covers tender process and optimization methodologies. However, the alignment between the needs of professionals in a logistic company and forecasting methodologies used by researchers aims to allow the development and



implementation of new solutions on tender process to be efficient and hence competitive.

- **Strict definition of the inputs involved in the project.** The achievement of the two first sub-objectives listed will allow having a precise idea of the information and regulations involved in tender management. In other words, which information the algorithm will need to give a powerful and applicable solution.
- **To develop models, based on historical data, that can forecast the time evolution of transport costs associated with legs.** These models should be able to consider both trend and seasonality factors (Figure 1.2). These models will be then implemented as Python, R, or Java applications.
- **Using the previous forecasts, to develop models that can provide 'risk-aware' intervals for the values of the long-term fare prices to be used in calls for tenders.** In other words, these intervals will provide decision makers a 'low-risk' range in which long-term fare prices can be established. Obviously, the closer the price to the lower bound of the interval the more competitive the price will be but the more risk (inside the 'safety' zone) will be incurred, and vice versa.



*Figure 1.2. Historical data (circles), model (squares), and forecasted values (triangles)*

- **Design and execution of a series of experiments.** This last sub-objective is meant to identify and quantify the contribution of the project. To do it, historical data will be used to forecast next periods prices; but still past data, so that the accuracy of the methodology can be checked.

### 1.3 Approach and method followed

Once the objectives are fixed, it is essential to clarify the adequate processes to achieve them.

Firstly, as most projects, a wide documentation at the beginning of the work will have its results during the execution of the project. Before deepening on the subject and the programming side, it will be necessary to have a theoretical background. A bibliographical research on the topic will also show if models to treat the problem have been developed and the gaps between them and the real challenges in the industry.

To find all the documentation available, public electronic resources will be used as well as the documents provided by the own Final Master Project Directors Alberto García Villoria together with Angel A. Juan for the more technical part. Universitat Oberta de Catalunya library and other public libraries will also be visited. The material that is expected to find are articles from universities or other research centres, master thesis and different kind of technical reports. Another useful material in terms of project structuration, format will be some Final Master Projects and Ph.D. from other UAB and UOC students which projects were also directed by Angel A. Juan.

Moreover, there is a direct approach on the logistic industry thanks to a job in Tender Management for Air & Ocean European and Global bids in a logistic company since September 2018. The different tasks of a tender coordinator will be further detailed in following sections of the master thesis. These tasks are linked to different challenges which justify the project and will conduct itself the **first part** of the project.

A **second part** of the Final Master Project will be the programming side to reach the global objective of generating *risk-aware* fare tariffs that can be valid for one year or more, and thus can be employed in tender calls, divided in the following points and different strategies for each of them:

1. **To make sure enough historical data is available** (the more the better) to be able to generate reliable forecasts on the evolution of transport costs. After an initial analysis of the ocean freight rates data provided by the logistic company (**Figure 5.2, page 58**), it was decided to use artificial data (**Figure 5.3, page 59**) to overcome the problem of not having sufficient historical data. Additionally, the real data includes a sharp increase due to several factors which would hinder the method to provide accurate forecasts.

2. [Development of Python, R, or Java applications that allow to automate the use of time series analysis](#) (e.g., Winter's method or ARIMA models) in order to forecast the future evolution of transport costs for each leg and transport mode. After an analysis of different methods in the sub-section 5.5, Winter's method is used and ARIMA models are mentioned as future work.
3. [Development of a computation method, based on the forecasted cost values, that can generate the \*low-risk\* interval](#) for the decision maker to establish the fare tariff. In this point of the methodology also different strategies could be followed, and the details of the generation process and therefore values obtained is explained in the sub-section 5.6.
4. [Partial validation of the pilot project results \(forecasted values\)](#) from the Tender Manager position and other stakeholders who have the knowledge about the cost evolution inside the logistic company.
5. [In future stages, this method will then be implemented](#) as a Python / Java application, so it can be easily used by tender managers at the logistic company.
6. [Final validation of the pilot project](#) from the Tender Manager position and other stakeholders who are involved in tender the process inside the logistic company.
7. [Technical and easy-to-understand documentation](#) will be provided once the validation of the final application is done, so that company colleagues can use the developed modules.

During the development of both parts of the project, there will be a set of deliveries and personal or online meetings every 15 days, as a maximum, with the director of the project Ángel A. Juan. Additionally, emails commenting and showing the progress will be sent at the deadlines set in the next point of the section to the academical director of the project Alberto García Villoria. Furthermore, there will be a personal contact with workgroups inside the company who will be the main link between the university and current tools used and in development inside the logistic company.

To sum up, the methodology followed aims to overcome the gaps between the researches on the tender process and forecasting methods to be linked taking into account the complexity that have to handle logistic companies in real-life. The close vision of the challenges and limitations can only be seen from the insides of a company; receiving requests and coordinating tenders until its final compilation.

## 1.4 Planning of the Work

	TASK	DURATION (days)	START	END
<b>1</b>	<b>INTRODUCTION</b>	<b>16</b>	<b>16.03.19</b>	<b>31.03.19</b>
1.1	Context and justification of the Work			
1.2	Aims of the Work			
1.3	Approach and method followed			
1.4	Planning of the work			
1.5	Brief summary of products obtained			
1.6	Brief description of the other chapters of the memory			
<b>2</b>	<b>LITERATURE REVIEW</b>	<b>15</b>	<b>01.03.19</b>	<b>15.03.19</b>
2.1	Introduction			
2.2	Literature by project stages			
2.3	Summary of literature review			
<b>3</b>	<b>FORMAL DESCRIPTION OF THE PROBLEM</b>	<b>30</b>	<b>01.04.19</b>	<b>30.04.19</b>
3.1	Logistics and logistic providers			
3.2	Spot quotation vs reverse auction			
3.3	Forecasting ocean freight rates			
<b>4</b>	<b>OUR APPROACH</b>	<b>46</b>	<b>01.05.19</b>	<b>15.06.19</b>
4.1	Implementation Plan			
4.2	Smart Tenders Management with Artificial Intelligence			
<b>5</b>	<b>COMPUTATIONAL EXPERIMENTS</b>	<b>61</b>	<b>16.06.19</b>	<b>15.08.19</b>
5.1	Introduction			
5.2	Reading Time Series Data			
5.3	Plotting Time Series			
5.4	Decomposing Time Series			
5.5	Forecast using Exponential Smoothing			
5.6	Establishing low-risk intervals for long-term tariffs for tenders			
<b>6</b>	<b>CONCLUSIONS AND FUTURE WORK</b>	<b>16</b>	<b>16.08.19</b>	<b>15.09.19</b>

*Table 1.1. Chronological order of the Work*

## 1.5 Brief summary of products obtained

This sub-section summarizes the products obtained during the experimentation section. Initially, the real ocean freight rates data is read and plotted (Figure 5.2, page 58) and finally decided to use artificial data (Figure 5.3, page 59) from an example in Hanke et al (2001), as explained during the methodology and more in detail during the experimentation.

Different parameter combinations are analysed (Table 5.4, page 78) and the RStudio optimized option is used to fit the model (Figure 5.20, page 83). The parameters alpha, beta and gamma for Winters' method are the smoothing constants for the level, trend and seasonality. RStudio has the possibility to optimize these parameters by minimizing the squared error of the forecast, which are also shown in Table 5.4 (page 78) as alpha 0,7994528; beta 0,0796615; gamma 0,9536164.

Finally, the low-risk confidence interval for 1-year validity is obtained (Figure 5.22, page 86) for a concrete origin and destination ocean freight and the procedure explained in the subsection 5.6. The minimum value of the interval is calculated with the average of the forecasted values. The maximum value of the interval is calculated as the average of the values that define the upper bound of the 95% confidence intervals associated to each forecasted value. Concretely, for the data used in this Final Master Project, the minimum value of the interval is 590.2019 and the maximum is 777.0406

## 1.6 Brief description of the others chapters of the memory

In this Section 1. Introduction the Final Master Project on ocean freight rates forecasting is introduced. The work is contextualized and justified and the objectives, methodology and planning detailed. The Section 2. Literature review uses the stages structure of the whole project to analyse related literature. Finally, a summary of the literature review highlights the most important findings. The stages of the Smart Tenders Management with Artificial Intelligence project will be defined and developed in detail in our approach to solve the problem in Section 4.

The problem of rating forecasting for ocean freight rates is described in detail in Section 3. Formal description of the problem. To do so, there are three subsections which go deeper into the subject step by step. Firstly, the term logistics is defined and also what logistic providers does or which is their structure. Secondly, it is explained what a tender or reverse auction is in comparison to spot quotations. Thirdly and finally, the section focuses on rating forecasting ocean freight rates, which is a challenge for tender

managers.

Two sub-sections can be found in Section 4. [Our approach](#); a general Implementation Plan to deal with different challenges faced on Tender Management department inside a logistic company and Smart Tenders Management with Artificial Intelligence project, which includes ocean freight rates forecasting using Winters' Method in RStudio. The second sub-section treats in detail the second stage of the project: ocean freight rates forecasting using Winters' Method in RStudio; and the third stage: Modelling Success Probabilities in Tenders.

The Smart Tenders Management with Artificial Intelligence project will be treated in [Section 5. Computational experiments](#), focusing on the second stage: ocean freight rates forecasting using Holt Winters' Method in RStudio. The steps will go from the reading of the data to the creation of a low-risk interval for ocean freight rates forecast, passing through plotting and decomposing the data, evaluating different forecasting methods and verifying, fitting, forecasting and analysing the chosen method; Holt Winters' Exponential Smoothing.

To sum-up, [Section 6. Conclusion and future work](#) includes a description of the conclusions of the work and lessons learned. Additionally, a critical reflection on the attainment of the aims initially posted and a critical analysis of the follow-up of the planning and methodology along the development of the project. It also highlights the lines of future work that have not been able to explore in this work and have remained slopes. Finally, [Section 7. Glossary](#) defines the terms and acronyms more notable used inside the Memory and [Section 8. Bibliography](#) quotes all the references used.

## 2. Literature review

This Section 2 uses the stages structure of the whole project to analyse related literature. Finally, a summary of the Literature Review highlights the most important findings. The stages of the Smart Tenders Management with Artificial Intelligence project will be defined and developed in detail in our approach to solve the problem in Section 4.

### 2.1 Introduction

One of the first steps for the development of the project is the review of the existing literature, the theoretical background that will support the Work and introduce the contribution. Next, the Literature Review presents as a theoretical base for the rest of the Final Master Project.

Initially, while working as a Tender Manager and even before the start of the Final Master Project, the research began with the current procedures used to improve the tender process. The research led to companies (such as *Tender Tool* or *Tendereasy*), focused on providing services to logistic companies to improve their processes of tariffs fixation for tenders, but specially to manage in a more coordinated way the general process or having a better understanding of the market and their own information. Some of these service companies show papers in their web portals regarding their products and services, but no information has been found regarding the analytical methods they use to provide an advantage to their customers.

In a next step, there is an investigation on the studies up to date regarding the logistic industry and more concretely data analytics and the methods applied in the model to develop the contribution of this project. There are different tools, techniques and methodologies for data analysis. The most common is the hypothesis contrast, but depending on the objective other techniques can be used, such as forecasting, classification or cause-effect methods (e.g. correlation analysis, regression analysis, data visualization, scenario analysis, data mining, text semantic analysis, patent analysis and scientific literature, Montecarlo simulation, mathematical programming and optimization, neuronal network), among others. Depending on the nature of the problem and the scope of the project, one or another statistical test will be applied.

The goal of the project is to have information about the costs' evolution in tenders to ease the decision-making. These types of problems make use of forecasting tools of pattern recognition and Machine Learning, where forecasting models and algorithms are used. In the following sub-section scientific literature of the analytical models and methodologies will be reviewed by project stages.

## 2.2 Literature by project stages

Initially, the idea around the project was to follow the work done for the [Stage 3. Modelling Success Probabilities in Tenders](#) in [Figure 4.8 \(page 42\)](#), showing the whole tender project that will be explained in following sections. Concretely, **Casassas et al. (2018) [5]** analyse data and generate predictive models, or in other words model the success probabilities in tenders.

To start with, **Wan and Beil (2006) [33]** considered prices and qualification to select a provider. Also, **Watt, Kayis and Willey (2008) [35]**, identified a representative and manageable set of variables relevant to the selection of suppliers. In **1995, Flanagan, D'Mello and O'Shaughnessy [11]** analysed data to study the factors that influence the results of a tender. Some years before but with a similar goal, **Hoffmeister and Dyl (1982) [19]** developed four models to help companies with the selection of target firms to start tender offer where the probability of success is high.

In **2008, Albano [1]** states that participation of small firms in the market for public contracts is a key policy issue. E-procurement through MEPA (Electronic Public Administrations Marketplace) is enhancing such participation and making it effective. In this field, the project studies aspects such as the relationship between the size and geographic place of the company and its success. Finally, **Ralph A. Walkling (1985) [31]** develops and tests a model that predicts the result of a tender offer.

A more detailed analysis continues with documentation related to the next phase [Stage 4. Optimization of the Tenders Portfolio](#) of the tender project, which was supposed to be the focus of the current Master Project, where optimization models are being developed to allow efficient pricing politics that increase the competitiveness in the bidding process.

Most literature on bidding projects move on the field of construction companies, which want to estimate and bid consistently in a way that reports benefit. Concretely **McGuire, Joseph (2006) [26]** prepares a project for McGuire, Kenneth; the Marketing Manager of Odyssey International. First, it can be found some theory about the preparation before bidding and secondly factors that contribute to the success of the bid are studied (as Xènia et al. during the first phase of the tender project). Third, some strategies that improve the possibility and profitability of a successful bid are presented in a way that success rate of competitive construction bidding is optimized.

Some years before, **Wanous et al. (1999) [34]** already studied the important decision whether or not to bid for a project in the construction industry. Similarly to the first phase of the tender project performed by Xènia et al., 38 factors affecting the decision are analysed and only the most important are considered. Then, 162 real bidding situations are used to optimize the



model and another 20 real projects to test it proving 85% accuracy in simulating the actual decisions.

From another point of view, the company offering the tender may want to maximize the result in terms of performance of the project. In **2016, Liang et al. [21]** developed a bidding strategy to support decision-making based on comprehensive information in construction projects. The model takes into account the set of design units, construction units and suppliers and not their individual performance. The problem is formulated as a 0-1 multi-objective optimization problem optimizing the collaborative performance of the different agents.

Curiously, the literature that sets a base for the optimization stage of the tender project are two papers not directly related to tenders. The first one by **Kizys et al. (2015) [20]** introduces an original algorithm for realistic portfolio optimization (ARPO). It quickly generates near-optimal asset-allocation solutions in terms of minimum risk for a given expected return rate under realistic conditions; including cardinality and quantity constraints. Moreover, the algorithm uses an iterated local search metaheuristic with quadratic programming. Additionally, specific operators and biased randomization strategies are used to reduce the risk of getting trapped in a local minimum and to make the search more efficient. The main shortcoming is the fact that it relies upon expected returns and variance data. Next will be reviewed a model which adds stochastic elements to the problem formulation, including random effects.

The second paper on in which the tender project is mainly based is from the recent **2018; Panadero et al. [30]** dealt with the relevant problem of maximizing net present value in project portfolio selection problem (PPSP) under uncertainty and rich conditions. Randomness is applied to periodical cash flows and also to discount rates, reducing the gap between theory or previous projects and practice. Firstly, a formal model for a stochastic version of PPSP is provided and then it is proposed a sim-heuristic approach to obtain efficient solutions. In general terms, the algorithm integrates Monte Carlo simulation inside a variable neighbourhood search framework, using other components such as simulated annealing or biased randomization techniques. Research concludes with findings on relation between expected NPV and risk, project interdependencies and comparisons between stochastic and deterministic solutions.

As mentioned before, the initial goal of the Final Master Project was to work on Stage 4 following the project by Casassas et al. (2018) in Stage 3. However, once inside the logistic company and from the position of tender coordinator, the company has not even reached Stage 1. The development and therefore literature review has also to be moved to the topics around the first stages.

First, [Stage 1. Automation of Transport Cost Analysis](#) is not very applicable in terms of research, as the main work is to combine the outputs from different tools in a platform without applying any calculation. That is the reason why this Final Master Project and literature coming

next is based on the development of a model to perform [Stage 2. Predicting Cost Evolution and Establishing Long-Term Tariffs](#). The costs' forecast is an interesting topic for many industries; information is power and as more information, better decisions, including decisions about the future, could be made. Currently, there are more than 70 prediction models, and for this reason the review of the articles will focus on the efficiency of the different methods to apply for these predictions.

The first article strictly focused in the stage of the whole tender project that this Final Master Project will develop is titled *Forecasting spot prices in bulk shipping using multivariate and univariate models* from **Geomelos et al. (2014) [13]**.

As in our case, the justification of the project is the volatility and uncertainty of spot market of shipping carriers. The aim of the research is to reduce this uncertainty by using estimates of ex-post and ex-ante forecasts. The methodology used is based on econometric analysis to compare different econometric models from two important categories of econometrics: (1) multivariate models (VAR and VECM) and (2) univariate time series models (ARIMA, GARCH and E-GARCH) in order to derive the best predicting model for each ship type. The results show that multivariate models give more precise forecasts; which can be used for the formation of efficient economic policy from decision-makers of shipping industry reducing spot markets' risk.

The second paper around the predictive stage is *Forecasting container shipping freight rates for the Far East-Northern Europe trade lane* by **Munim, Z. H., Schramm, H-J. (2017) [29]**. As in the previous case and in our own Final Master Project, the study introduces a state-of-the-art volatility forecasting method for container shipping freight rates. The difficulty to predict the container shipping industry during the last decade can be explained due to the demolition of the shipping conference system in 2008 for all trades calling a port in the European Union (EU) and the global financial crisis in 2009, which have affected the container shipping freight market adversely.

After the introductory justification of the project and literature review, the authors discuss contemporary container freight rate dynamics in an attempt to forecast for the Far East to Northern Europe trade lane. The methodologies employed in this case are Autoregressive Integrated Moving Average (ARIMA) as well as the combination of ARIMA and Autoregressive Conditional Heteroscedasticity (ARCH) model. It can be observed that the last combination provides comparatively better results than the existing freight rate forecasting models while performing short-term forecast on a weekly as well as monthly level. They also observed remarkable influence of recurrent general rate increases (GRIs) on the container freight rate volatility.

Additionally, **Coghlan, Avril (2018) [6]** is a little booklet with some useful information on how to use R for time series analysis. Other projects are also reviewed and have contributed to the final result of the Work. For example, **Bowen and Edwards (1985) [4]** comment that the use of methods based on historical data and deterministic do not represent reality in an honest way because factors such as price variability are not considered. For this reason, these investigations should move to stochastic simulation, so that the models represent reality more accurately.

**Gog (1998) [14]** compares the exactitude of the demand forecast between three forecasting models using three error measures. The results obtained by applying regression on this type of cases is not as efficient as applying other methods as the analysis of time series.

In the project of **Mentzer y Schroeter (1994) [27]** is commented that, even if a forecast focus mainly on software techniques, the development of an efficient forecasting process implies the integration of forecasting techniques together with systems and administration. However, there is not much investigation on the forecasting area with the combination of these three factors. The authors also establish that the analysis of time series is the main technique to forecast, as it includes tendency and seasonality

**Da Veiga et al. (2014) [7]** argue that the analytical methods of time series as Holt-Winters or ARIMA provide exact results in an easier way than other more complex techniques, even though it has to be considered that the application examples probably show cases where they are the most convenient methods to apply.

**Veenstra and Franses (1997) [31]** develop the initial stage of the process to do a long-term forecast of the volatile transport costs. In this new way to model the tariffs, they consider a multivariate time series model. It will be used to generate forecasts and identify the long-term tendency of costs.

## 2.2 Summary of literature review

To sum up the literature review, there are various projects on [Stage 3. Modelling Success Probabilities in Tenders](#) of the tender project ([Figure 4.8, page 42](#)), where the most relevant variables to achieve a tender are found. Being capable to know if the tender will be or not successful is a great advantage for the Company, because working on a tender and failing has an economic, time and effort waste by the logistic provider. In Flanagan, D'Mello and O'Shaughnessy (1998), Hoffmeister and Dyl (1981) or Wan and Beil (2009) projects it is explained the main variables that lead a tender to be accepted or not; highlighting the qualitative evaluation of possible providers, the size of the company and the geographical situation. Additionally, most authors make use of logistic regression techniques, which Walkling (1985)

shows that overcomes some of the limitations of other techniques. Therefore, this project will use the same as the author to predict the probability of success in tenders based on historical data, where initially and as in the previous investigations, an analysis of the available data will be made to prove which are more related to success. Later, logistic regression models will be developed until the one that best fits the data is achieved.

When it comes to the [Stage 4. Optimization of the Tenders Portfolio](#) with the attempt to maximize the benefit choosing the margin applied to each possible tender, there is not so much documentation. Some projects related with tenders that treat in some way optimization topics take different point of views to the one in which the current project focuses. For example, McGuire adjusts the model generated in the first phase using real data maximizing the number of outputs where the model result (success or fail) is equivalent to the result in reality. Another example is Liang et al., who aim to maximize the good performance of a project. Also, some similar characteristics are found between the tender optimization problem and the project portfolio selection project (PPSP). Therefore, projects such as Panadero's or Kizys' will be adapted and used in a tender optimization problem.

Finally, as explained during the literature review and previous introduction, the focus of the Final Master Project has moved to a preliminary stage to the fourth and even third stages. First, [Stage 1. Automation of Transport Cost Analysis](#) is not very applicable in terms of research, as the main work is to combine the outputs from different tools in a platform without applying any calculation. Therefore, the Final Master Project focuses on the development of a model to perform [Stage 2. Predicting Cost Evolution and Establishing Long-Term Tariffs](#). The forecast. Cost forecast is an interesting topic for many organizations, as information is power and, as much information, better decisions can be made, including decisions about the future. Given the dynamism in shipping costs and the difficulties involved, little literature has been found related to new techniques or methods to meet this challenge, and the vast majority focuses on prediction methods using econometric models, which have been very limited.

Most projects, even not being focused on the tenders, show that methods such as time series analysis are very useful for prediction. For example, in the work of Munim and Schramm (2017) shows how time series (ARIMA, GARCH and E-GARCH) have been used for the formation of efficient economic policies to make decisions on costs and prices in an industry with very dynamic costs and uncertain, as is the ocean transport case.

Currently, there are more than 70 prediction models, so some of the reviewed articles focus on the effectiveness of the different methods to apply for these predictions. Goh (1998) proves that applying time series models is much more efficient for these cases, and according to Da Veiga et al. (2014), time series models, specifically Holt-Winters or ARIMA, provide more accurate results more easily than other more complex methods. For this project, after the review

of existing literature, time series will be used. Obviously, different methodologies will be studied to decide which is the most appropriate the case of study.

It is interesting to mention that almost all projects found for all phases are in the field of construction. That is, after analysing existing literature it can be concluded that the use of tenders in the logistic field is a practice adapted from other industries and is being increasingly used. Therefore, the different stages of the whole tender project will be increasingly studied. In a near future it will be especially interesting the study of this fourth phase of the optimization of the tender benefit for a logistic company focusing on the decision of the margin applied to each proposal. Finally, the contribution of this project will be to create a methodology with few parameters and which are automatically adjusted, based on Winters' method (also called Triple Exponential Smoothing), which will allow to calculate *low-risk* forecasted intervals for the long-term tariffs to be offered in a tender.

## 3. Formal description of the problem

The problem of rating forecasting for ocean freight rates is described in detail in Section 3. To do so, there are three subsections which go deeper into the subject step by step. Firstly, the term *logistics* is defined and also what logistic providers does or which is their structure. Secondly, it is explained what a tender or reverse auction is in comparison to spot quotations. Thirdly and finally, the section focuses on rating forecasting ocean freight rates, which is a challenge for tender managers.

### 3.1 Logistics and logistic providers

The first step to treat the focus of this Final Master Project, which is the forecast of ocean freight rates, is to understand what is logistics, what a logistic company does and which is the structure.

A logistic company plans, implements and controls the movement and storage of goods, services or information within a supply chain and between the points of origin and final destination. Depending on each customer logistical requirements, only one or more logistic companies handle these supply chain functions.

To go back to the origins of the usage of “logistics” term it is required to go to military. In times of war supplying troops with the proper equipment and other supplies was a logistics function. Since 1950s, the evolution of logistics has been a result of the rise of consumerism and subsequent growth of more complex supply chains.

Nowadays, the coordination of supplies and materials movements is a globalized process. The supply chain is currently a vital part of the logistics process: efficient flow and storage of goods from point of origin to the point the final delivery point. A complete supply chain may include transportation, shipping, receiving, storage and management of all or one of these functions. Additionally, logistics may involve information, transportation, inventory, warehousing, packaging, disposal and security within the business sector.

In other words, the final goal of business logistics aims to achieve the optimal demand-service level at the lowest cost possible. The main function of business logistics is control over the movement of resources and supplies. Businesses need to ensure that required items are in the right place at the right time. A business suffers if it doesn't synchronize its supply chain.

### 3.1.1 3PL and 4PL vs. freight broker

Logistics companies provide logistical solutions to other organizations. Specifically, these services involve the management of freight, warehousing and inventory management. Businesses often outsource certain functions to logistical organizations to lower costs and increase the efficiency of their supply chains. Logistics companies are usually third-party logistics service providers and specialize in a specific area of logistics.

The terms 3PL and freight broker are sometimes used interchangeably. Both of them are intermediaries between the shipper and the carrier, however, their roles and services provided differ. In general terms, a broker focuses mainly on executing single shipments and meeting specific costs and service requirements. On the other hand, a 3PL provider considers a bigger picture in a more strategic way. Their scope goes from the early planning in the supply chain anticipating the needs of a client through every step of the supply chain, transport, and delivery process.

A [freight broker](#) will often focus on one single shipment and act as a mere intermediary (with no assets) between a shipper with goods to transport and the carrier which own the ships which will move the freight. In other words, a freight broker works with owner-operators to coordinate individual shipments. Currently, freight brokers also handle a significant amount of trucks in North America.

A [3rd party logistic company \(3PL\)](#) applies, in its broadest form, to any service contract that involves the storage or shipping of commodities; is the outsourcing of the logistics function to external organizations. The warehousing and transportation services are adapted to each client needs. A 3PL usually owns assets and establishes relationships with shippers on a more focused way. Third-party logistic companies typically make investments in physical and human capital including facilities and vehicles.

While 3PLs may only offer one service, such as transportation, warehouse storage or inventory management, [fourth-party logistics companies \(4PLs\)](#) offer a wide range of services and management of the whole client's supply chain, integrating also operations. In fact, 4PLs design and implement a comprehensive supply-chain solution for an organization that typically includes the use of a number of different third-party logistics providers. While a third-party logistics organization performs a specific logistics function, a fourth-party logistics company uses its resources and capabilities to manage the entire logistics process for clients.

For those shippers searching for a logistic company with a collaborative transportation network will see 3PL as the best option. However, freight brokers also have a role in 3PL relationships. When a shipper requires more flexibility in their shipping operations, it might include one or more brokerage firms to supplement its needs. From a shipper perspective, long-

standing relationships developed with 3PLs can mean an advantage when shipping capacities are scarce. Therefore, a mixture of relationships with 3PL companies and freight brokerages is the most interesting approach of delivering a good and reliable service while reducing shipping costs.

### 3.1.2 Logistic company structure

Regardless of the type of logistic company, the general goal of such kind of organizations is to provide inbound and outbound transportation of raw products and finished products from one end to another. To reach this and the rest of the wide range of tasks and goals, the hierarchy of a logistics company is divided on the basis of the responsibilities and roles of each of the employees. The hierarchy structure of a logistic company can be described as follows:

**Strategic level** executives comprise the highest level of any logistics company. These positions are meant to develop strategies so that the company reaches its goals following a correct functioning. In other words, the work positions in the strategic level ensure implementation of long terms in addition to setting budgets and determining targets.

Supply chain management includes, among others, logistics. As it is a very important part, the strategic level executives must have knowledge about logistics and other operations inside the supply chain. Within the described level, the following the chairman, director (executive and non-executive), chief finance officer, head of marketing, market analyst and project director designations can be found.

The mid-level of a logistic company, also called **managerial level**, looks at all the managerial aspects and makes sure that all tasks are correctly organizes and are working smoothly. Working on this level can be found the executives that coordinate the activities of the transport system and maximize the revenue expanding the client base. The activities to reach such goals are divided into duties reflected in the managerial titles of sales, branch, marketing, block development, HR, engineering, warehouse, transport, fleet, inventory, assistant and officer.

The lower level of any logistics company is the **executive level**. Even if it is on the bottom of the hierarchy, it should not be considered as less important. Related tasks are the implementation of the strategies defined in upper levels and the most common executive titles working at this level are technical support, IT, computer repair, in-house logistics, warehouse and material handling.

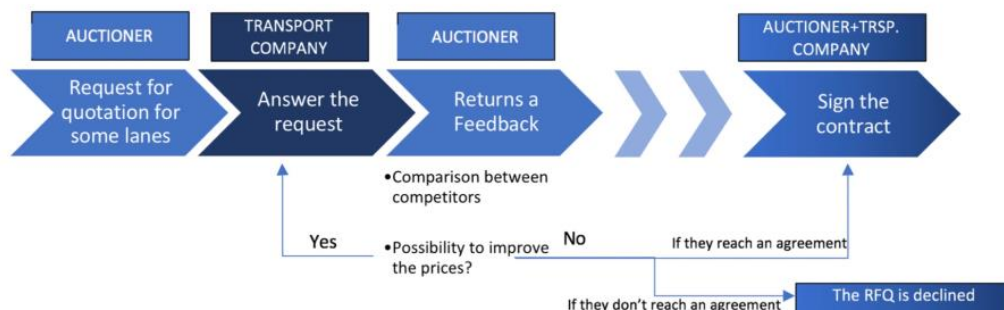


### 3.2 Spot quotation vs reverse auction

Usually the customer requests for information and costs to the provider in two ways, according to the requirements.

- **Spot quotation:** It is the price quoted and information provided for immediate settlement on a commodity, a security or a currency. The spot rate, also called “spot price,” is based on the value of an asset at the moment of the quote. This value is in turn based on how much buyers are willing to pay and how much sellers are willing to accept, which depends on factors such as current market value and expected future market value. As a result, spot rates change frequently and sometimes dramatically (Hafizoglu et al., 2016).
- **Tenders or also known as reverse auction:** By using reverse auctions across the value chain, companies can dramatically improve their sales, speed and efficiency metrics while also ensuring that supply chain functions will have a significant and demonstrable impact on the bottom line. Some of the important areas where companies use this process are: (1) Procurement; (2) acquisition of necessary services; (3) logistics and distribution.

In traditional auction process also known as forward auctions or tenders, the seller offers a product or goods that is in demand or requested by more than one buyer who participate in bidding process, bidding up the price (Manoochehri and Lindsay, 2008). The highest bidder wins the auction and buys the product. In reverse auction process, the buyer gets to choose the best quotation offered by the sellers. The sellers quote their lowest price and provide their best value added services in order to win the bid; the complete process has been explained in the introduction and added in **Figure 3.1**.



**Figure 3.1.** A scheme of a general RFQ process.

The reverse auction process starts with the buyers inviting and requesting several sellers to provide the information about them, this is called request for information RFI. In this process the sellers provide the following information: (1) Company presentation; (2) financial stability; (3) capacity of volumes that company can handle; (4) IT capabilities and connectivity to buyer's network; (5) global presence and strengths in regional trade lanes. After submission of RFI, buyers get to select several sellers according to best suit. Then the buyer prepares and invites the selected participants for request for proposal RFP or request for quotation RFQ. Here the sellers provide their best proposal, this includes: (1) lowest price; (2) fastest service; (3) other added services like custom clearance, intermodal service, consolidation service, etc. The buyer selects the best seller according to the best suited service.

In recent years the use of reverse auction is drastically increasing. The 2006 Purchasing's Annual Benchmark E-Sourcing Survey showed that 31% of respondents saying they use reverse auctions (Hannon, 2006). In this project we explore the benefits of reverse auction and the challenges faced by freight forwarding companies to provide strategic logistics and supply chain solutions for a long term period.

### 3.2.1 Advantages of reverse auctions

The most beneficial player in the reverse auction is the buyer while the supplier gets some benefits with a fixed contract for long term. The key benefits of reverse auctions include reduction of purchasing costs, increased efficiency in the market, enhancement of the procurement process, and access to a global supplier base. Some of the major factors that benefit buyer are; (1) reduction in purchasing cost; (2) increased market access; (3) efficiency in procurement process; (4) access to large supplier base.

### 3.2.2 Disadvantages of reverse auctions

The major disadvantages that a logistics service provider or supplier face in reverse auction are; (1) risk of new entrants; (2) very less profit margin; (3) many competitors; (4) forwarders may face loss during the peak seasons especially in terms of price quoted remains constant for the contracted period, usually one year.

### 3.2.3 Challenges

One of the major challenges of a logistics service provider in reverse auction process is

to procure the cost of container or a space in a vessel (ship, aircraft, rail) for long term from the carriers. This is because; the logistics service provider cannot promise the carrier that they will win the tender. On the other hand, due to uncertainties that may affect the logistics world (e.g. strict environmental regulation IMO2020, political aspects of certain countries, war, embargos etc.).

The other challenges that the logistic service providers face is to provide the best possible solution for the request made by customers, in terms of service and cost. This solely depends on the incoterms. For instance if the customer requests ex-works (EXW) service for full container load (FCL) of ocean freight transport and provides origin and destination locations, the service providers has to analyse the best possible route and mode of transport (by truck, rail, or barge) for pick-up of goods from origin to nearest port of loading, the fastest and cheapest carrier from port of loading to port of destination, best possible route and mode of transport for delivery of goods to the destination, and has to consider all other costs such as custom clearance, documentation, terminal handling fees, preparation of invoice, etc.

In the later stages of this proposal we analyse these challenges in detail and with real life data, the solution proposal is made by introducing and developing research methods and models.

### 3.3 Forecasting ocean freight rates

The literature review shows many attempts to develop forecasting models, but there has been no success in forecasting shipping freight rates. Several segments of the shipping industry have been and are interested, such as shipping companies but also banks, shipbuilders or logistic companies. The appearance of market slumps, like the one from 2007 crisis and following recession, was not known by ship owners, or at least they were not prepared to deal with it. Actually, some economists in the industry believe that the complexity of the problem makes it impossible to forecast how the market will develop.

The development of a forecasting model that works reliably for freight rates would provide a very valuable tool for stakeholders in the industry. They would be able to anticipate to the changes in the levels of freight rates and therefore to take actions to avoid, reduce the negative effects or even take advantage of it. To have a clearer picture of what is the problem to solve, it is interesting to highlight that the inputs and outputs expected from the model to develop. Concretely, the model will receive past years' ocean freight rates from one port to the other. Initial experimentations will be done with only one port pair (port of origin to port of destination) and can be extended in future work to all the port pairs operated by the logistic provider. After the model performs the required calculations using past years' data, it will

provide the expected ocean freight rate considering the past levels and evolution.

The most relevant stakeholders involved considering the scope of work of this project are shipping companies and consequently logistic providers who make use of shipping companies assets. These stakeholders need freight forecasting to estimate cash flows and profits, to develop budgets and to make decisions about the strategic direction of the company. Freight rates level result on the immediate revenue of the company within a financial year. Most of the decisions and actions in all companies depend on earning forecasting. Parallely, most of the decisions taken by shipping companies and logistic providers require forecasting of freight rate development to perform and make decisions, as their financial performance and strategy totally depend on it.

The revenue perspective of vessels leads to the decision about the purchase or lay-up. In other words, while a ship generates benefit, it will keep on service. From a shipping company perspective, if freight rates are not high, lay-up costs are, but freight rates are expected to recover, the company might decide to wait and see if ocean freight rates increase so that the vessel can generate profit. In case the market is strong and expected to continue growing, the shipping carrier will probably decide to buy new ships due to positive future expectations. The actions of shipping companies will, for sure, have an effect on logistic providers which make use of their vessels.

In addition to the decisions about the physical assets, financial decisions are clearly influenced by the levels of ocean freight rates. The revenue generated determines the flows of cash and therefore the free capital available to invest or servicing debt. Also, the value of the assets depends on the evolution of the financial situation. Therefore, reliable forecast of ocean freight will lead to a high probability to make correct financial decisions. Not only that, but freight rate development can influence a company's strategic focus regarding services offered and markets served. These facts will be of special interest in times when freight rates are declining and there are no signs of recovery, as happened during global recessions.

To sum up, it has been explained why shipping companies and their shareholders, as well as logistic providers have a strong interest in reliable forecast of freight rate developments. The freight rate level influences all aspects of the company operations in terms of share performance, debt service and new orders. This Final Master Project will contribute to find a way to develop a model using freight rate data. A good forecasting methodology will be a very valuable tool for decision makers in the industry to set a long-term strategy and also short-term tactics.

## 4. Our approach

In the previous section it was explained what is logistic industry, what is a logistic company and its structure and the focus was put on ocean freight rates forecasting problem. This fourth section is divided in 2 subsections; a general Implementation Plan to deal with different challenges faced on Tender Management department inside a logistic company and Smart Tenders Management with Artificial Intelligence project. The second sub-section treats in detail the second stage of the project: ocean freight rates forecasting using Winters' Method in RStudio; and the third stage: Modelling Success Probabilities in Tenders.

### 4.1 Implementation Plan

#### 4.1.1 First steps

Tender Management functions are to coordinate, design and deliver proposals to submit Requests for Information (RFI) and Requests for Quotation (RFQ), among others, as has been introduced in previous Section 3. The ultimate goal is achieving revenue and volume growth within the logistic service provider overall.

The motivation of the first part of Section 4 is to propose solutions to get the right line for an organization model in Tender Management and also try to reduce unproductivity and double works. The project is based in a real case, where the new Tender Management Air & Ocean organization is a real department inside an existing logistic service provider. Finally, the build-up methodology wants to be reflected in a better organization and effectiveness increase, focusing on Europe and Asia thanks to the coordination of Tender Management A&O in Europe and Asia.

##### 4.1.1.1 Standard Operating Procedure

A standard operating procedure (SOP) is a set of step-by-step instructions compiled by an organization to help workers carry out complex routine operations. The justification of the SOP creation is to avoid situations where the whole tender process was performed following different steps, which generated confusion and double-work across the organization. Therefore, the aim can be summarized as achieving efficiency, quality output and uniformity of performance, while reducing miscommunication and failure to comply with industry regulations.

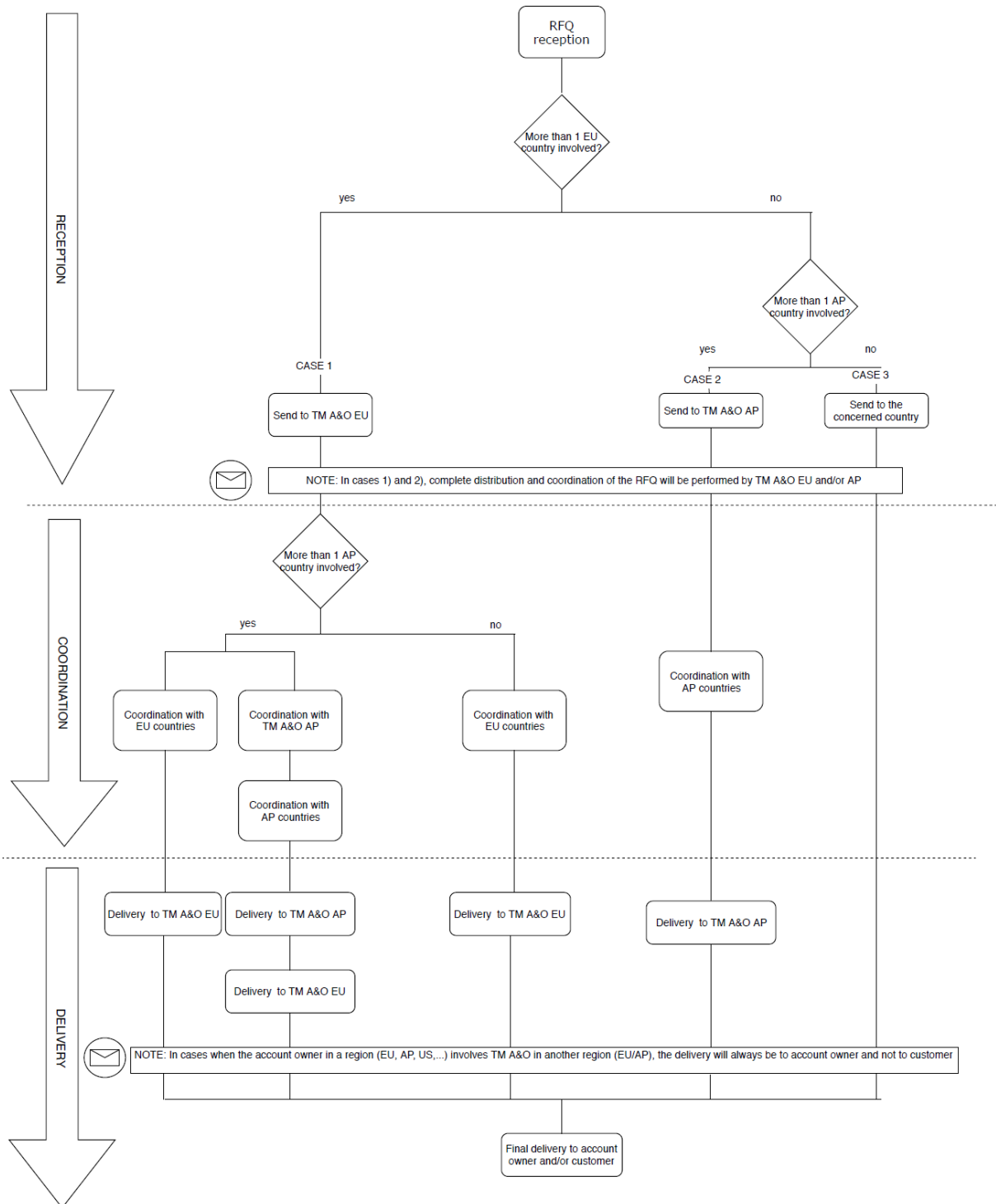
Concretely, the general purpose of Tender Management Standard Operating Procedure is to outlay the department responsibilities within the logistic service provider commercial organizations in all respective regions actively leading and steering well-defined Air & Ocean

related tenders. The SOP describes the general requirements, operating manner, deliverables and reporting lines. Therefore, it is updated with the newly implemented processes included in this Implementation Plan.

#### 4.1.1.2 Distribution Process

Preliminary steps to go through all the points described in this Implementation Plan are to settle distribution and management flow rules. For this reason, the Process Flow Distribution in [Figure 4.1](#) aim is to have a clear picture of communication for tenders from their reception to delivery. Similarly to the SOP point, the justification is to avoid past situations when the process and communication have not been efficient and straightforward, which has caused double-work as well as lack of reply from the people involved in the tender process. Therefore, according to the distribution process established, once the RFQ/RFI is received communication flows involve Tender Management A&O Europe and/or Asia, and concerning countries.

Important notes are firstly, the distribution only by Tender Management Europe and/or Asia when they are involved. Other Tender Management teams to be established within the company network in other regions e.g. US. Therefore, such teams must have the control from the beginning to the end. Secondly, in general terms the final delivery will be done by TM A&O EU when it involves more than one European country and TM A&O AP when it involves more than one Asian country (and not more than one European country); always keeping in copy the account owner. Only in cases when an account owner in a region involves TM in another region the delivery will be to the account owner and not to the customer. For example, it would be the case of an account owner involving TM A&O AP for on-carriage in different Asian countries for a RFQ only origin one European country.



*Figure 4.1. Distribution Process Flow for RFQ/RFI*

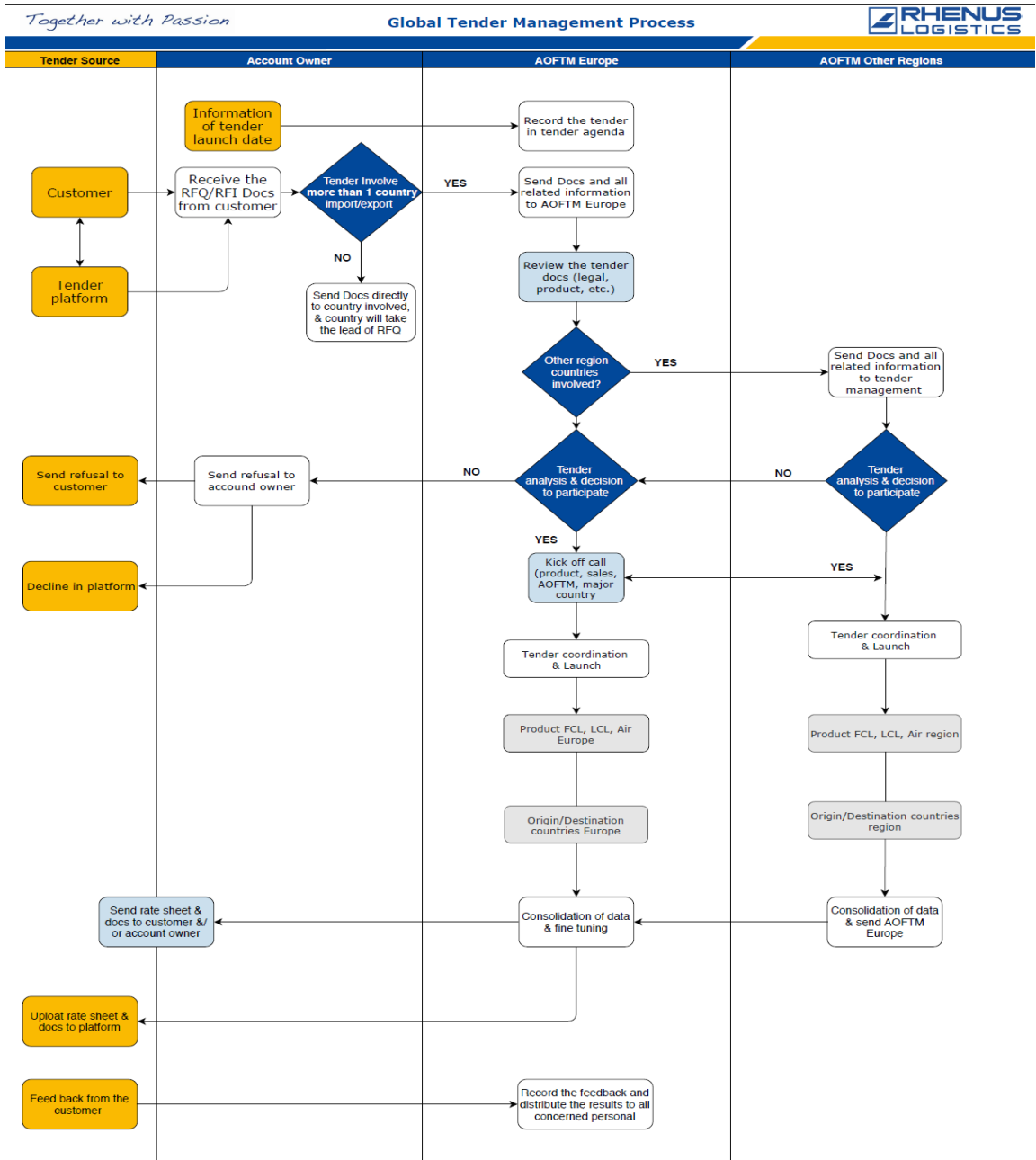
Together with the Distribution Process Flow, it is required a **global contact list for RFQ/RFI**. The contact list includes a dedicated mail address for Tender Management A&O in the different regions where the logistic provider has presence: Europe, Asia Pacific or some main countries in the organization with also a Tender Management Team, such as USA or

Germany.

#### 4.1.1.3 Tender Management Process Flow

After the Distribution Process Flow set up, the Tender Management Process Flow is also a key point to be implemented in parallel in the different regions. While the Distribution Process referred to the flow of information and coordination, in this [Figure 4.2](#) we can see the different steps and links between Tender Management A&O Europe and Asia and the sales person in charge of the corresponding RFI/RFI. The process flow is included in the SOP, sharing its justification and final goal: avoid unstructured and inefficient situations to achieve structured performance, efficiency, quality in the output and finally success.





*Figure 4.2. Tender Management Process Flow*

### 4.1.2 Rating automation

Tender sheets can be received in different formats, some of them more and some of them less structured. A logistic provider faces multiple challenges from various sections of the company, such as focus on sales and pipeline, organically growth and integration of new RFQ/RFI from acquired companies. A key section inside Tender Management A&O Europe Implementation Plan is the rating automation in different steps; forecasting Ocean Freight with al alignment between academic research and the knowledge of Product teams. The rating

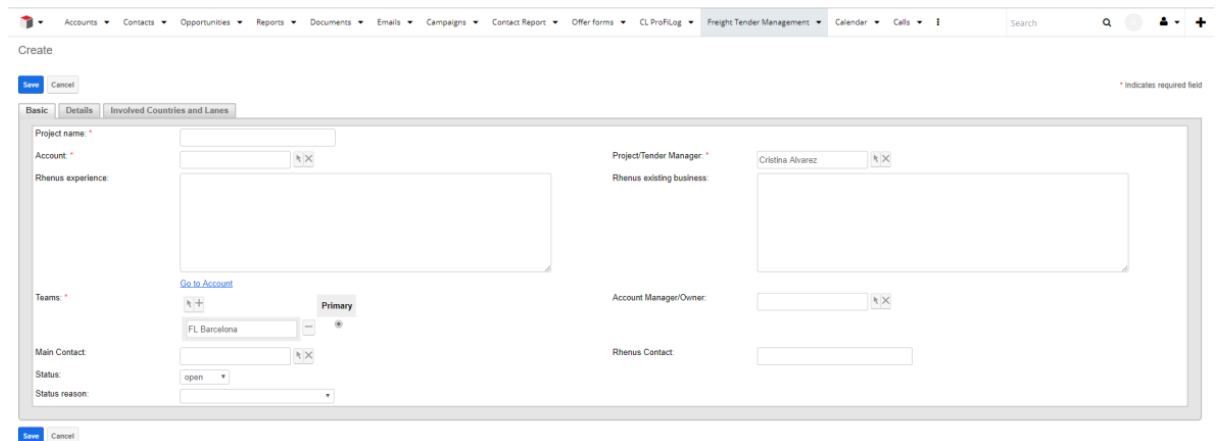
automation, including Ocean Freight forecast will be the main body of this Final Master Project and will be further developed in point 4.2 and experimental analysis performed in section 5.

### 4.1.3 Reporting and analysis

#### 4.1.3.1 Sugar CRM reporting and analysis

Sugar provides a simple CRM interface for a wide range of companies, including logistic providers. Therefore, it is the solution to manage the relationship with the clients in three basic fields: the commercial management, marketing and the clients' follow-up.

The justification of this project development is that, in the example company used as a case of study, it is not used in all countries and where it is used, it is not used in the same way. Therefore, the objective or main focus of the Tender Management A&O Europe regarding Sugar is the RFQ/RFI reporting to customize and automate business processes. To reach this goal the process will be (1) to define the fields and new module requests adapting Freight Tender Management module in **Figure 4.3**, (2) to create a manual for use of TM A&O in Sugar, (3) to distribute and implement firstly in Germany as it is the more advanced country within the organization and (4) finally to distribute and implement it globally.

The image shows a screenshot of the Sugar CRM 'Create' form for the 'Freight Tender Management' module, specifically the 'Basic' tab. The form is titled 'Create' and has a navigation bar at the top with various menu items like 'Accounts', 'Contacts', 'Opportunities', etc. The form itself is divided into several sections: 'Project name' (text input), 'Account' (lookup field), 'Project/Tender Manager' (lookup field with 'Cristina Alvarez' selected), 'Rhenus experience' (text area), 'Rhenus existing business' (text area), 'Teams' (lookup field with 'FL Barcelona' selected and a 'Primary' checkbox), 'Account Manager/Owner' (lookup field), 'Main Contact' (lookup field), 'Rhenus Contact' (lookup field), 'Status' (dropdown menu with 'open' selected), and 'Status reason' (dropdown menu). There are 'Save' and 'Cancel' buttons at the top and bottom of the form. A small note '\* Indicates required field' is visible in the top right corner of the form area.

**Figure 4.3.** Sugar CRM, Freight Tender Management Module, Basic tab

The final aim is to integrate the different documents that need to be shared worldwide for tender reporting, including historical data, tenders on hand and finally the agenda for the following months. This provides sales people the flexibility to build great, personal customer

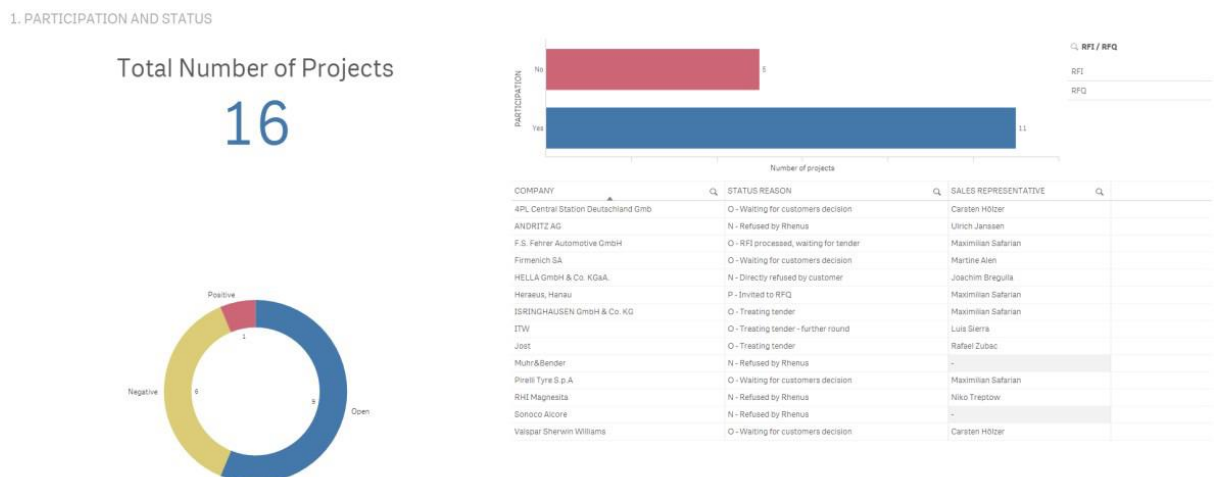
relationships in an easy way while people and teams working in tenders encourage strategic growth.

#### 4.1.3.2 QlikSense tender analysis

QlikSense is a Business Intelligence tool which allows to freely search and explore across data, instantly pivoting analysis when new ideas surface. It is ideal for a global enterprise to take advantage of machine suggestions and automation powered by Qlik cognitive engine and get total flexibility with a multi-cloud data analytics platform that supports the full spectrum of BI use cases.

Similarly that the justification and final goal explained in Sugar CRM, QlikSense BI is available in the case of study company (at least desktop version), but it is not used commonly across it. To reach the Implementation Plan goals in reference to QlikSense the steps will be (1) to create applications as seen in the examples from **Figures 4.4** and **4.5**, (2) request QlikSense server version and link Sugar report, (3) to implement in Germany and (4) to implement globally.

**Figure 4.4** shows the total number of RFQs and RFIs that the tender management team has received and analysed in a period of time. Colour blue represents open projects, e.g. *treating tender* or *waiting for customer decision*; colour yellow represents negative projects, e.g. *too expensive* or *no feedback*; and colour red represents positive projects, e.g. *won parts*.



**Figure 4.4.** QlikSense application for tender analysis, participation and status

Figure 4.5 shows the FCL, LCL and airfreight volumes for each project or tender. That is, the expected containers or kilograms that the awarded logistic provider would be moving during one year. In general terms, as bigger the volume is, the more turnover and therefore interesting is the request.

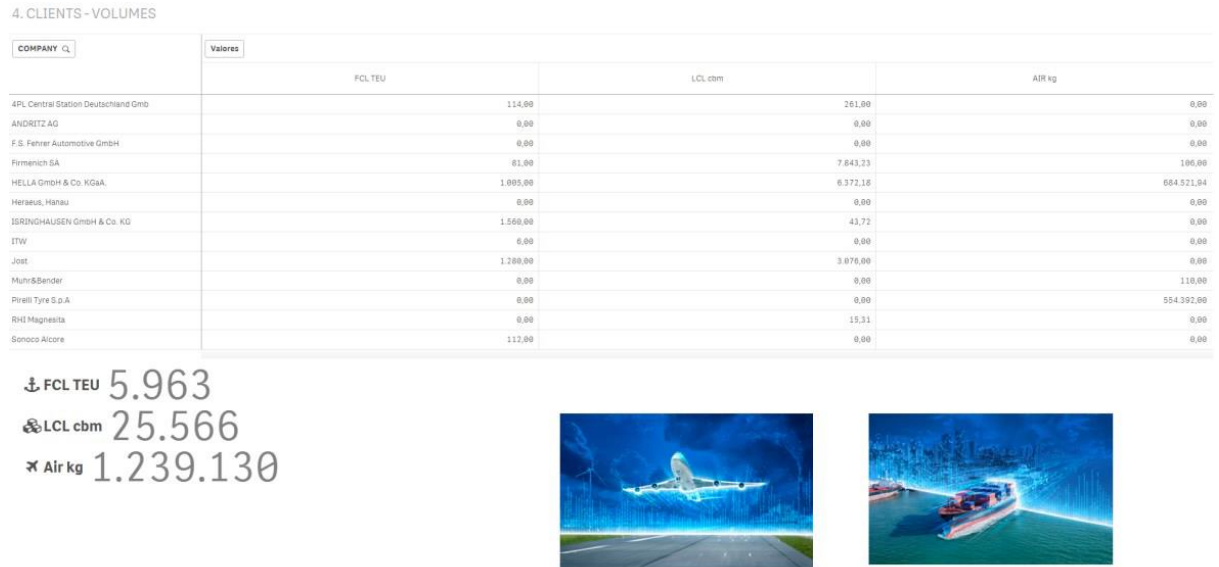


Figure 4.5. QlikSense application for tender analysis, clients and volumes

The final goal is to use the Sugar reports explained in the previous point to generate analysis for historical data, current situation and upcoming RFQs/RFIs. This will allow having a clear picture of the relevant information of tenders for all the interested parties involved.

#### 4.1.4 Scoring system

A key decision in a Tender Team when a RFQ/RFI is received is the participation or not, as detailed in Tender Management Process Flow Figure 4.2. The need of the tool can be justified with the argument that not all tenders can be delivered to the customer due to the workload during some seasons and resources available, at least until the rating automation explained in previous points is further developed. Therefore, the objective is to create a tool that can support the decision on which opportunities or tenders is better to invest efforts and resources.

To develop the tool, the Tender Scoring System created by Road Department will be adapted to Air & Ocean requirements following the next steps: (1) definition of required data for A&O, (2) data collection from countries and tuning, (3) model creation and (4) the final

implementation of the model.

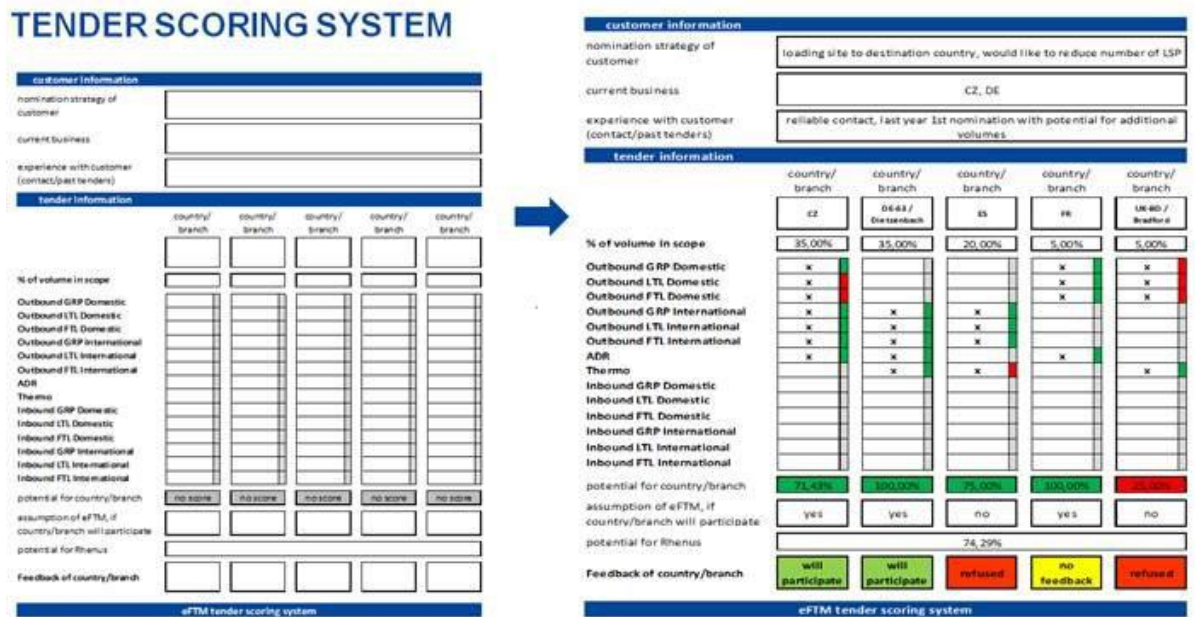


Figure 4.6. Tender Scoring System from Road Division, general template and example

The model to be adapted from Road Division in Figure 4.6 contains the different criteria with different weighing, based on importance inside the evaluation. The target of the Tender Scoring System for A&O is to have a general overview of a RFQ to see in short time the potential for the logistic provider. As can be seen in the right side of Figure 4.6, the score shows how good the customer/tender fits in the existing company network and therefore helps in the participation decision. Finally, the Scoring System will be distributed to all involved parties, so that all participants have this picture.

#### 4.1.5 Database creation for RFIs

Among the main tasks of Tender Team A&O Europe as detailed in the Standard Operating Procedure is to analyse, coordinate and respond Requests for Information. One of the challenges faced by Tender Team while filling RFIs is to address the right person among the company departments such as IT, Financial or Product to get the information required by the customer. In some occasions, new information has to be provided but in some others, the data has already been requested to fill past RFIs. In other words, the project justification is the lack of a database and the goals to make the process efficient and provide solid and correct information to the customers.

Therefore, the steps to complete the database for RFI will be (1) to decide the most appropriate platform to be used, (2) to incorporate previous data and finally (3) to define the system usage and how new data will be incorporated.

The main goal of the database creation is to have in a defined platform all the historical points requested by companies. Therefore, it will be accessible for defined teams so that new RFIs can be efficiently filled avoiding double work.

The RFI database is currently being developed in Wiki page of Air & Ocean Freight Tender Management shown in **Figure 4.7**. Accessing this page, one can provide the information to RFI of customers.



*Figure 4.7. RFI database in Wiki page of Air & Ocean Freight Tender Management*

## 4.2 Smart Tenders Management with Artificial Intelligence

### 4.2.1 Project overview

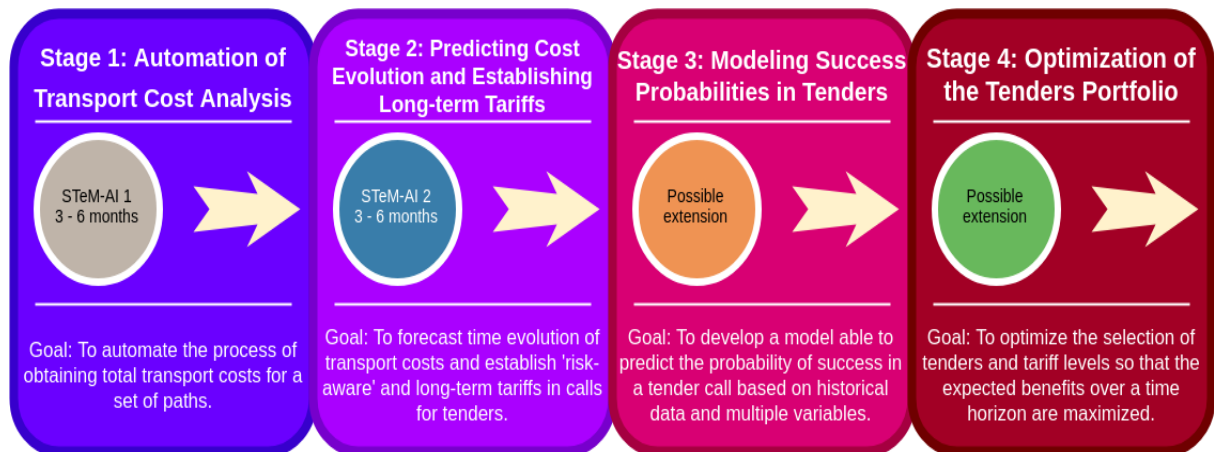
In the real-life case that motivates this study, the computation of the cost associated with a given transportation path (i.e., a route connecting an origin with a destination) is done using several tools. These tools depend on several factors, e.g.: the transportation mode (aerial, terrestrial, ocean, etc.); the service requirements (e.g., full container load or less than container load, country specification charges, etc.); transport related charges (custom clearance, documentation, etc.); and the information related to the transportation mode (transit times, carrier name, vessel, demurrage and detention, etc.). In other words, each tool only provides a partial cost for the given path, but costs from different tools need to be integrated at a later step in order to obtain the total transportation cost.

The integration of data coming from different tools needs to be done in an automatic way in order to avoid unnecessary processing and waste of human resources. This is not always easy to achieve due to the fact that, typically, each software tool uses a different data

format and each country or branch might have specific characteristics that makes difficult the existence of a single 'universal' tool. This will be the first stage in the complete STeM-AI project displayed in **Figure 4.8**.

Secondly, the transportation costs associated with a path are only valid for short-term periods. Hence, they cannot be used to estimate the evolution of costs throughout longer periods (e.g., one year or even more). Without a reliable cost estimation on the long run, it is not possible to provide 'low-risk' fare tariffs for tender calls. Thirdly, there is a lack of tools that allow for identifying the cause of success or failure in tender calls, as well as methods that can predict the probability of success depending on the offered price.

Fourthly and finally, it is difficult to manage tender selection and pricing policies, i.e.: how could managers configure an 'optimal' portfolio of tenders and what prices can they offer during the corresponding bidding processes? (with the ultimate goal of maximizing the expected benefits in the long run and without exceeding their service capacity).



*Figure 4.8. Structure of the STeM-AI project.*

## 4.2.2 Predicting Cost Evolution and Establishing Long-term Tariffs using Time Series

### 4.2.2.1 Goals

Time Series is a technique that a scientist must have in his or her back to predict future events. Basically forecasting is predicting future events based on historical

activities. The prediction of values is based on the assumption that events in the past will continue to have an impact in the future.

The two main goals of time series analysis are:

- **Identify the pattern** represented by the sequence of observations.
- **Forecast** future values of the time series variable.

#### 4.2.2.2 Use cases

Among the possible uses cases can be found:

- **Predicting revenues.** All companies are highly interested in predicting their revenues based on the activities they are doing. Sometimes the companies have to give guidance to the Wall Street or the respective stock market in reference to the revenue they will generate the next quarter. For example, they could have to give a forecasted higher and lower range, e.g. between 50 million and 55 million dollars; not an actual one.
- **Predicting earthquakes.** Several scientists are collecting the data based on the magnitude, the initial point, the circumstances why the earthquake might have occurred. With this information they try to predict future earthquakes. Occasionally it appears in newspapers that whenever a new earthquake is coming, there is advice from the government to the public that this event could happen.
- **Predicting new jobs.** Based on the events that the industry has seen, data scientists can predict the new jobs that are coming overall in a market or in a concrete segment of a market.
- **Predicting student enrolments.** The direction board of a college can try to predict the new enrolments to catch based on the past performance.

#### 4.2.2.3 Process

Some usage examples have been listed and explained, now the process of time series forecasting will also be listed. Note the following process has to be considered a very high level process, as there will be a range of sub-processes:

1. **Plot the data.** The very first step is plotting the data in a certain period of time. As mentioned before, intervals in the whole period of time must be equal: monthly,



quarterly, yearly,...

2. **Analyze if data is stationary.** If it is not, it has to be made stationary. As explained, data not stationary is due to trend, seasonal, cyclical or irregular components.
3. **Technique identification.** Identify the model technique or method that best suited the forecasting; Moving Average, Exponential Moving Averages, Least Square Regression, Quadratic Regression, Exponential Regression and ARIMA Models.
4. **Build the model.**
5. **Do the prediction.**

#### 4.2.2.4 Data and components

One of the most complex parts of prediction is the collection of valid and reliable data. A prediction cannot be more precise than the data it is based on. The most sophisticated model will fail if data is not reliable. The following 4 criteria can be applied to know if data will be useful:

- **Reliable and precise.** Data has to be collected from a reliable and precise source.
- **Relevant.** Data needs to be representative for the circumstances it will be used.
- **Consistent.** When the collection process is modified, it has to be adjusted to keep the consistency in the historical patterns.
- **Opportune** data will be more valuable

A time series consists on data that is collected, saved or observed through successive increments of time. The analysis of time series assumes that the data consist of a systematic pattern (a set of identifiable components) and of random noise (error) that makes the pattern difficult to identify. Most techniques have some way of filtering this noise to be able to identify the most essential way of the pattern.

As introduced previously, a times series is a set of numerical data obtained at a regular period of time. It is important to highlight that the periods of time must be regular. For example, times series can be monthly, quarterly or yearly, but the period cannot change between points. Based on such periods, data can be plotted and observe what type of time series it is. Once the plot is ready, the analysis can be based on different components:

- **Trend component.** Emphasis has to be put into trend, being one of the aspects to be observed. The fact of data following some kind of trend or not, upward or downward, is an important component of a time series.
- **Seasonal component.** If the line plotted is making some kind of peaks after regular interval of times, that means that there is some sort of seasonal component.
- **Cyclical component.** The nature of those upward and downward peaks differentiate

seasonal versus cyclical component. When the peaks only depend on time, the component to consider is a cyclical one. However, if for example, hotel occupancy is increasing due to summer season, seasonality is making effect. Therefore, it only depends on the data and expertise of the analyser to say if it is a seasonal or cyclical component.

- **Irregular component.** Any unusual spike or data point is very high or very low which the analyser is not able to explain without additional information about the data. For example, treating data from a company that experienced a big wage can be reflected into a decrease of revenue during a concrete month. Therefore, with a certain level of business expertise the effects on the plot can be explained and adjustments made before going ahead and creating a model.

These components are important because, if there is a real aim to predict, e.g. the next revenues, sales, or anything of interest, the components need to be removed and series made stationary. When any of the components trend, seasonal, cyclical, or irregular is present in the data means that it is not stationary or not normally distributed. A Normal distribution is one of the assumptions of doing the regression or forecasting, so the components must be removed to be able to forecast. In next steps, it will be explained how to make data stationary or normal removing these components during the process of creating time series models.

### 4.2.3 Winters' Method

#### 4.2.3.1 Technique justification

To select the most appropriate forecasting method for a set of data, existing patterns have to be identified. To do it, auto-correlation can be used, as it is common that observations of different periods of time are related. The autocorrelation will be measured using autocorrelation coefficients.

In our problem, the data presents tendency patterns and for this reason smoothing methods or ARIMA have to be chosen. In general terms, ARIMA is considered more sophisticated and precise in the cases that patterns are not easily visible. It is more a complex method to apply, so it is important to highlight that a more sophisticated and complex method does not mean that it will generate more precise results than a simpler one.

The comparative result of [Da Veiga et al., 2014](#) between ARIMA and Holt-Winters' models, comments that ARIMA methodology is proved to be much more precise and more

complex, but depending on the case of study the assumptions can change. In other words, Winters' method can give more precise results than more complex methods, considering a not very long forecasted time horizon. Additionally, each technique must also be evaluated at the level of reliability and applicability to the problem, the length of the forecast horizon or the acceptance of decision-makers who are going to make use of the tool in the organization.

For all the above, one of the smoothing methods will be chosen to predict the costs of transport of the tenders. Smoothing methods, model the components in time series which are usually easy to identify in a time series data chart. These methods break down the data into trend and seasonal components, and then extend the estimates of the components to the future to give predictions. They are divided in static methods, which have patterns that do not change over time, and dynamic methods (Moving Averages, simple, double and triple exponential smoothing) that have patterns that do change over time and estimates are updated using neighbouring values.

In this case of study, the values change over time, therefore dynamic methods will be used. Among these, the average methods, such as moving averages, generate, as the name implies, predictions based on the averages of the most recent observations. The smoothing methods do the same as the previous ones but with a series of exponential decreasing weights.

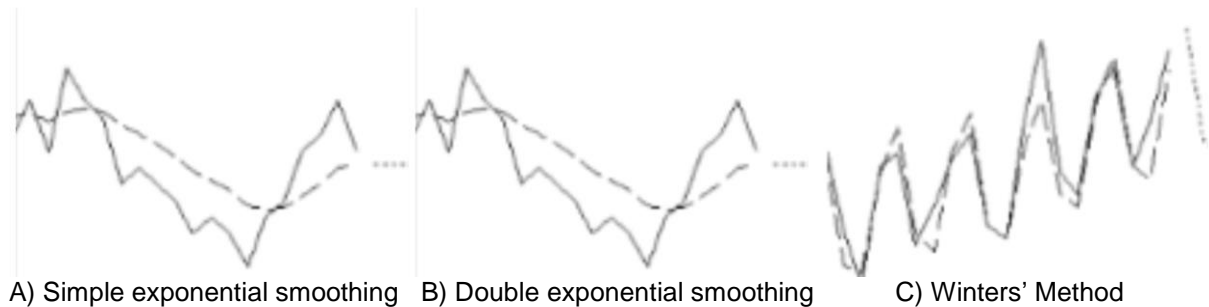
The previously mentioned are more suitable for data in which the increase or decrease in the trend is not predictable, therefore they are more robust methods to predict values. That is why they are the most interesting for the contribution of this project.

The **Table 4.1** defines the main characteristics of each exponential smoothing model (**Figure 4.9**):

	Model	Definition	Prediction
A	<p>SIMPLE EXPONENTIAL SMOOTHING</p> <p>For series without tendency and seasonal component</p>	This procedure works better without tendency and seasonality. The simple dynamic component in a moving average model is the level.	<p>Length: short-term</p> <p>Profile: flat line</p>
B	DOUBLE EXPONENTIAL	This procedure is adequate when there is a tendency but there is not	Length: short-term

	<p>SMOOTHING</p> <p>For series with tendency but without seasonal component</p>	<p>seasonality. It calculates the dynamic estimator for two components: level and tendency.</p>	<p>Profile: straight line with increasing / decreasing slope equal to the last estimated tendency</p>
C	<p>TRIPLE EXPONENTIAL SMOOTHING (WINTERS')</p> <p>For series with tendency and seasonal component</p>	<p>This method to smooth data using Holt Winters' is adequate when both tendency and seasonality are present, which can be additive or multiplicative. Winters' method calculate dynamic estimators for three components: level, trend and seasonal.</p>	<p>Length: short / mid-term</p> <p>Profile: tendency with seasonal pattern</p>

*Table 4.1. Exponential smoothing methods definition and prediction*



*Figure 4.9. Exponential smoothing methods graphics*

Since the transport cost data in the tenders show seasonality, the most suitable model of exponential smoothing is triple, or also called as Winters' method. The reason is that in addition to serving as a general smoothing method, it takes into account the seasonality and can represent data better and reduce error (Hanke, Reitsch and Wichern, 2001) [16].

#### 4.2.3.2 Winters' method definition

The triple smoothing method or Winters' is used in series that present both tendency and seasonality components. That means, the forecasts are not only calculated from the previous observations but also independent smoothing tendency and seasonality have to be considered.

The seasonal component can be additive or multiplicative. For example, during December the sales of a particular toy can be increased by 1 million dollars every year. So, an amount of 1 million dollars can be added to the forecasts for each December (with respect to the annual average) to take into account this seasonal fluctuation. In this case the seasonality is additive, and the series show stable seasonal fluctuations. On the contrary, during the month of December sales can grow by 40%, and in this case the seasonal component is multiplied by nature. Therefore, the measure of seasonal fluctuations varies, depending on of the general level of the series (Figure 4.10).

- **Additive model** has to be used for data with or without constant trend and seasonal pattern constant (with measurement not proportional to the data). Forecast Profile: straight line with the seasonal pattern added.
- **Multiplicative model** has to be used for data with or without constant trend and seasonal pattern constant (with measure proportional to the data). Forecast Profile: straight line with seasonal pattern multiplied.

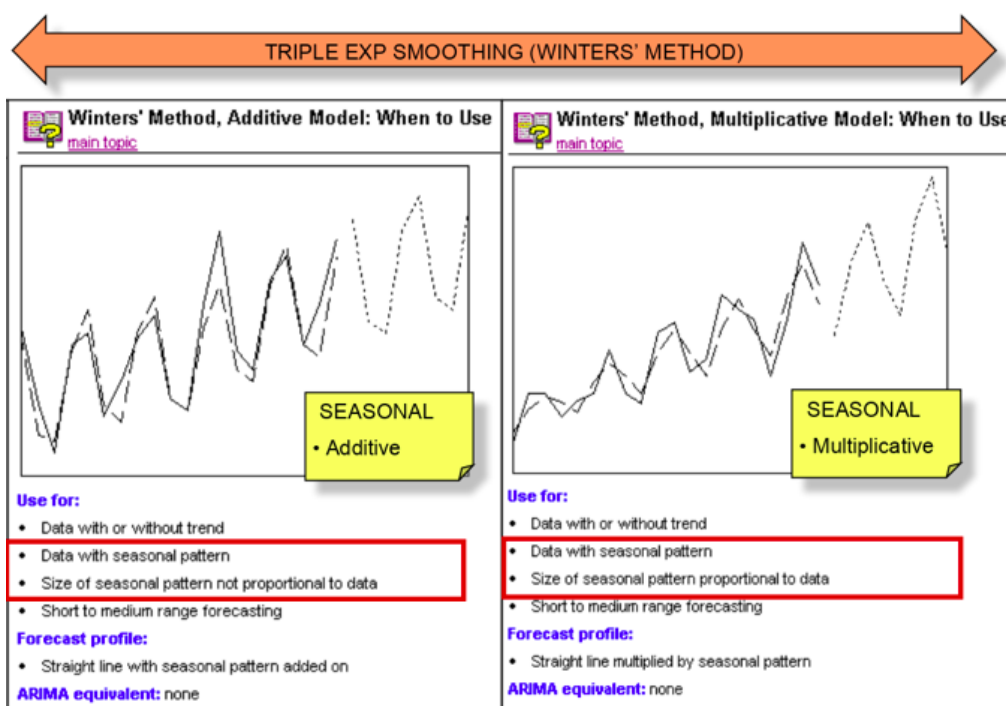


Figure 4.10. Example of Winters' method with additive and multiplicative seasonality

In general terms, as will be explained in Section 5, data presents multiplicative seasonality. For this reason, from now on the reference will be to multiplicative models. The method calculates dynamic estimators for short and mid-term components: (i) level ( $L_t$ ); (ii) tendency ( $T_t$ ); and (iii) seasonality ( $S_t$ ); and also for future forecasted periods ( $\hat{Y}_{t+p}$ ).

#### 4.2.3.3 Method equations

The formulas used in the multiplicative Winters' smoothing to calculate estimations and predictions:

- Estimation of the level or exponential smoothing series

$$L_t = (\alpha * (Y_t / S_{t-s})) + (1 - \alpha) * (L_{t-1} + T_{t-1})$$

- Estimation of tendency

$$T_t = \beta * (L_{t-1} + T_{t-1}) + (1 - \beta) * (T_{t-1})$$

- Estimation of seasonality

$$S_t = (\gamma * (Y_t / L_t)) + (1 - \gamma) * S_{t-s}$$

- Forecast of future periods

$$\hat{Y}_{t+p} = (L_t + pT_t) * S_{t-s+p}$$

where:

$Y_t$  = new observation or actual value in period t

$L_t$  = new smoothed value or actual estimated level

$T_t$  = tendency estimation

$S_t$  = seasonality estimation

$\alpha$  = smoothing constant for the level

$\beta$  = smoothing constant for the tendency estimation

$\gamma$  = smoothing constant for the seasonality estimation

$S$  = length of the seasonality

$p$  = periods to forecast in the future

$\hat{Y}_{t+p}$  = forecast for future periods

#### 4.2.3.4 Method evaluation, error measures

Several investigations (Makridakis and Hibon, 1997; Fildes et al., 1998; Makridakis et al., 1993) demonstrate that the precision of simple methods is almost as good as that of more complex and sophisticated techniques. This means that, for each problem, there is a technique that fits better than others and does not necessarily have to be the most sophisticated and complex. For checking if the chosen method works, MAPE, MAD and MSD error measures will be used. The statistical software calculates three measures of accuracy of the fitted model:

- The **MAPE (Mean Absolute Percentage Error)** measures the accuracy of time series adjusted values (in %).
- The **MAD (Mean Absolute Deviation)** measures the accuracy of time series adjusted values (in the same unit as the data - it helps conceptualize the amount of error).
- The **MSD (Mean Square Deviation) or MSE** is always calculated using the same denominator,  $n$ , whatever the model is, so that it can be compared the value of MSD between models. MSD is a more sensitive measure to an unusually larger forecast error than MAD.

A difference between MSD and MAD is that the MSD measure is much more influenced by **large adjustment errors** than by small errors, since for MSD the errors are squared.

For all error measures, at a lower value, better model fit. These statistics are used to compare the settings of different methods or different parameters in the same method. As a part

of Winters' Method, the values of the parameters  $a$ ,  $b$ ,  $g$  (or level smoothing constants, trend and seasonality respectively), are values that can be chosen according to some subjective criteria. Optimization can also be used to find the optimal values that minimize the error. The constants are the ones that best fit the model, and therefore if more accurate or with less error predictions are required, the values that best fit the model have to be chosen. In other words, unknown parameters are determined by minimizing the squared prediction error. If optimization is not used for some reason, the most appropriate values can be selected taking into account the values of the error measures. In this way, the values of the parameters are changed in a subjective way so that the MAPE, MAD or MSD values are decreasing and therefore the model fits better.

## 4.2.4 Modelling Success Probabilities in Tenders

### 4.2.4.1 Introduction

The methodology proposed in Stage 3 is based on the use of historical data and machine learning methods. In particular, we make use of logistic regression models to depict how the probability of obtaining a tender might vary for different values of decision variables (e.g., the long-term tariff provided, which should cover the transport costs plus a certain profit margin) and exogenous variables (e.g., region, transport mode, volume, customers' fidelity, etc.).

First, a descriptive and correlation analysis of the set of data is completed. This analysis allows for identifying the main variables that affect, in a direct way, the probability of success. It also allows us to obtain information about the relationship among these variables. Next, binomial regression models are built by adding one variable at a time until a proper adjustment is found. For these models, two different link functions are taken, the logit and the probit ones (Borooah, 2002).

As explained before, the goal is to obtain a model that can allow to predict the probability of success in the tender as a function of the level we assign to the decision variables, in particular the long-term tariff. Notice that a higher tariff will lower the competitiveness in the market, but will also reduce the risk of suffering losses due to offering tariffs that fall below the future transportation costs. Also, knowing the behaviour of the logistic model might be useful to select the right value for the tariffs. For example, imagine a scenario in which a tariff of 620 USD has an associated success probability of 0.75, while a tariff of 690 USD has an associated success probability of 0.73 and a tariff of 700 USD has an associated success probability of 0.60. In such a scenario, it seems clear that most managers will chose a tariff of 690 USD due to the trade-off between potential revenue and success probability.



#### 4.2.4.2 Data

In order to approximate the response by using a model, artificially generated data will be needed. However, to make the model as more realistic as possible, random data will be avoided as it does not represent a company real data. So as to prepare a realistic setting, the variables that may affect the win or lose of the tender are provided by the logistic company in which this Final Master Project is based. Some of these variables are the geographic zone where the transport will be managed from, the importance for the company to win the tender, the number of years of relationship with this client, the number of annual shipments, the base cost of each shipment and the margin beyond the base cost. This last one, plays the main role in taking the decision about the acceptance/refusal of the tender. In addition, margin cost is the most interesting for a decision maker as choosing its value is its responsibility. The variables are divided between company and tender variables:

##### **Company:**

- CompanyID: company identification.
- GeoZone: location of the client.
- yearsC: years of relationship with the client.
- levelCompany: prestige of the client.
- levelRelationship: value of the relationship with the client.
- #Tenders: number of tenders offered by the client.
- #Tenders\_1: number successful of tenders.
- % Tenders\_1: percentage of successful tenders.

##### **Tenders:**

- Num\_Shipments: number of shipments included in a tender.
- Cost\_per\_Shipment: cost per shipment of a tender.
- CompanyID: -
- Geo\_Zone: -
- yearsC: -
- levelTender: value of the tender.
- Margin: benefit margin assigned to the proposal.

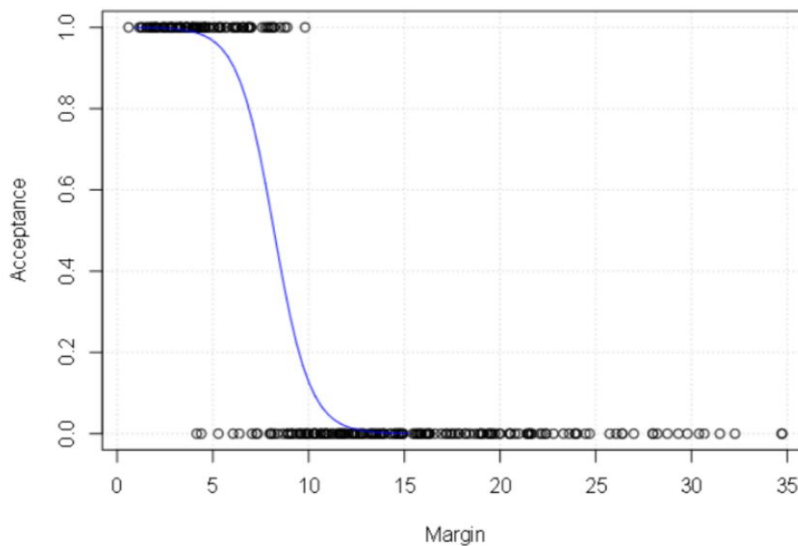
Moreover, the relationship between the margin and the probability of success is an inverse relation, as occurs with the price per shipment. It means that an increase of the margin (or price) implies a reduction of the probability of success. Otherwise, the other variables have a direct relationship with the probability of success. After the generation of a set of data taking into

account the considerations described before, now a descriptive analysis is made to check the relationships between each variable and the response and to study their behaviour. First of all, the calculation of the correlation coefficient between each numeric variable and the response and the obtained results are the following:

Variable	Correlation with acceptance
Num Shipments	0.2836
Cost per Shipment	-0.5093
Margin	-0.7294
Importance of Tender	0.5935
Level of Relationship	0.3987

#### 4.2.4.3 Results

After an initial data analysis, it is found that the most relevant characteristic that affects the acceptance of a tender is the margin applied to the base cost, in USD. **Figure 4.11** shows that there is an inflexion point located at a margin level of 8%. This means that a high number of tenders with a marginal benefit greater than this threshold will be lost, while most of the tenders with a marginal benefit lower than this threshold will be won. Although the margin is the main variable, other variables should also be taken into account in the predictive model.



**Figure 4.11.** Relation between the margin applied and the success of tenders

Finally, **Table 4.2** shows the best model we have been able to obtain using the real-life data available for this study. The model includes three variables: marginal benefit, level of relationship with the company offering the tender, and geographical zone. All in all, the model is able to predict with high

accuracy the result of the tender ( $R\text{-adj} > 76\%$ ) based on these variables. Notice that the decision maker has the possibility to influence at least on one of the variables, the benefit margin that is applied on the base cost.

Model	R	Radj	Deviance
result ~1	0.000000	0.000000	305.11077
result ~Margin	0.6899681	0.6888841	98.70451
result ~Margin + levelRelationShip	0.7524614	0.7507242	78.64421
result ~Margin + levelRelationShip + GeoZone	0.7674005	0.7641128	74.34025

*Table 4.2. Logistic regression model to estimate the probability of success*

## 5. Computational experiments and analysis of the results

This section will treat on the Smart Tenders Management with Artificial Intelligence project, focusing on the second stage of the project; ocean freight rates forecasting using Holt Winters' Method in RStudio. The steps will go from the reading of the data to the creation of a *low-risk* interval for ocean freight rates forecast, passing through plotting and decomposing the data, evaluating different forecasting methods and verifying, fitting, forecasting and analysing the chosen method; Holt Winters' Exponential Smoothing.

### 5.1 Introduction

The contribution of this project focuses on the development of a methodology of Stage 2 from the Smart Tenders Management with Artificial Intelligence project for tenders' price setting (Figure 5.1).



*Figure 5.1. Structure of the STeM-AI project, stage 2*

Winters' Method technique will be applied in this part of the project for Stage 2: Predicting Cost Evolution and Establishing Long-term Tariffs to create a methodology that solve the challenge providing solutions to support the decision of tariff to offer.

## 5.2 Reading Time Series Data

For computational experiments the first data set used is coming from shipping lines tariffs given to the logistic provider used for this case of study. Random component is avoided because results would not be according reality and it would not be possible to apply the model. However, a second data set is extracted from [Hanke et al. \(2001\) \[16\]](#), which uses the quarterly sales of saws from 2000 to 2006 for the Acme Tool Company to use predictive analytics as a business tool. The reason is that real data from the logistic provider is only recorded since 2015, and also that a key fact in ocean freight transportation during 2017 generates a peak, which would result in inaccurate forecasts.

The real first data set used, which appears in [Figure 5.2](#), represents the cost for a logistic provider to move a container of cargo from one point of origin to one point of destination without considering extra costs. Two lines appear in the plot almost overlapped, representing 20 and 40' containers. In total 104 periods are used, representing half-month rates for 4 years and 4 months, from 1st January of 2015 to 30th of April 2019. The different objects *myData20* and *myData40* are created to save the costs for the two container sizes after reading the data from the csv file.

However, as explained before, the data used will finally be the sales of saws in Acme Tool Company from [Hanke et al. \(2001\) \[16\]](#), which is shown in [Figure 5.3](#). In this case, a total of 28 periods are used, representing trimestral rates for 7 years, from 2012 to 2018. Only the object *myDataV2* is created to save what will be used from now on as ocean freight rates. The data can be read by typing:

```
myData_V2 <- read.csv(file="D:/Configuraciones/cristina.alvarez/Mis documentos/R/FCL
forecast/DATA/SeafreightData_V2.csv", header=TRUE, sep=";")

myDataV2 <- myData_V2$Cost
```

Once time series data is read into R, the next step is to store the data in a time series object in R, so that many functions for analysing time series data can be used. To store the data in a time series object, the *ts()* function in R is used. For example, to store the data in the variable *myData2V2timeseries* as a time series object in R, is done:

```
myData2V2timeseries <- ts(myDataV2, start = c(2012, 1), frequency = 4)
```

Sometimes the time series data set is collected at regular intervals that were less than one year, for example, monthly or quarterly. In this case, the number of times that data was collected per year can be specified by using the 'frequency' parameter in the `ts()` function. For trimestral time series data, frequency has been set at 4.

The first year and interval that the data was collected can also be specified by using the `start` parameter in `ts()` function. For example, if the first data point corresponds to the third quarter of 2015, `start` would be set as `start=c(2015,3)`.

And the output obtained from `myDataV2timeseries` is:

```
> myDataV2timeseries  
  
   Qtr1 Qtr2 Qtr3 Qtr4  
2012  500  350  250  400  
2013  450  350  200  300  
2014  350  200  150  400  
2015  550  350  250  550  
2016  550  400  350  600  
2017  750  500  400  650  
2018  850  600  450  700
```

### 5.3 Plotting Time Series

Once a time series is read into R, the next step is usually to make a plot of the time series data, which can be done with the `plot.ts()` function in R. For example, the following lanes plot the time series of the ocean freight transportation costs of 20' and 40' container in the same graph (Figure 5.2). The lines will be plotted in different colours and showing the different rate levels in two different y axis, on the left and right side.

```

plot.ts(myData20timeseries, ylab = "myData20timeseries", xlab = "Time (years)", col = "red", mar =
c(4, 4, 4, 4))

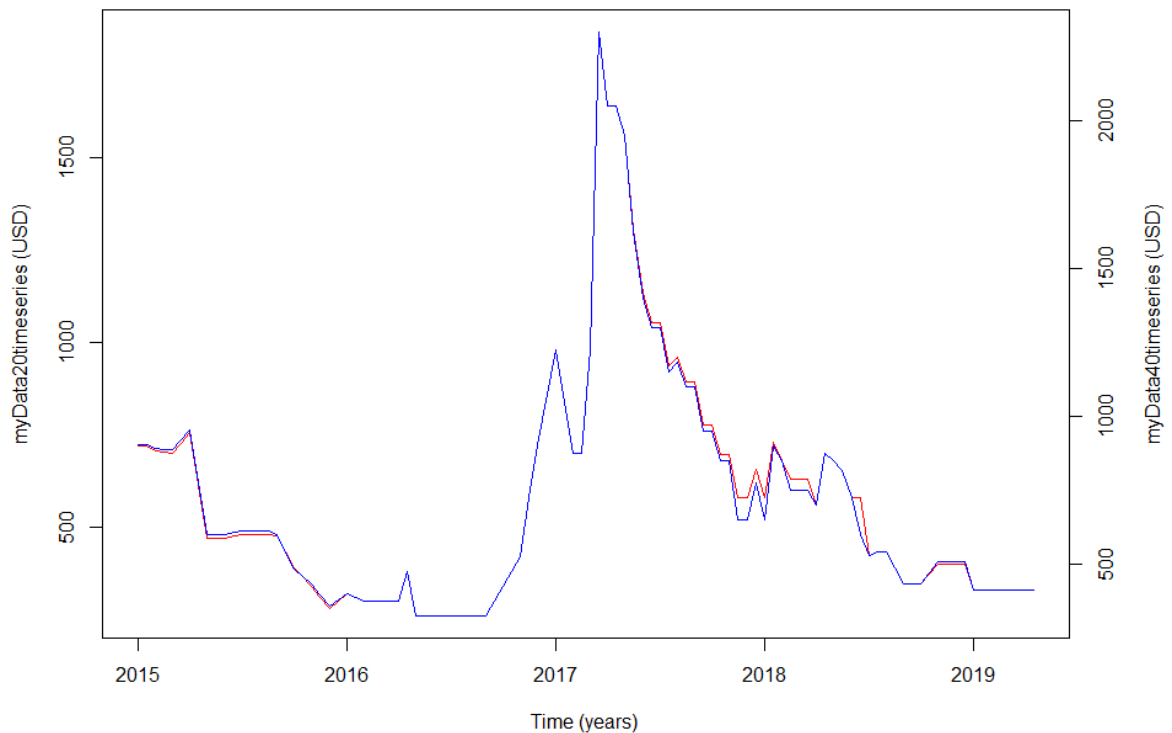
par(new=TRUE)

plot.ts(myData40timeseries, xaxt="n", yaxt="n", ylab = "", xlab = "", col = "blue", mar = c(4, 4, 4, 4))

mtext("myData40timeseries", side=4, line=3)

axis(side = 4)

```



*Figure 5.2. Ocean freight costs real data*

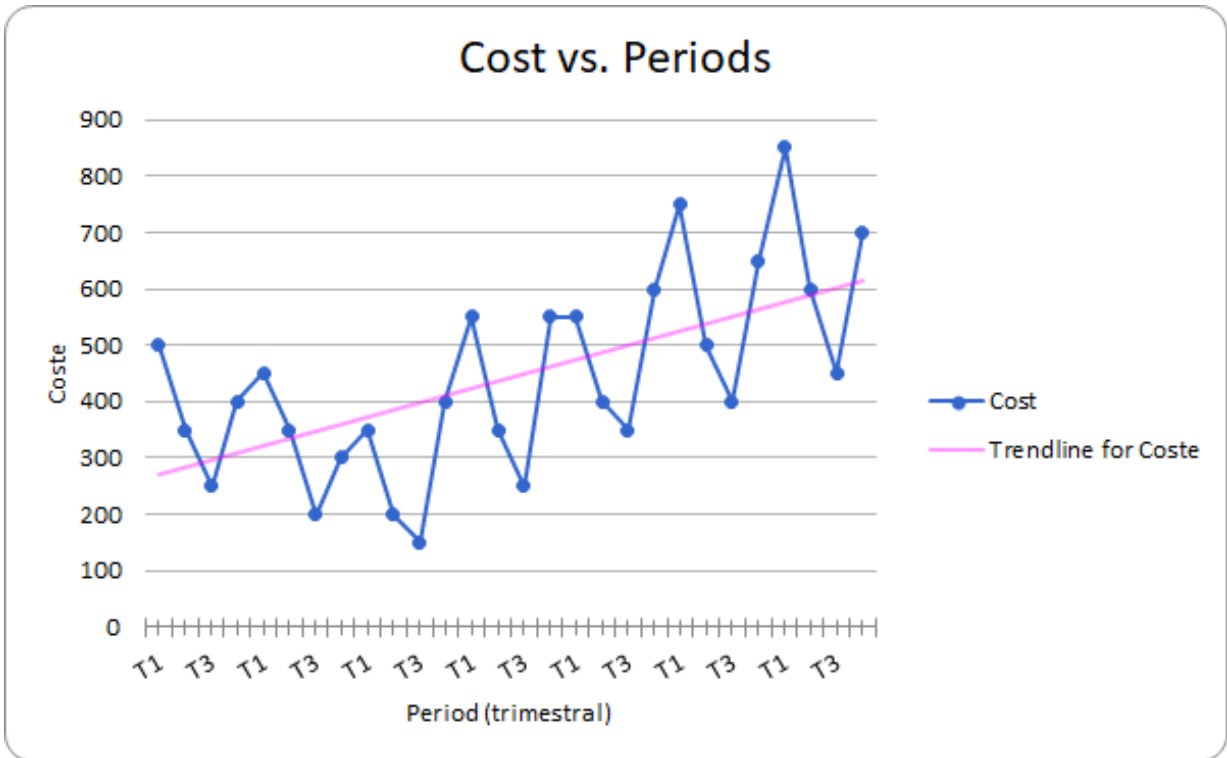
The real ocean freight data in [Figure 5.2](#) from Rotterdam to Shanghai port only includes years 2015 to 2019. Moreover, there is a peak in 2017 rates that wouldn't allow the Time Series method to provide accurate forecasts. The increase of the rates in 2017 is explained due to an increase in eastbound demand, driven by several factors, according Maersk carrier. These

included domestic policies in China stimulating demand and hence requiring raw materials (industrial input like chemicals, wood, paper and recyclables); a decline in Asian imports from Europe for the last 12-18 months now requiring re-stocking; and strong demand continuing in areas like food and beverages – commodities such as meat, milk powder and alcoholic beverages.

In adapted ocean freight rates plotted in **Figure 5.3** can be seen how data has a growing tendency (values keep increasing along time), which is confirmed with the trend-line option of Excel. A seasonal pattern can also be appreciated since the first and fourth quarters have relatively high values (repeated annually). Given the nature of the used data, it can be confirmed that Winters' technique is the most appropriate, since it will take seasonality into account.

```
par(mar = c(5, 5, 5, 5))
```

```
plot.ts(ylab = "myDataV2timeseries (USD)", xlab = "Time (years)", col= "blue",
myDataV2timeseries)
```



**Figure 5.3.** Ocean freight costs (adapted), excel



## 5.4 Decomposing Time Series

Decomposing a time series means separating it into its constituent components, which are usually a trend component and an irregular component, and if it is a seasonal time series, a seasonal component.

### 5.4.1 Decomposing non-seasonal data

A non-seasonal time series consists of a trend component and an irregular component. Decomposing the time series involves trying to separate the time series into these components, that is, estimating the trend component and the irregular component. This decomposition of non-seasonal data will not be further developed in this project because, as seen in [Figure 5.3](#), the nature of data used is seasonal.

### 5.4.2 Decomposing seasonal data

A seasonal time series consists of a trend component, a seasonal component and an irregular component. Decomposing the time series means separating the time series into these three components: that is, estimating these three components.

The function `decompose()` in RStudio first determines the trend component using a moving average (if filter is NULL, a symmetric window with equal weights is used), and removes it from the time series. Then, the seasonal figure is computed by averaging, for each time unit, over all periods. The seasonal figure is then centred. Finally, the error component is determined by removing trend and seasonal figure from the original time series.

How these three components interact determines the difference between a multiplicative and an additive time series. In a multiplicative time series, the components multiply together to make the time series. If the trend is increasing, the amplitude of seasonal activity increases. Everything becomes more exaggerated. In an additive time series, the components add together to make the time series. If the trend is increasing, the size of peaks and troughs throughout the time series will be similar. In this case ([Figure 5.3](#)) it can be considered that seasonality is multiplicative; as we are talking about data with seasonal pattern proportional to data ([Figure 4.10, page 48](#)).

To estimate the trend component and seasonal component of a seasonal time series that can be described using an additive or multiplicative, the `decompose()` function in R can be used. This function estimates the trend, seasonal, and irregular components of a time series

that can be described using an additive or multiplicative model, modifying the *type* parameter.

The function *decompose()* returns a list object as its result, where the estimates of the seasonal component, trend component and irregular component are stored in named elements of that list objects, called *seasonal*, *trend*, and *random* respectively.

For example, as discussed above and seen in [Figure 5.3](#), the time series of the ocean freight rates is seasonal with a peak every summer and trough every winter, and can probably be described using a multiplicative model since the seasonality is more highlighted over time.

The following command is used to estimate the trend, seasonal and irregular components of this time series:

```
myDataV2timeseriescomponents <- decompose(myDataV2timeseries, type = "multiplicative")
```

The estimated values of the seasonal, trend and irregular components are now stored in variables *myDataV2timeseriescomponents\$seasonal*, *myDataV2timeseriescomponents\$trend* and *myDataV2timeseriescomponents\$random*. The estimated values of the different components are printed by typing:

```
> myDataV2timeseriescomponents
```

```
$x
```

```
  Qtr1 Qtr2 Qtr3 Qtr4
```

```
2012 500 350 250 400
```

```
2013 450 350 200 300
```

```
2014 350 200 150 400
```

```
2015 550 350 250 550
```

```
2016 550 400 350 600
```

```
2017 750 500 400 650
```

2018 850 600 450 700

*\$seasonal*

Qtr1 Qtr2 Qtr3 Qtr4

2012 1.3404070 0.8907843 0.6335143 1.1352943

2013 1.3404070 0.8907843 0.6335143 1.1352943

2014 1.3404070 0.8907843 0.6335143 1.1352943

2015 1.3404070 0.8907843 0.6335143 1.1352943

2016 1.3404070 0.8907843 0.6335143 1.1352943

2017 1.3404070 0.8907843 0.6335143 1.1352943

2018 1.3404070 0.8907843 0.6335143 1.1352943

*\$trend*

Qtr1 Qtr2 Qtr3 Qtr4

2012 NA NA 368.75 362.50

2013 356.25 337.50 312.50 281.25

2014 256.25 262.50 300.00 343.75

2015 375.00 406.25 425.00 431.25

2016 450.00 468.75 500.00 537.50

2017 556.25 568.75 587.50 612.50

2018 631.25 643.75 NA NA

*\$random*

	<i>Qtr1</i>	<i>Qtr2</i>	<i>Qtr3</i>	<i>Qtr4</i>
2012	NA	NA	1.0701670	0.9719491
2013	0.9423689	1.1641842	1.0102376	0.9395508
2014	1.0189843	0.8553190	0.7892481	1.0249645
2015	1.0941950	0.9671684	0.9285272	1.1233759
2016	0.9118292	0.9579573	1.1049474	0.9832508
2017	1.0058994	0.9869066	1.0747209	0.9347571
2018	1.0045715	1.0463126	NA	NA

*\$figure*

[1] 1.3404070 0.8907843 0.6335143 1.1352943

*\$type*

[1] "multiplicative"

*attr(,"class")*

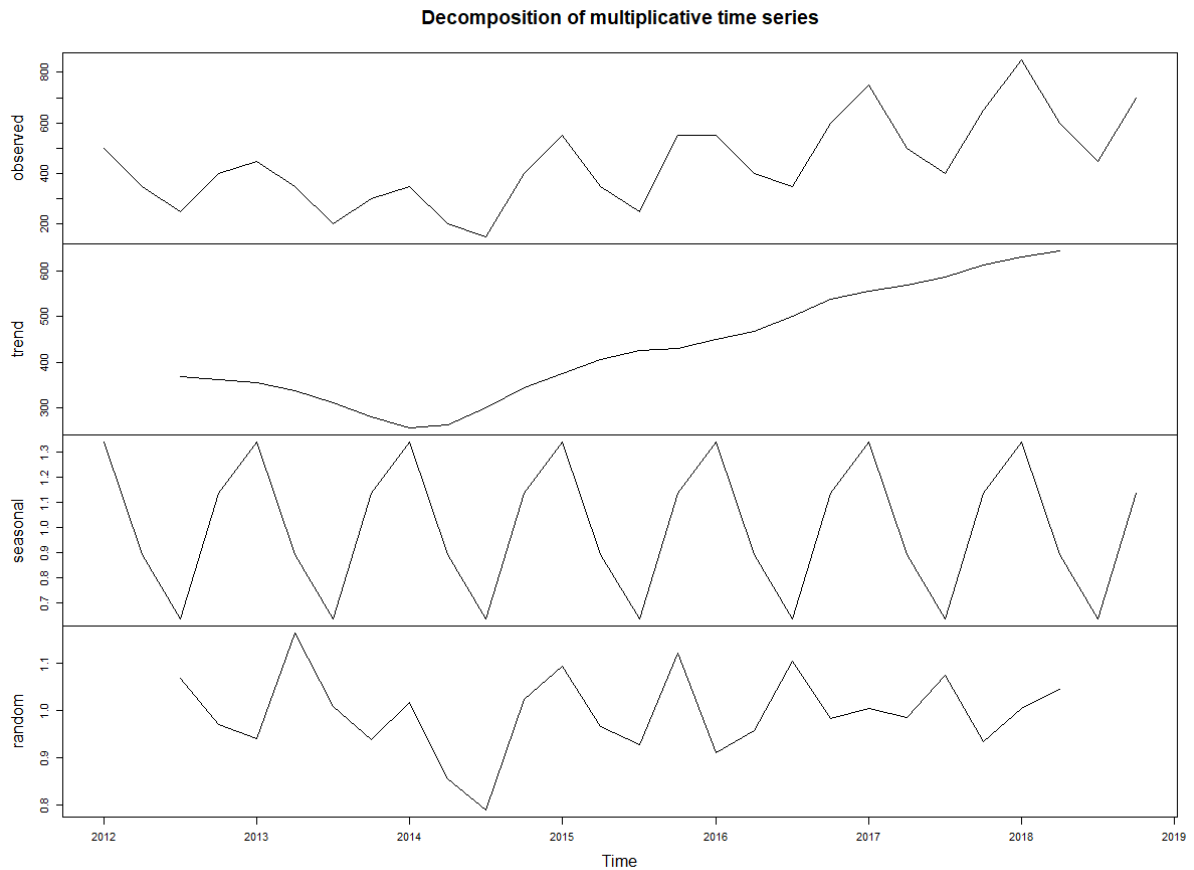
[1] "decomposed.ts"

The estimated seasonal factors are given for quarter 1 (Q1) to quarter 4 (Q4), and are the same for each year. The largest seasonal factor is for the first quarter (about 1.34), and the lowest is for the third (about 0.63), indicating that there seems to be a peak in ocean freight

rates in the first quarter and a trough in the third one each year.

The estimated trend, seasonal, and irregular components of the time series (Figure 5.4) can be plotted by using the `plot()` function, for example:

```
plot(myDataV2timeseriescomponents)
```



**Figure 5.4.** *Decomposition of multiplicative time series*

The plot above shows the original time series (top), the estimated trend component (second from top), the estimated seasonal component (third from top), and the estimated irregular component (bottom).

We see that the estimated trend component shows a decrease from 2012 to 2014, followed by an increase from then on to 2018. The seasonality is very stable in her peaks during the analysed years and the randomness varies from 0,78 to 1,16.

### 5.4.3 Seasonally adjusting

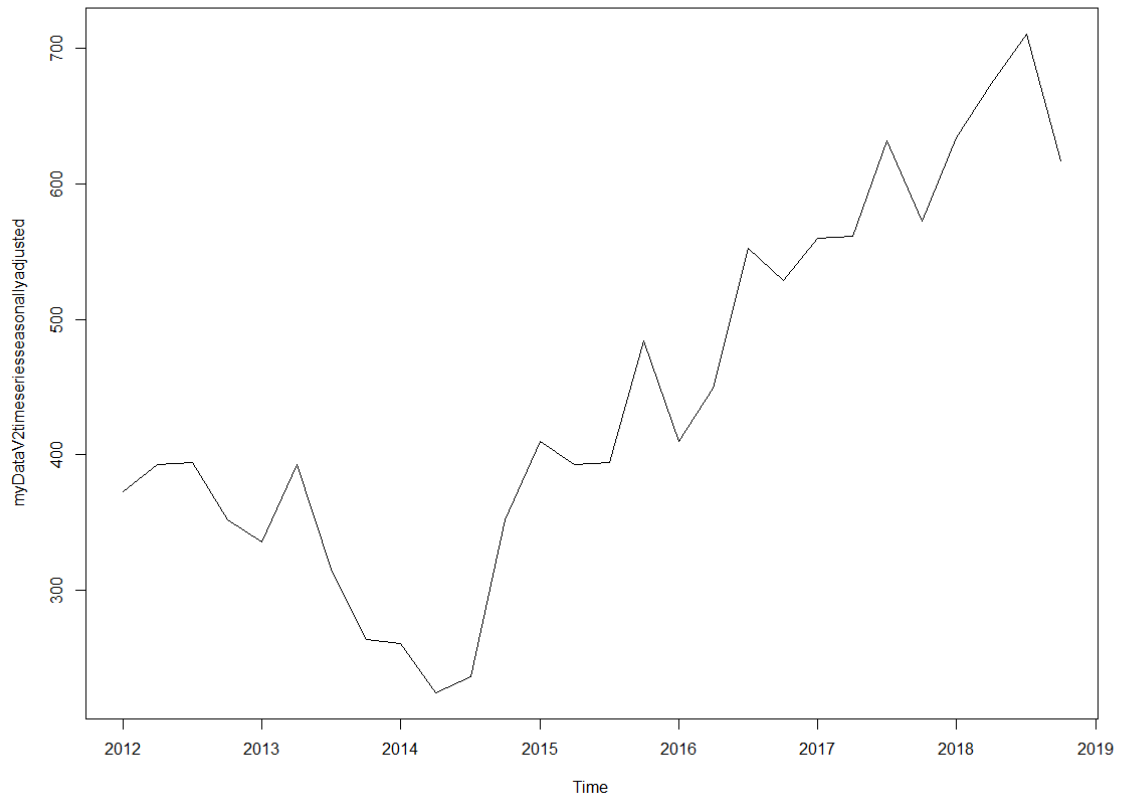
A time series can be adjusted by estimating the seasonal component and dividing the estimated seasonal component from the original time series if the concrete time series can be described is seasonal and can be described using a multiplicative model. To do it the estimate of the seasonal component calculated by the *decompose()* function can be used.

For example, to seasonally adjust the time series of the ocean freight rates, *decompose()* function is used and then the seasonal component is divided from the original time series:

```
myDataV2timeseriesseasonallyadjusted <- myDataV2timeseries /  
myDataV2timeseriescomponents$seasonal
```

Then, the seasonally adjusted time series can be plotted (Figure 5.5) by using the *plot()* function, by typing:

```
plot(myDataV2timeseriesseasonallyadjusted)
```



*Figure 5.5. Ocean freight time series seasonally adjusted*

Figure 5.5 shows how the seasonal variation has been removed from the seasonally adjusted time series. The seasonally adjusted time series now just contains the trend component and an irregular component.

## 5.5 Forecast using Exponential Smoothing

Exponential smoothing can be used to make short-term forecasts for time series data.

### 5.5.1 Simple Exponential Smoothing

If the time series that could be described using an additive model with constant level and no seasonality, simple exponential smoothing can be used to make short-term forecasts.

The simple exponential smoothing method provides a way of estimating the level at the current time point. Smoothing is controlled by the parameter alpha; for the estimate of the level

at the current time point. The value of alpha lies between 0 and 1. Values of alpha that are close to 0 mean that little weight is placed on the most recent observations when making forecasts of future values. As it is not the case of ocean freight rates, Simple Exponential Smoothing will not be further developed.

### 5.5.2 Holt's Exponential Smoothing

When time series can be described using an additive model with increasing or decreasing trend and no seasonality, Holt's exponential smoothing can be used to make short-term forecasts.

Holt's exponential smoothing estimates the level and slope at the current time point. Smoothing is controlled by two parameters, alpha, for the estimate of the level at the current time point, and beta for the estimate of the slope  $b$  of the trend component at the current time point. As with simple exponential smoothing, the parameters alpha and beta have values between 0 and 1, and values that are close to 0 mean that little weight is placed on the most recent observations when making forecasts of future values. Like in the previous 5.5.1 Simple Exponential Smoothing method, it is not the case of ocean freight rates, si Holt's Exponential Smoothing will neither be further developed.

### 5.5.3 Holt-Winters' Exponential Smoothing

To make short-term forecasts of time series that can be described using an additive model with increasing or decreasing trend and seasonality, Holt-Winters' Exponential Smoothing can be used.

Holt-Winters' Exponential Smoothing estimates the level, slope and seasonal component at the current time point. Smoothing is controlled by three parameters: alpha, beta, and gamma, for the estimates of the level, slope  $b$  of the trend component, and the seasonal component, respectively, at the current time point. The parameters alpha, beta and gamma all have values between 0 and 1, and values that are close to 0 mean that relatively little weight is placed on the most recent observations when making forecasts of future values. Therefore, the ocean freight rates time series can be described using an additive model with a trend and seasonality.



### 5.5.3.1 Verification of the application of the Winters' method

Before establishing long-term tariffs, it is required to know how costs will evolve based on historical data previously read and plotted. Before using Winter's Method's predictions, it should be demonstrated that the prediction is reliable. Therefore, in this first experiment results from RStudio software for the last year 2018 will be compared to the real historical values. To do so, the 2018 data is deleted from a new data set, and the following lines are typed:

```
myData_V3 <- read.csv(file="D:/Configuraciones/cristina.alvarez/Mis documentos/R/FCL
forecast/DATA/SeafreightData_V3.csv", header=TRUE, sep=";")

myDataV3 <- myData_V3$Cost

myDataV3timeseries <- ts(myDataV3, start = c(2012, 1), frequency = 4)

HoltWinters(myDataV3timeseries, seasonal = "multiplicative")

myDataV3timeseriesforecasts <- HoltWinters(myDataV3timeseries)

library("forecast")

myDataV3timeseriesforecasts2 <- forecast(myDataV3timeseries, h=4)

> myDataV3timeseriesforecasts2

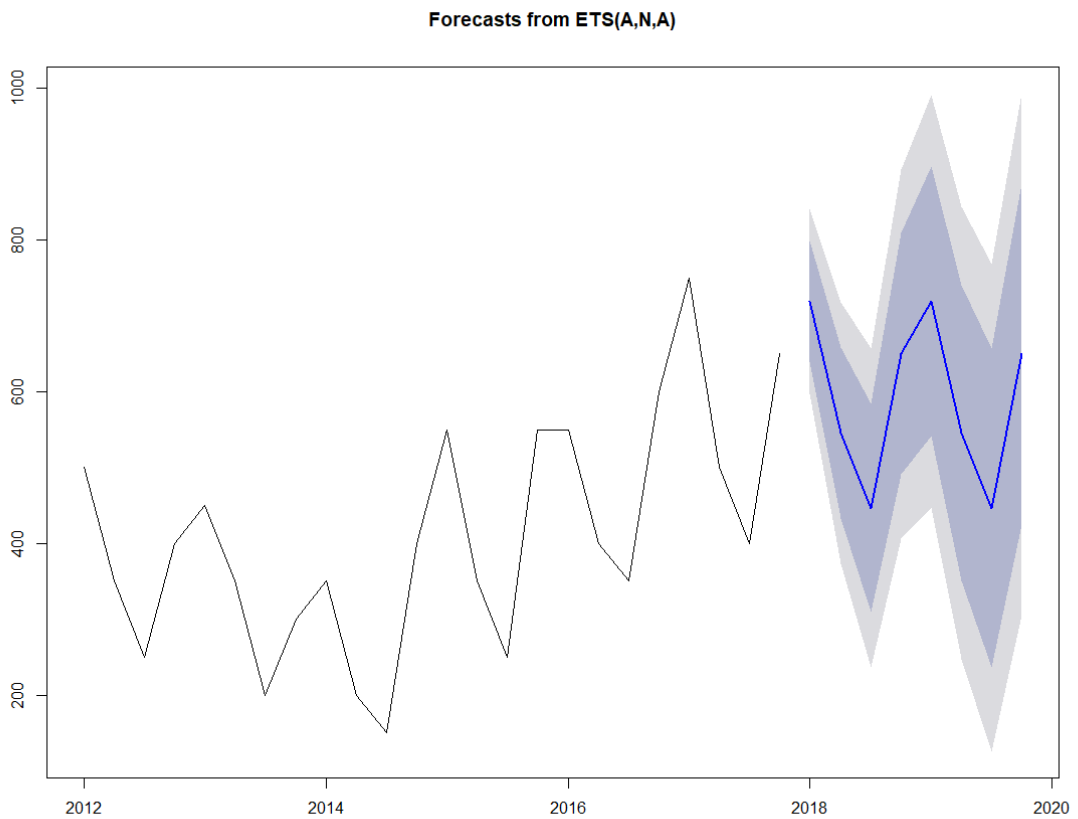
      Point Forecast  Lo 80  Hi 80  Lo 95  Hi 95
2018 Q1    718.8660 639.3247 798.4072 597.2182 840.5138
2018 Q2    545.3439 432.9081 657.7798 373.3882 717.2997
2018 Q3    446.5978 308.9140 584.2816 236.0287 657.1669
2018 Q4    649.9999 490.9917 809.0081 406.8179 893.1820
```

The following [Table 5.1](#) and [Figure 5.6](#) show the quarterly real costs for year 2018 and

the 80% and 95% interval confidences. Both in 80 and 95%, all the forecasted values are within the intervals, except for the first quarter. Therefore, as expected the 95% confidence interval assumes less risk, as the forecasted value which lies out of the interval is closer to the 95% interval upper bound. For sure, it can also be concluded that the ocean freight costs evolution is highly unpredictable and probably some months or quarters where the offer will be lower to the cost will be compensated with the majority of cases where there will be a benefit margin.

Year	Quarter	Real data	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2018	Q1	850	718,866	639,3247	798,4072	597,2182	840,5138
2018	Q2	600	545,3439	432,9081	657,7798	373,3882	717,2997
2018	Q3	450	446,5978	308,914	584,2816	236,0287	657,1669
2018	Q4	700	649,9999	490,9917	809,0081	406,8179	893,182

*Table 5.1. Forecast and confidence intervals 2018*



*Figure 5.6. Ocean freight forecast using Holt Winters' Method 2018-2019*

### 5.5.3.2 Model fit

Once the methodology has been tested in the previous point, next year ocean freight rates can also be forecasted. Before the forecasting itself, the *HoltWinters()* function is used to fit the predictive model:

```
> myDataV2timeseriesforecasts <- HoltWinters(myDataV2timeseries)

> myDataV2timeseriesforecasts

Holt-Winters exponential smoothing with trend and additive seasonal component.

Call:
HoltWinters(x = myDataV2timeseries)

Smoothing parameters:

alpha: 0.4745088

beta : 0.1384328

gamma: 1

Coefficients:

[1]

a 603.22988

b 12.87664

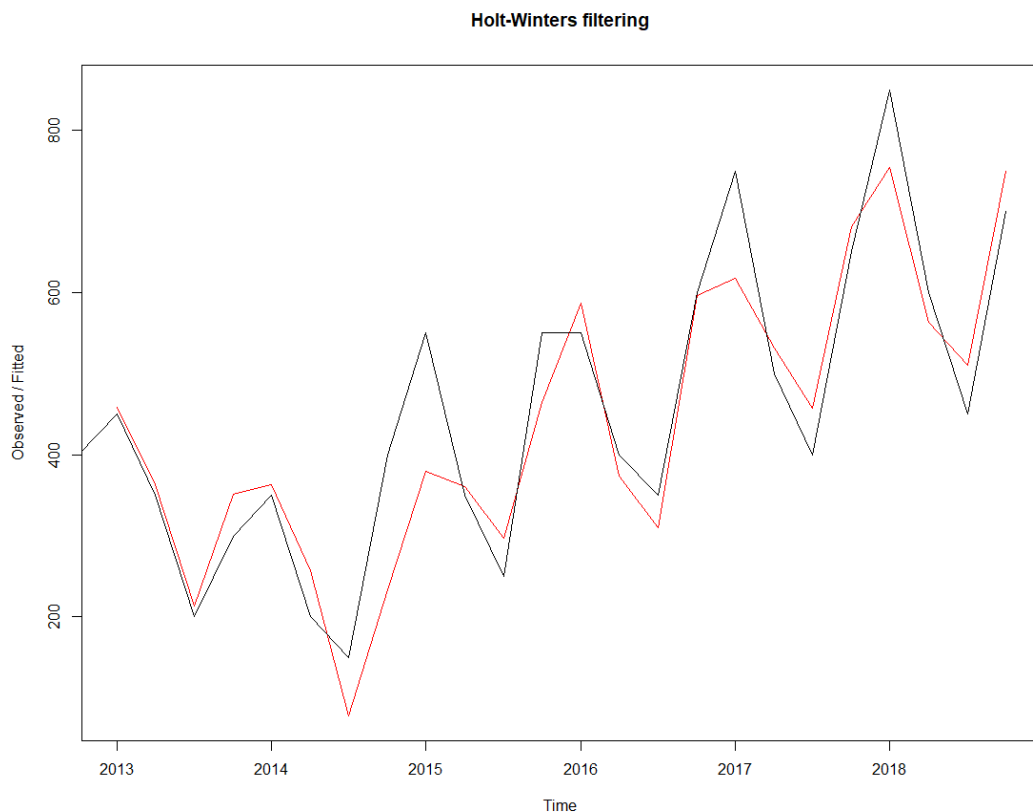
s1 265.99678

s2 -19.21102

s3 -160.54602
```

The estimated values for ocean freight rates of alpha, beta and gamma are 0.4745088, 0.1384328, and 1, respectively. The value of alpha is very high or low, indicating that the estimate of the level at the current time point is based upon both more and less recent observations. The low value of beta indicates that the estimate of the slope  $b$  of the trend component is based mostly on not recent observations. The value of gamma indicates that the estimate of the seasonal component at the current time point places most of the weight on recent observations when making forecasts of future values. Finally, coefficients are in a vector with named components  $a$ ,  $b$ ,  $s_1$ , ...,  $s_p$  containing the estimated values for the level, trend and seasonal components.

**Figure 5.7** shows the plot of the original time series as a black line, with the forecasted values as a red line on top of that:



**Figure 5.7.** Ocean freight Holt-Winters' fitting

The plot in [Figure 5.7](#) shows that the Holt-Winters' exponential method is very successful in predicting the seasonal peaks, which occur roughly in summer every year. To make forecasts for future times not included in the original time series, the `forecast.HoltWinters()` function is used in the `forecast` package. In the case of study, the data for the ocean freight from beginning of 2012 to the end of 2018.

### 5.5.3.3 Forecast

If the goal is to make forecasts for the following 2 years (8 trimesters), and plot the forecasts:

```
library("forecast")

HoltWinters(myDataV2timeseries, seasonal = "multiplicative")

myDataV2timeseriesforecasts <- HoltWinters(myDataV2timeseries)

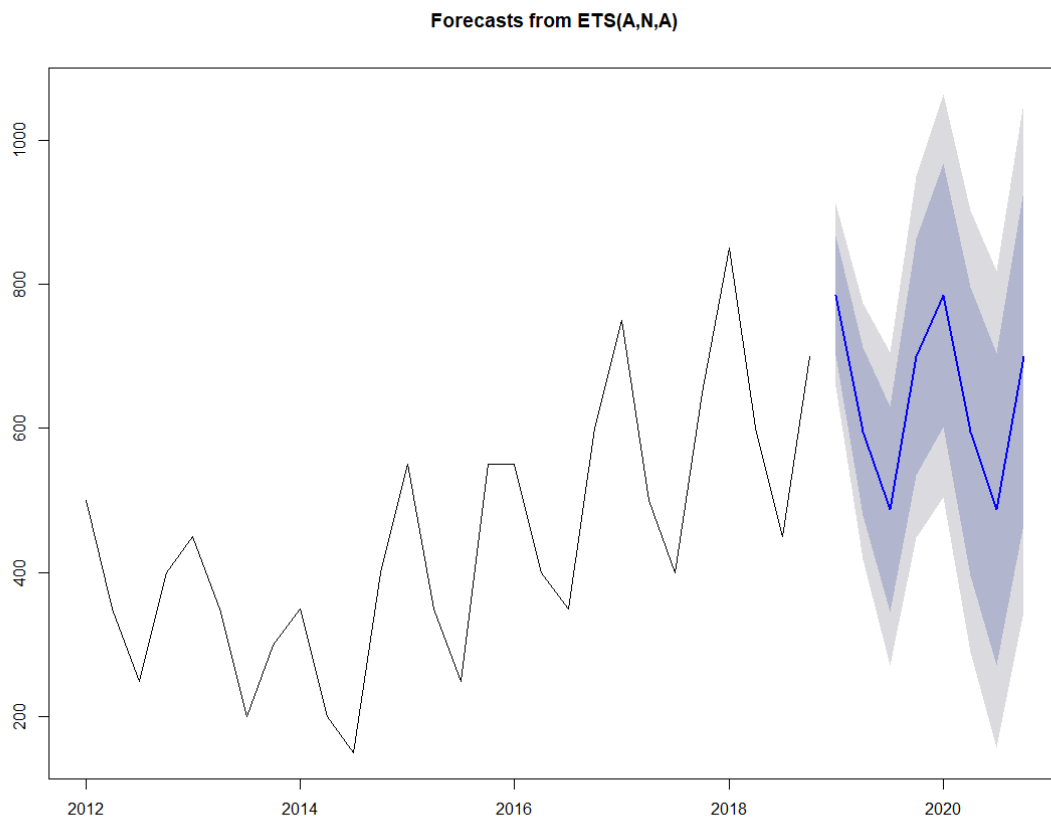
myDataV2timeseriesforecasts2 <- forecast(myDataV2timeseries, h=8)

plot(myDataV2timeseriesforecasts2)
```

The [Table 5.2](#) shows the forecasted values and confidence intervals for 80% and 95% for ocean freight rates during 2019 using Holt Winters' method. The forecasts are shown as a blue line in [Figure 5.8](#), and the other blue shaded areas show 80% and 95% prediction intervals, respectively. As has been commented during the analysis of the data, the seasonality is reflected in peaks in summer and low points in winter. On the other hand, the trend is what makes 2019 rates go higher with respect to 2018 rates.

Year	Quarter	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2019	Q1	784,3468	701,5705	867,123	657,7514	910,9421
2019	Q2	595,6714	479,5194	711,8234	418,0322	773,3105
2019	Q3	488,1365	346,2539	630,0191	271,1458	705,1272
2019	Q4	699,409	535,7893	863,0286	449,1742	949,6437

*Table 5.2. Forecast and confidence intervals 2019*



*Figure 5.8. Ocean freight forecast 2019-2020 using Holt Winters' Method*

### 5.5.3.4 Confidence intervals

The confidence intervals represent the values within the unknown value can be found, providing a security level equal to the defined percentage (generally 95%) that the real value will be found within the interval. It has to be taken into account that such interval (Figure 5.8) becomes wider while the prediction goes to a longer term. How is it possible to make confidence intervals narrower? As more historical data is included, the model is better adjusted to reality, because it has more information in which can be based. Therefore, as more historical data, better predictions can be done and narrower the interval becomes. This experiment will show how the quantity of data affects the confidence intervals. The same data will be used but, in this case, historical data will be added in blocks of 7 periods.

To perceive clearly the differences, the average of both bounds that form the confidence interval has been calculated. Next, the distance between the lower and upper bounds of the interval is calculated, has shown in Table 5.3.

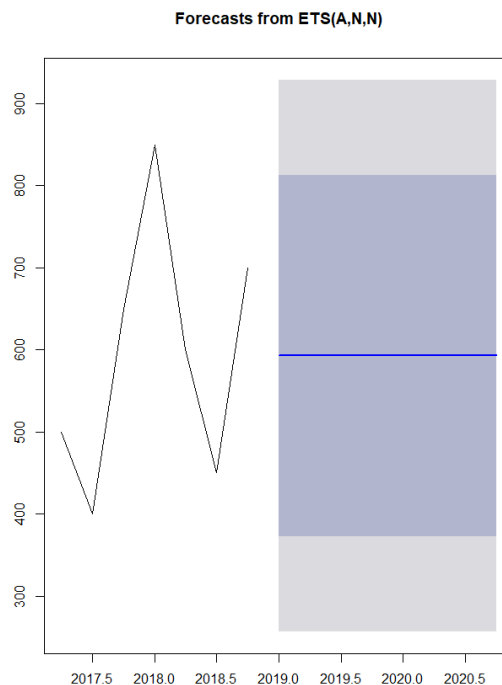
	Period of the forecast	Forecast	Lo 95	Hi 95	Difference
7 Periods	29	592,573	256,3285	928,8175	672,489
	30	592,573	256,3285	928,8175	672,489
	31	592,573	256,3285	928,8175	672,489
	32	592,573	256,3285	928,8175	672,489
	<b>Average</b>			<b>256,3285</b>	<b>928,8175</b>
14 Periods	29	813,1686	664,528	961,8092	297,2812
	30	581,208	393,6409	768,7751	375,1342
	31	441,1774	221,4758	660,879	439,4032
	32	703,2861	455,5775	950,9947	495,4172
	<b>Average</b>			<b>433,80555</b>	<b>835,6145</b>
21 Periods	29	861,5234	696,2009	1026,8459	330,645
	30	642,9912	517,3695	768,6129	251,2434
	31	552,4161	439,282	665,5502	226,2682
	32	803,4711	639,3838	967,5584	328,1746
	<b>Average</b>			<b>573,05905</b>	<b>857,14185</b>
28 Periods	29	784,3468	657,7514	910,9421	253,1907
	30	595,6714	418,0322	773,3105	355,2783
	31	488,1365	271,1458	705,1272	433,9814
	32	699,409	449,1742	949,6437	500,4695
	<b>Average</b>			<b>449,0259</b>	<b>834,755875</b>

*Table 5.3. Confidence intervals of the forecasts for different sizes of data sets*

To conclude with this experiment, as commented before, as more data available, as narrower the confidence interval becomes, which makes the prediction much more reliable in regards to what will happen in reality. It has to be highlighted that the confidence interval of the

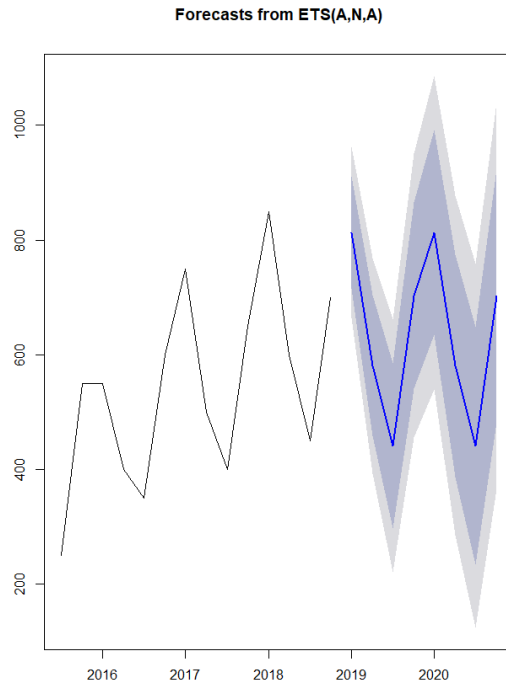
forecast with 21 periods has been clearly lower than the rest, including the 28 periods interval. It can be seen in table 5.3, where the interval bounds are closer to the values forecasted than in the rest of the figures. This fact occurs because in this case the model has been better adjusted, with better parameters for this data set generating a more suitable model. Even with this exception, it can be confirmed that the intervals tend to become narrower with more historical data.

The previous statements are supported by the following graphs (Figures 5.9-5.12), where the narrowest interval is the one from Figure 5.11, which uses 21 periods to generate the forecast and confidence interval. The case with 7 periods in Figure 5.9 leads to a straight forecast and confidence interval due to the fact that 7 periods is less than 2 complete series or years, as each series is composed by 4 trimesters.

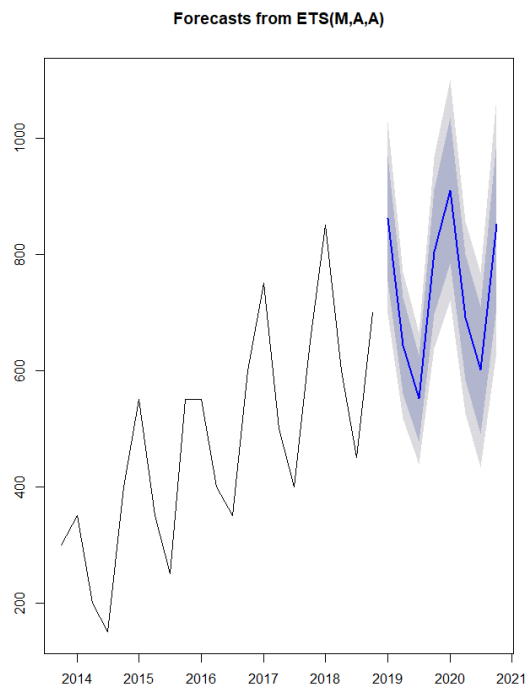


*Figure 5.9. Winters' Method forecast using 7 periods*

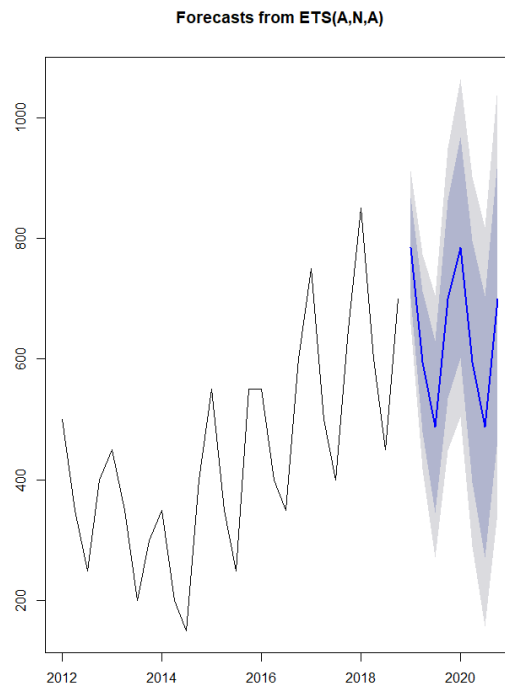




*Figure 5.10. Winters' Method forecast using 14 periods*



*Figure 5.11. Winters' Method forecast using 21 periods*



*Figure 5.12. Winters' Method forecast using 28 periods*

### 5.5.3.5 Finding the best fit of the model

In the previous experiment it was shown how a model with more data and better fitted (more adequate parameters) generates much more reliable results. For this reason, this third experiment will find the best fit of the model for our data. The fit of the model can be evaluated by: (a) analysing if in the graph that generates RStudio the data is overlapped (for example, in [Figure 5.7](#) from the first experiment the real values and the fitted are not overlapped specially in years 2014 and 2016, so it is not totally accurately fitted) and (b) the error measures MAPE, MAD and MSD, which smaller values imply better fit. In this third experiment, it will be found the best fit of the model through the subjective selection of parameters and also through the application of Excel Solver to optimize and compare the results.

The initial reference in the [subjective selection](#), the parameters  $\alpha = 0.4$ ,  $\beta = 0.1$  and  $\gamma = 0.3$  will be used ([Figure 5.13](#)). The reason is that they are used in [Hanke et al. \(2001\) \[16\]](#) for the sales of saws in Acme Tool Company. Keeping the rest of the parameters unmodified; increasing  $\alpha$  ([Figure 5.14](#)) the result improves, increasing  $\beta$  ([Figure 5.15](#)) the

result worsens and increasing gamma (Figure 5.16) the result improves. Therefore, a logical subjective selection would be to increase alpha and gamma (Figure 5.17) The concrete values can be seen in Table 5.4, including also the RStudio optimized selection parameters (Figure 5.18)

Parameters	Parameters values	SSE	MSD	MAD	MAPE
Initial	alpha 0,4; beta 0,1; gamma 0,3	137262,2	4902,221429	70,01586555	15,81003416
Alpha increase	alpha 0,9; beta 0,1; gamma 0,3	134512,3	4804,010714	69,3109711	16,30846379
Beta increase	alpha 0,4; beta 0,9; gamma 0,3	139673,7	4988,346429	70,62822685	17,12199439
Gamma increase	alpha 0,4; beta 0,1; gamma 0,9	122269,1	4366,753571	66,08141623	16,37415623
Alpha and gamma increase	alpha 0,9; beta 0,1; gamma 0,9	119783,1	4277,967857	65,40617599	16,80158649
RStudio optimization	alpha 0,7994528; beta 0,0796615; gamma 0,9536164	114487	4088,821429	63,94389282	16,42595412

Table 5.4. Measure errors for different parameter combinations

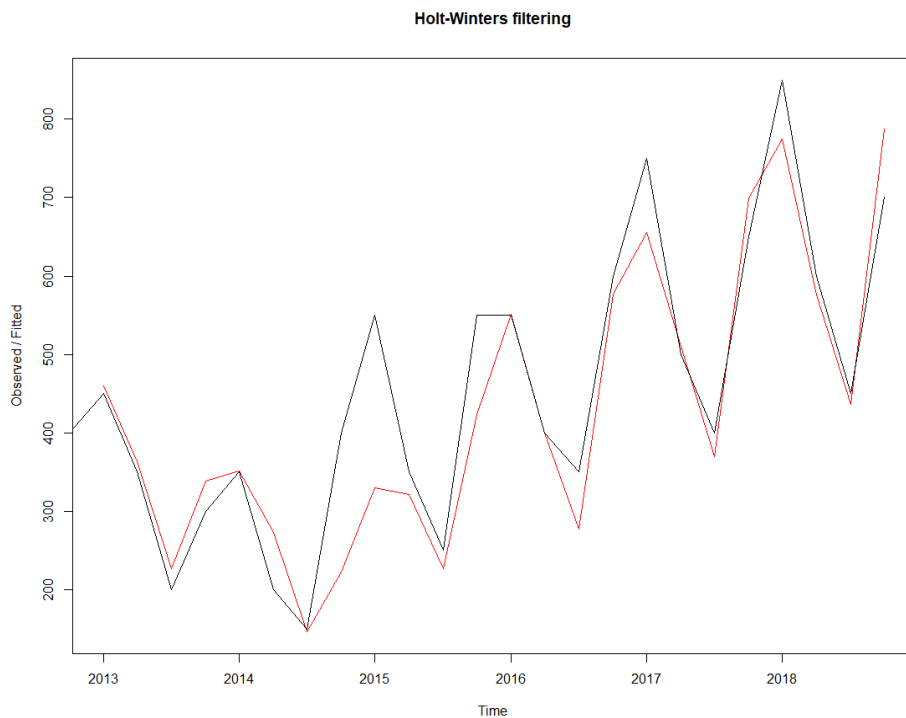


Figure 5.13. Forecast graph (alpha = 0.4, beta = 0.1, gamma = 0.3)



*Figure 5.14. Forecast graph (alpha = 0.9, beta = 0.1, gamma = 0.3)*



*Figure 5.15. Forecast graph (alpha = 0.4, beta = 0.9, gamma = 0.3)*

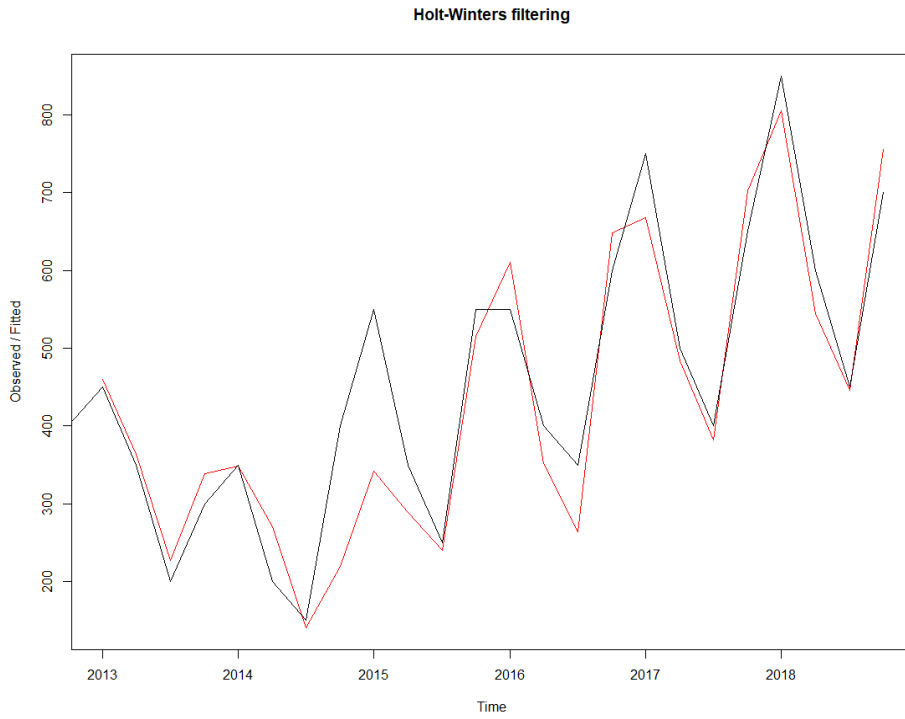


Figure 5.16. Forecast graph ( $\alpha = 0.4$ ,  $\beta = 0.1$ ,  $\gamma = 0.9$ )

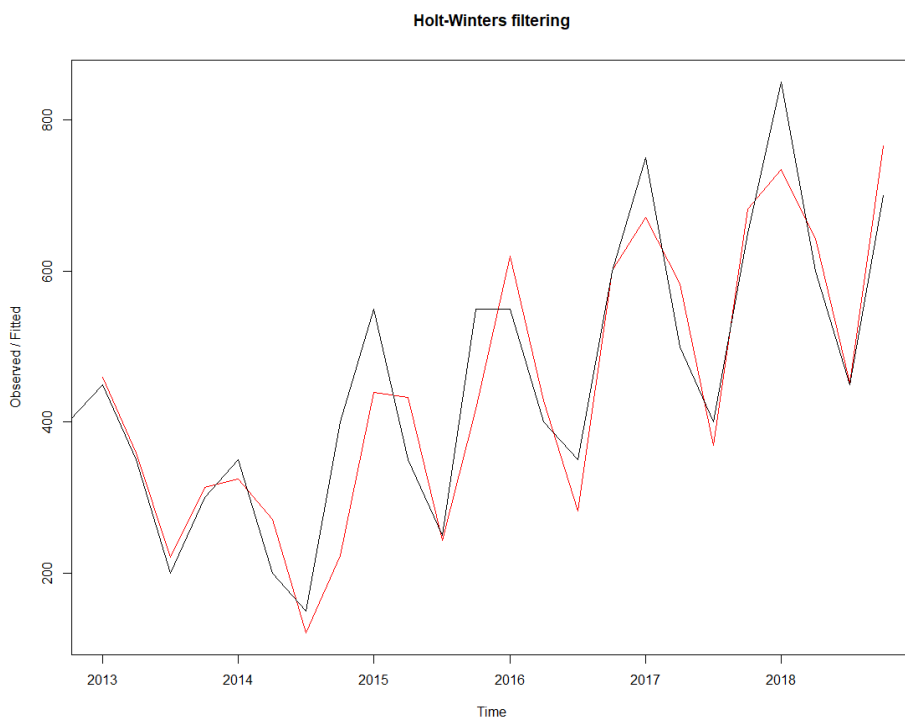
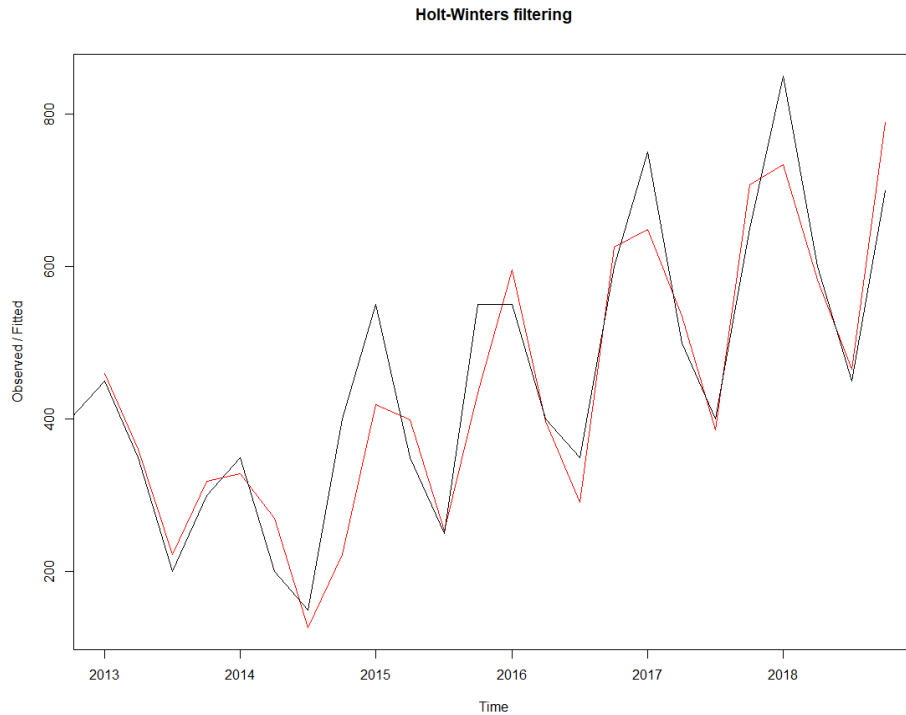


Figure 5.17. Forecast graph ( $\alpha = 0.9$ ,  $\beta = 0.1$ ,  $\gamma = 0.9$ )



*Figure 5.18. Forecast graph RStudio optimized parameters*

To conclude, the low-risk interval generation for the tender tariff, firstly the prediction needs to be checked. At this point, it can be concluded that with the actual data, the parameters that minimize the error, optimized by RStudio will be used (alpha 0,7994528; beta 0,0796615; gamma 0,9536164). It is important to highlight that, for each data set, a different model will be required, and therefore, different parameters and fit. In this section it has also been concluded that, in general terms, the more historical data available, the more precise predictions the model will provide

#### 5.5.3.6 Ljung-Box test

It is interesting to investigate if predictive model can be improved upon by checking whether the in-sample forecast errors show non-zero autocorrelations at lags 1-20, by making a correlogram and carrying out the Ljung-Box test:

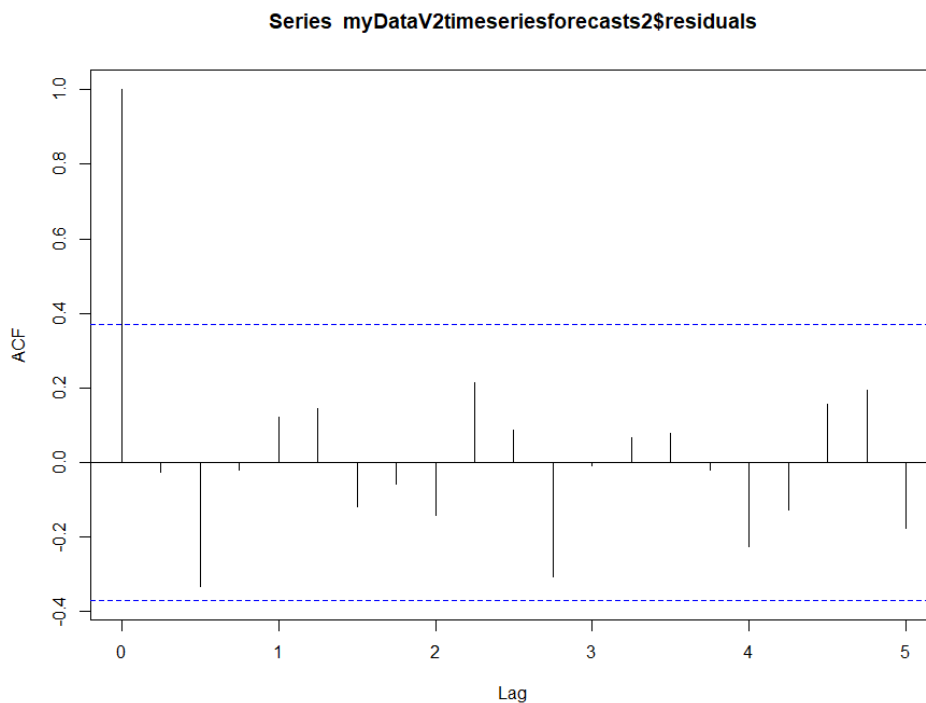
```
> acf(myDataV2timeseriesforecasts2$residuals, lag.max=20)
```

```
> Box.test(myDataV2timeseriesforecasts2$residuals, lag=20, type="Ljung-Box")
```

*Box-Ljung test*

*data: myDataV2timeseriesforecasts2\$residuals*

*X-squared = 27.802, df = 20, p-value = 0.1142*



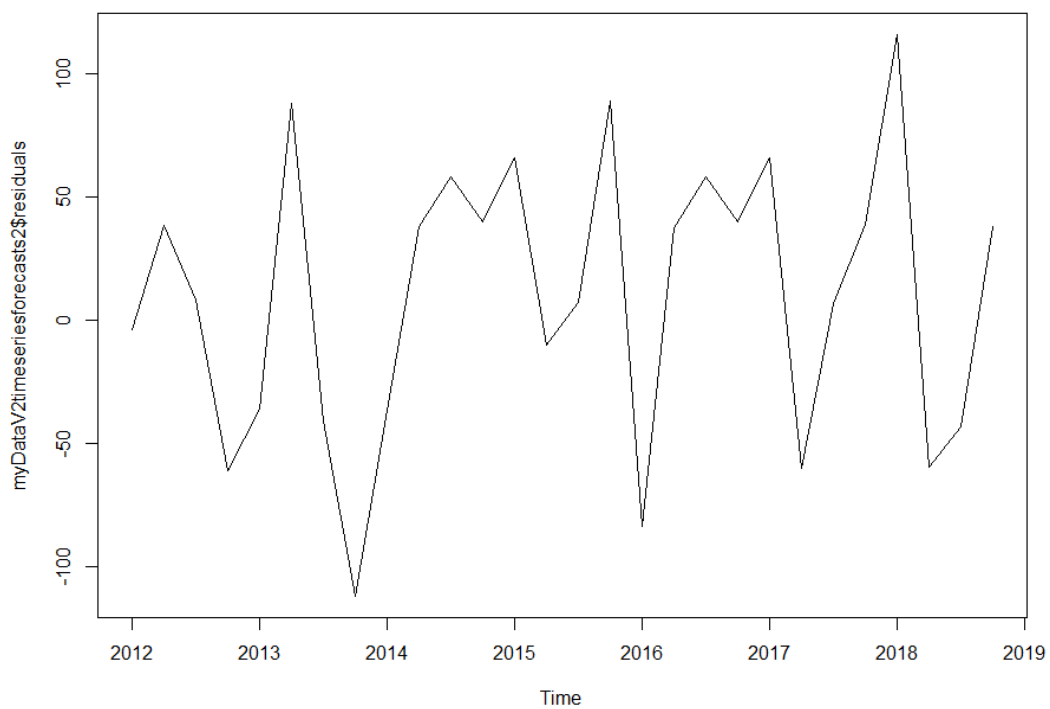
*Figure 5.19. Ljung-Box test, ocean freight forecast residuals*

The correlogram in [Figure 5.19](#) shows that the autocorrelations for the in-sample forecast errors do not exceed the significance bounds for lags 1-20. Furthermore, the p-value for Ljung-Box test is 0.1142. If the p-value is higher than a selected level of significance (in this case 0.05) there is little evidence of non-zero autocorrelations at lags 1-20. That is, there is little evidence to reject the null hypothesis.

We can check whether the forecast errors have constant variance over time, and are

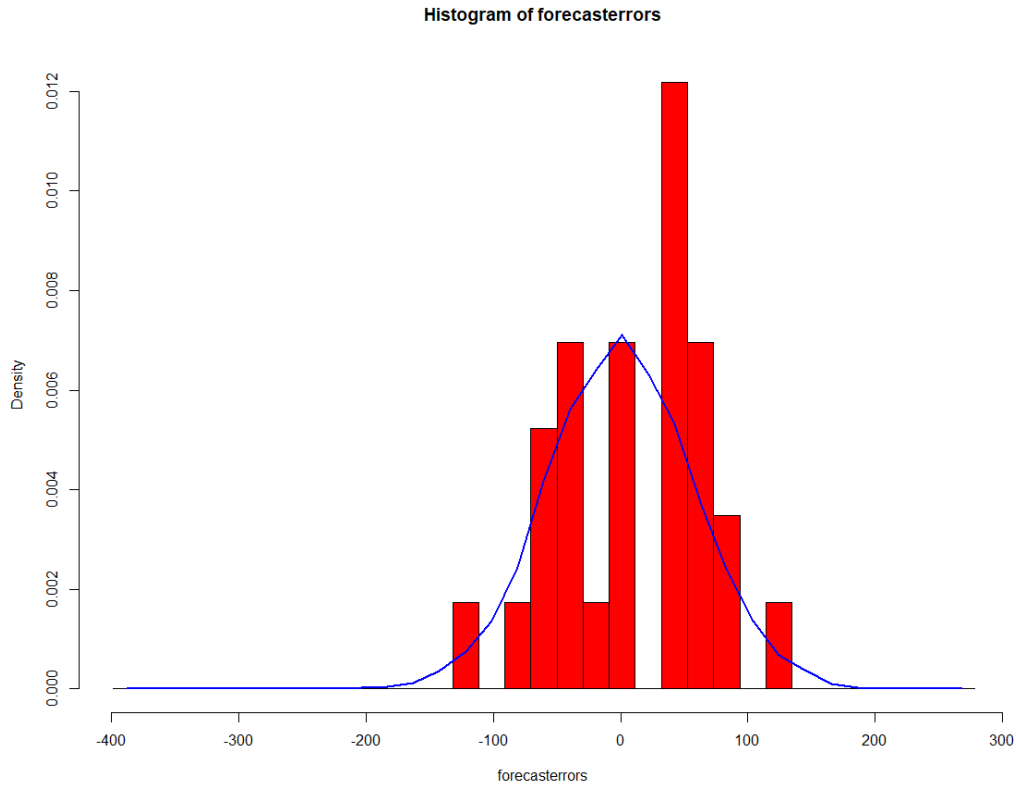
normally distributed with mean zero, by making a time plot of the forecast errors (Figure 5.20) and a histogram (with overlaid normal curve) (Figure 5.21):

```
plot.ts(myDataV2timeseriesforecasts2$residuals) # make a time plot  
plotForecastErrors(myDataV2forecasts2$residuals) # make a histogram
```



*Figure 5.20.* Time plot of the forecast errors





*Figure 5.21.* Histogram of the forecast errors

From the time plot, it appears plausible that the forecast errors have constant variance over time. From the histogram can be seen that the distribution of forecast errors is roughly centred on zero, and is more or less normally distributed, although it seems to be slightly skewed to the right compared to a normal curve. However, the right skew is relatively small, and so it is plausible that the forecast errors are normally distributed with mean zero.

Thus, there is little evidence of autocorrelation at lags 1-20 for the forecast errors, and the forecast errors appear to be normally distributed with mean zero and constant variance over time. This suggests that Holt-Winters' exponential smoothing provides an adequate predictive model of the ocean freight rates, which probably cannot be improved upon.

## 5.6 Establishing low-risk intervals for long-term tariffs for tenders

Once data has been processed to obtain the ocean freight costs evolution for the following months, results will be used to compute the "low-risk" intervals for the average cost during the time horizon planned (generally between 6 and 18 months in the case of tenders) as

follows:

- The **minimum value** of the interval is calculated with the average of the forecasted values.
- The **maximum value** of the interval is calculated as the average of the values that define the upper bound of the 95% confidence intervals associated to each forecasted value.

Actually, the forecasted values provide support information for decision-making by themselves. However, it is important so realize that these value by themselves are not reliable to fix a tariff, because it cannot be ensured 100% that the cost will have that value in that instant, and therefore the scenario wouldn't fit the reality, where events that the company cannot control can happen. In such cases, the cost could increase unexpectedly and therefore the forecast would be obsolete. It would neither be appropriate to choose a lower value for the fixation of tariffs' interval lower bound, because it would be an optimistic thought in a future scenario where it is more likely that the costs go down, which would neither fit reality. For that reason, the chosen range for the interval which will support the tariffs' fixation ensures the risk. In other words, creates a margin to face a possible increase in future costs in real life due to an unexpected reason such as the fuel increase or a peak on demand not programmed. The *low risk* interval obtained for the cost of the tariff defined by the averages of the values obtained during the forecast can be seen in [Figure 5.22](#). That interval will provide support to decisions as it will define a range of long-term costs, that, with a later add of the marginal benefit desired, will enable to establish the tariff for a tender.

It has to be taken into account that, as closer the price is to the lower bound of the interval, the more competitive it will be, but it also implies to assume more risk. On the opposite, as closer the price is to the upper bound, less competitive and less risky it will be.

```
minvalue <- mean(myDataV3timeseriesforecasts2$mean[1:4])  
  
maxvalue <- mean(myDataV3timeseriesforecasts2$upper[9:12])  
  
  
plot(myDataV3timeseriesforecasts2)  
  
abline(h=minvalue, col="green")  
  
abline(h=maxvalue, col="green")
```

```
> minvalue
```

```
[1] 590.2019
```

```
> maxvalue
```

```
[1] 777.0406
```

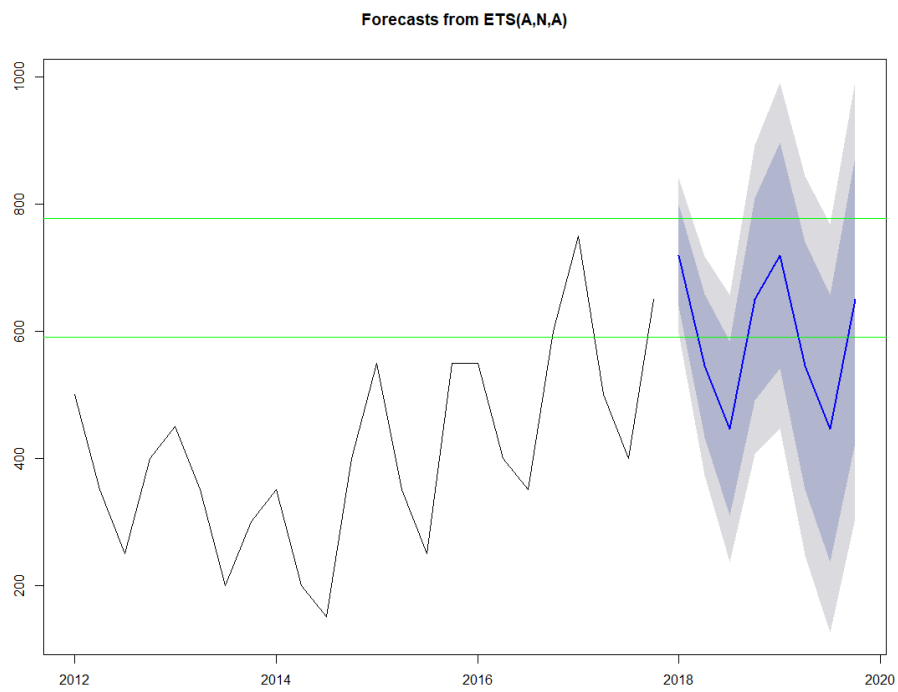


Figure 5.22. Ocean Freight forecast with the corresponding low-risk interval

## 6. Conclusions

The concept of reverse auction or tender is gaining importance in the logistics and transportation industry, as explained during the [introduction](#). The number of services contracted through this process is increasing, and many provider companies need to provide competitive prices to gain them. After a study of [existing literature](#), it is interesting to mention that almost all projects found in the subject are based in the field of construction. This supports the statement that tenders in the logistic field is a practice adapted from other industries and is being increasingly used. The [formal description of the problem](#) further develops the introduction on the logistic industry as well as the differences and structure of 3PL, 4PL and freight brokers. The study goes deeper into the subject by moving to the differences between spot quotation and reverse auctions and finally to forecasting ocean freight rates.

Tender Managers deal with challenges divided inside [our approach](#) in this project as organizational and analytical. The first group, organizational challenges and solutions include first steps, creation of a Standard Operating Procedure, distribution process and process flow, rating automation, reporting and analysis with Sugar and QlikSense, a scoring system to determine how interesting is a tender according to a logistic company network and capabilities, and the creation of a RFI database.

The second group, analytical challenges, is addressed in *Smart Tenders Management with Artificial Intelligence* project ([Figure 4.8, page 42](#)). Concretely, the main focus has been put in the second stage of the STM-AI project. That is, the development of a methodology with few parameters, which are automatically adjusted, based on time series and more concretely Winters' Method. It allows to calculate low-risk forecasted intervals, giving the decision-maker information regarding which is the most adequate long-term tariff to offer in the tender. Initial results have also been obtained for the third stage of modelling success probabilities in tenders.

The [computational experiments and analysis of the results](#) focus on ocean freight rates forecasting using Winters' Method. The inputs to the model are adapted from Hanke et al. (2001) [16] and are used to represent the costs from a logistic provider during 7 years divided by trimesters ([Figure 5.3, page 59](#)). After reading, plotting, analysing the components and fitting the model ([Figure 5.18, page 81](#)), the next year rates are forecasted. Beyond the forecasted values for the following periods, 80% and 95% confidence intervals are given by the model.

As expected the 95% confidence interval assumes less risk. During the experimentation, only one of the forecasted values lied out of the interval, and it is much closer to the 95% interval upper bound than to the 80%. For sure, it can also be concluded that the ocean freight costs evolution is highly unpredictable and probably some months or quarters where the offer will be lower to the cost will be compensated with the majority of cases where

there will be a benefit margin. The parameters alpha, beta and gamma that best fit the model (Table 5.4, page 78) are obtained thanks to RStudio optimization by minimizing the squared error.

Considering that the higher the offered rate is, the less risk it is assumed but also the less success probability, a value, or interval in this case, has to be agreed to be input in the offer. To do it, the upper bound of the final interval is obtained by calculating the average of the upper bounds along the forecasted periods. However, to keep the risk low, the lower bound is obtained by calculating the average of the forecasted values along the forecasted period; instead of using the lower bound values. Therefore, the *low-risk* confidence interval for 1-year validity is obtained (Figure 5.22, page 86) for a concrete origin and destination ocean freight.

## 7. Self-assessment

After going through the lessons learned following the different sections of the Final Master Project, it is interesting to make a self-assessment or critical reflection on the attainment of the aims initially posed. After the validation taking into consideration past data and also after a final alignment of Tender Management and Product FCL teams, the [global objective](#) to generate 'risk-aware' fare tariffs that can be valid for one year or more, and thus can be employed in tender calls is achieved.

The [partial objectives](#) are listed as the study of the current bibliography related to the theme, meeting professionals in the logistic industry and align with researchers, strict definition of the inputs involved in the project, the development of models based on historical data that can forecast the time evolution of transport costs associated with legs, the usage of the previous forecasts to develop models that can provide 'risk-aware' intervals for the values of the long-term fare prices to be used in calls for tenders, and finally the design and execution of a series of experiments. In general terms, it can be concluded that all of them are achieved.

Following the same self-assessment line, the planning and methodology are critically analysed. When it comes to [methodology](#), the forecasting part is listed as making sure enough historical data is available, developing Python, R, or Java applications that allow to automate the use of time series analysis, developing of a computation method based on the forecasted cost values that can generate the low-risk interval and validating the pilot project results (forecasted values). As will be mentioned in the following section, the rest of the methodology points could be developed, by implementing the method inside the logistic company after a final validation of the pilot project and preparation of technical and easy-to-understand documentation

Finally, the [planning](#) of the work had to be rescheduled moving the final delivery from June to September due to the workload during Tender Management high season. Additionally, there have been different meetings and discussions to obtain real ocean freight data which can be used for the forecast, so that the development of the project has been pending from different departments inside the logistic company. However, the follow-up process during the development of the project was maintained, with a set of deliveries and personal or online meetings every 15 days with the director of the project Ángel A. Juan and the academical director of the project Alberto García Villoria.

The objection in terms of objectives and methodology, which will for sure be mentioned in the *future work*, is that the real ocean freight data provided by the logistic company ([Figure 5.2, page 58](#)) from Rotterdam to Shanghai port only includes years 2015

to 2019. Moreover, there is a peak in 2017 rates that wouldn't allow the time series method to provide accurate forecasts. The increase of the rates in 2017 is explained due to an increase in eastbound demand, driven by several factors, according Maersk carrier, including domestic policies in China stimulating demand and hence requiring raw materials.

In adapted ocean freight rates plotted in [Figure 5.3 \(page 59\)](#) can be seen how data has a growing tendency (values keep increasing along time), which is confirmed with the trend-line option of Excel. A seasonal pattern can also be appreciated since the first and fourth quarters have relatively high values (repeated annually). Given the nature of the used data, it can be confirmed that Winters' technique is the most appropriate, since it will take seasonality into account.

To sum up the self-assessment critical analysis, it can be stated that the objectives, methodology and planning had to be adapted to guarantee the success of the work due to the dependence to receive real ocean freight data from inside different departments from the logistic company.

## 8. Future work

This section presents the lines of future work that have not been able to explore in this work and have remained slopes.

On the one hand, the main focus of the project has been put in the second stage of *STeM-AI project* (Figure 4.8, page 42), which is the forecast of ocean freight rates. Following this line, other forecasting methodologies, such as ARIMA models, could be used and compared to Winters' method. Furthermore, the contribution of this Final Master Project will then be implemented as a Python / Java application, so it can be easily used by tender managers at the logistic company. Final validation of the pilot project from the Tender Manager position and other stakeholders who are involved in tender the process inside the logistic Company would be required. Once the validation of the final application is done, technical and easy-to-understand documentation would be provided so that company colleagues can use the developed modules.

It is important to highlight the contribution of our approach as a new methodology with few parameters, and adjusted automatically, based on Winters' method, that allows to calculate *low-risk* forecasted intervals for long-term ocean freight rates to be offered in a tender.

On the other hand, other stages of the *STeM project* could be developed. In particular, stage 3 modelling probability of success in tenders is introduced in *our approach* section, concluding that the margin is the main variable while succeeding in a tender. However, further experimentation could be done to develop a model able to predict the probability of success in a tender call based on historical data and multiple variables. When it comes to stage 4 of the methodology proposes the use of heuristics to optimize the selection of tenders and prices in a similar way as it would be done in a rich portfolio optimization problem for assets and investment levels.



## 9. Glossary

- **APAC:** Asia Pacific
- **ARCH:** Autoregressive Conditional Heteroskedasticity
- **ARIMA:** Autoregressive Integrated Moving Average
- **CRM:** Customer Relationship Management
- **BI&A:** Business Intelligence & Analytics
- **EU:** Europe
- **EXW:** "Ex works" incoterm means, in transportation, that the seller fulfils his obligation to deliver when he has made the goods available at his premises (i.e. works, factory, warehouse, etc) to the buyer.
- **FCL:** Full Container Load
- **IOT:** Internet Of Things
- **LCL:** Less than Container Load
- **Outliers:** extreme values
- **RFI:** Request For Information
- **RFP:** Request For Proposal
- **RFQ:** Request For Quotation
- **TM:** Tender Management

# 10. Bibliography

[1] Albano, Gian Luigi y col. (2008). «THE DETERMINANTS OF SUPPLIERS' PERFORMANCE IN E-PROCUREMENT: EVIDENCE FROM THE ELECTRONIC PUBLIC ADMINISTRATION'S MARKETPLACE (MEPA)».

[2] Angel A. Juan & Carles Serrat. PART II. Time Series Analysis: Introduction to TSA and Simple Time Series Methods & Moving Averages

[3] Bass, Brian. "The Definition of Logistical Organization." *Small Business - Chron.com*. [consulted: 28<sup>th</sup> of April, 2019] Available on the Internet:

<http://smallbusiness.chron.com/definition-logistical-organization-23276.html>

[4] Bowen and Edwards (1985)

[5] Casassas et al. (2018). «Rhenus – Tenders Optimization. »

[6] Coghlan, Avril (2018). «A Little Book of R For Time Series»

[7] Da Veiga, CLAUDIMAR PEREIRA y col. (2014). «Demand forecasting in food retail: A comparison between the Holt-Winters and ARIMA models». In: WSEAS transactions on business and economics 11.1, pages. 608-614.

[8] DeAngelis, SF (2015). Predictive analytics becoming a mainstream business tool.

[9] Econometrics with R. 11.2 Probit and logit regression. [consulted: 10<sup>th</sup> of May, 2019]. Available on the internet: <https://www.econometrics-with-r.org/11-2-palr.html>

[10] Fildes, Robert et al. (1998). «Generalising about univariate forecasting methods: further empirical evidence». In: International Journal of Forecasting 14.3, pages 339-358.

[11] Flanagan, D. J., D'Mello, J. P., and O'Shaughnessy, K. C. (1998). «Completing the deal: Determinants of successful tender offers.» *Journal of Applied Business Research*, 14, 21-32.

[12] FULLTILT Logistics. What is a Logistics Company? What Does Logistics Company Do?. [consulted: 28<sup>th</sup> of April, 2019]. Available on the Internet:

<https://fulltiltlogistics.com/what-is-a-logistics-company/>

[13] Geomelos et al. (2014). «Forecasting spot prices in bulk shipping using multivariate and univariate models».

[14] Gog (1998). «Evaluating forecast accuracy using error measures».

[15] Hafizoglu, A Baykal, Esmâ S Gel y Pinar Keskinocak (2016). Price and lead time quotation for contract and spot customers. In: *Operations Research* 64.2, pages. 406-415.

[16] Hanke, John E, Arthur G Reitsch and Dean W. Wichern (2001). Business forecasting. Vol. 9. Prentice Hall Upper Saddle River, NJ.

[17] Hannon, David (2006). e-Sourcing finds its place». In: *Purchasing Magazine*, November 2.

[18] HierarchyStructure. Logistic Company Hierarchy. [consulted: 29<sup>th</sup> of April, 2019] Available on the Internet:

<https://www.hierarchystructure.com/logistic-company-hierarchy/>

[19] Hoffmeister, J. R., and Dyl, E. A. (1981). Predicting outcomes of cash tender offers. *Financial Management*, 50-58.

[20] Kizys et al. (2015). A Biased-Randomized Iterated Local Search Algorithm for Rich Portfolio Optimization.

[21] Liang et al. (2016). Multi-objective optimal strategy for generating and bidding in the power market. In: *Energy Conversion and Management* 57:13–22 .

[22] Lloyd's Loading List. Ocean freight rates soar from North Europe to Asia. [consulted: 6<sup>th</sup> of July, 2019]. Available on the Internet:

<https://www.lloydsloadinglist.com/freight-directory/news/Ocean-freight-rates-soar-from-North-Europe-to-Asia/68798.htm#.XWOKqFwzZhE>

[23] Makridakis, Spyros and Michele Hibon (1997). «ARMA models and the Box–Jenkins methodology». In: *Journal of Forecasting* 16.3, pages 147-163.

[24] Makridakis, Spyros et al. (1993). The M2-competition: A real-time judgmentally based forecasting study. In: *International Journal of Forecasting* 9.1, pages 5-22.

[25] Manzanero A. L. D. (2009). *Forecasting Container Freight Rates – An Econometric Analysis*. [consulted: 25<sup>th</sup> of April, 2019]. Available on the Internet:

[https://studenttheses.cbs.dk/xmlui/bitstream/handle/10417/803/leopoldo\\_david\\_alonso\\_man\\_zanero\\_og\\_kristof\\_johannes\\_krupp.pdf?sequence=1](https://studenttheses.cbs.dk/xmlui/bitstream/handle/10417/803/leopoldo_david_alonso_man_zanero_og_kristof_johannes_krupp.pdf?sequence=1)

[26] McGuire et al. (2015). Marketing & Sales Big Data, Analytics, and the Future of Marketing & Sales.

[27] Mentzer, John T and Jon Schroeter (1994). Integrating logistics forecasting techniques, systems, and administration: the multiple forecasting system. In: Journal of Business Logistics 15.2, page. 205.

[28] Montgomery, Douglas C. et al (2015). Introduction to Time Series Analysis and Forecasting 2015.

[29] Munim, Ziaul Haque y Hans-Joachim Schramm (2017). Forecasting container shipping freight rates for the Far East–northern Europe trade lane. In: Maritime Economics & Logistics 19.1, pages 106-125.

[30] Panadero et al. (2018). A variable neighborhood search simheuristic for project portfolio selection under uncertainty. In: *Journal of Heuristics*.

[31] Veenstra, Albert Willem y Philip Hans Franses (1997). A co-integration approach to forecasting freight rates in the dry bulk shipping sector. In: *Transportation Research Part, policy and practice* 31.6, pages 447-458.

[32] Walkling, R. A. (1985). Predicting tender offer success: A logistic analysis. *Journal of financial and Quantitative Analysis*, 20(4), 461-478.

[33] Wan, Z., and Beil, D. R. (2009). RFQ auctions with supplier qualification screening. *Operations Research*, 57(4), 934-949.

[34] Wanous et al. (1999). PRACTICAL DECISION MAKING TOOL: BID NO BID CASE STUDY.

[35] Watt, DJ, B Kayis y K Willey (2010). The relative importance of tender evaluation and contractor selection criteria. In: *International Journal of Project Management* 28.1, pages 51-60.

# 11. Annexes

## ## 1. Reading Time Series Data

### ## 1.1 Real Data

```
myData <- read.csv(file="D:/Configuraciones/cristina.alvarez/Mis documentos/R/FCL
forecast/DATA/SeafreightData.csv", header=TRUE, sep=";")
```

```
head(myData)
```

```
myData20 <- myData$Total.20
```

```
myData40 <- myData$Total.40
```

```
myData20timeseries <- ts(myData20, start = c(2015,1), frequency = 24)
```

```
myData40timeseries <- ts(myData40, frequency = 24)
```

### ## 1.2 Adapted Data

```
myData_V2 <- read.csv(file="D:/Configuraciones/cristina.alvarez/Mis documentos/R/FCL
forecast/DATA/SeafreightData_V2.csv", header=TRUE, sep=";")
```

```
myDataV2 <- myData_V2$Cost
```

```
myDataV2timeseries <- ts(myDataV2, start = c(2012,1), frequency = 4)
```

## ## 2. Plotting Time Series

### ## 2.1 Real Data

```
par(mar = c(5, 5, 5, 5))
```

```
plot.ts(ylab = "myData20timeseries (USD)", xlab = "Time (years)", col= "red", myData20timeseries)
```

```
par(new=TRUE)
```

```
plot.ts(myData40timeseries, xaxt="n", yaxt="n", ylab = "", xlab = "", col = "blue", mar = c(5, 5, 5, 5))
```

```
mtext("myData40timeseries (USD)", side=4, line=3)
```

```
axis(side = 4)
```

```
logmyData20timeseries <- log(myData20timeseries)
```

```
logmyData40timeseries <- log(myData40timeseries)
```

```
plot.ts(logmyData20timeseries)
```

```
plot.ts(logmyData40timeseries)
```

## **## 2.2 Adapted Data**

```
par(mar = c(5, 5, 5, 5))
```

```
plot.ts(ylab = "myDataV2timeseries (USD)", xlab = "Time (years)", col= "blue", myDataV2timeseries)
```

## **## 3. Decomposing Time Series**

### **## 3.1 Decomposing Non-Seasonal Data**

```
library("TTR")
```

```
myDataV2timeseriesSMA3 <- SMA(myDataV2timeseries,n=3)
```

```
plot.ts(myDataV2timeseriesSMA3)
```

```
myDataV2timeseriesSMA8 <- SMA(myDataV2timeseries,n=8)
```

```
plot.ts(myDataV2timeseriesSMA8)
```

### **## 3.2 Decomposing Seasonal Data**

```
myDataV2timeseriescomponents <- decompose(myDataV2timeseries, type = "multiplicative")  
  
plot(myDataV2timeseriescomponents)
```

### **## 3.3 Seasonally Adjusting**

```
myDataV2timeseriesseasonallyadjusted <- myDataV2timeseries /  
myDataV2timeseriescomponents$seasonal  
  
plot(myDataV2timeseriesseasonallyadjusted)
```

## **## 4. Forecasts using Holt-Winters Exponential Smoothing**

### **## 4.1 Verification of the application of the Winters method**

```
myData_V3 <- read.csv(file="D:/Configuraciones/cristina.alvarez/Mis documentos/R/FCL  
forecast/DATA/SeafreightData_V3.csv", header=TRUE, sep=";")  
  
myDataV3 <- myData_V3$Cost  
  
myDataV3timeseries <- ts(myDataV3, start = c(2012,1), frequency = 4)
```

```
HoltWinters(myDataV3timeseries, seasonal = "multiplicative")  
  
myDataV3timeseriesforecasts <- HoltWinters(myDataV3timeseries)  
  
plot(myDataV3timeseriesforecasts)
```

```
library("forecast")  
  
myDataV3timeseriesforecasts2 <- forecast(myDataV3timeseries, h=8)  
  
plot(myDataV3timeseriesforecasts2)
```

### **## 4.2 Application of the Winters method**

```
HoltWinters(myDataV2timeseries, seasonal = "multiplicative")

myDataV2timeseriesforecasts <- HoltWinters(myDataV2timeseries)

myDataV2timeseriesforecasts

plot(myDataV2timeseriesforecasts)

myDataV2timeseriesforecasts2 <- forecast(myDataV2timeseries, h=8)

plot(myDataV2timeseriesforecasts2)

accuracy(myDataV2timeseriesforecasts2)
```

### **## 4.3 Confidence intervals**

```
myData_V4 <- read.csv(file="D:/Configuraciones/cristina.alvarez/Mis documentos/R/FCL
forecast/DATA/SeafreightData_V4.csv", header=TRUE, sep=";")

myDataV4 <- myData_V4$Cost

myDataV4timeseries <- ts(myDataV4, start = c(2017,2), frequency = 4)

HoltWinters(myDataV4timeseries, seasonal = "multiplicative")

myDataV4timeseriesforecasts <- HoltWinters(myDataV4timeseries)

myDataV4timeseriesforecasts2 <- forecast(myDataV4timeseries, h=8)

plot(myDataV4timeseriesforecasts2)
```



```
myData_V5 <- read.csv(file="D:/Configuraciones/cristina.alvarez/Mis documentos/R/FCL
forecast/DATA/SeafreightData_V5.csv", header=TRUE, sep=";")
```

```
myDataV5 <- myData_V5$Cost
```

```
myDataV5timeseries <- ts(myDataV5, start = c(2015,3), frequency = 4)
```

```
HoltWinters(myDataV5timeseries, seasonal = "multiplicative")
```

```
myDataV5timeseriesforecasts <- HoltWinters(myDataV5timeseries)
```

```
myDataV5timeseriesforecasts2 <- forecast(myDataV5timeseries, h=8)
```

```
plot(myDataV5timeseriesforecasts2)
```

```
myData_V6 <- read.csv(file="D:/Configuraciones/cristina.alvarez/Mis documentos/R/FCL
forecast/DATA/SeafreightData_V6.csv", header=TRUE, sep=";")
```

```
myDataV6 <- myData_V6$Cost
```

```
myDataV6timeseries <- ts(myDataV6, start = c(2013,4), frequency = 4)
```

```
HoltWinters(myDataV6timeseries, seasonal = "multiplicative")
```

```
myDataV6timeseriesforecasts <- HoltWinters(myDataV6timeseries)
```

```
myDataV6timeseriesforecasts2 <- forecast(myDataV6timeseries, h=8)
```

```
plot(myDataV6timeseriesforecasts2)
```

```
myData_V7 <- read.csv(file="D:/Configuraciones/cristina.alvarez/Mis documentos/R/FCL
forecast/DATA/SeafreightData_V7.csv", header=TRUE, sep=";")
```

```
myDataV7 <- myData_V7$Cost
```

```
myDataV7timeseries <- ts(myDataV7, start = c(2012,1), frequency = 4)
```

```
HoltWinters(myDataV7timeseries, seasonal = "multiplicative")
```

```
myDataV7timeseriesforecasts <- HoltWinters(myDataV7timeseries)
```

```
myDataV7timeseriesforecasts2 <- forecast(myDataV7timeseries, h=8)
```

```
plot(myDataV7timeseriesforecasts2)
```

#### **## 4.4 Evaluation of Winters method**

```
acf(myDataV2timeseriesforecasts2$residuals, lag.max=20)
```

```
Box.test(myDataV2timeseriesforecasts2$residuals, lag=20, type="Ljung-Box")
```

```
plot.ts(myDataV2timeseriesforecasts2$residuals) # make a time plot
```

```
plotForecastErrors(myDataV2timeseriesforecasts2$residuals) # make a histogram
```

#### **## 4.5 Finding the best fit of the model**

```
myDataV8timeseriesforecasts <- HoltWinters(myDataV2timeseries, seasonal = "multiplicative", alpha =  
0.4, beta = 0.1, gamma = 0.3)
```

```
myDataV9timeseriesforecasts <- HoltWinters(myDataV2timeseries, seasonal = "multiplicative", alpha =  
0.9, beta = 0.1, gamma = 0.3)
```

```
myDataV10timeseriesforecasts <- HoltWinters(myDataV2timeseries, seasonal = "multiplicative", alpha =  
0.4, beta = 0.9, gamma = 0.3)
```

```
myDataV11timeseriesforecasts <- HoltWinters(myDataV2timeseries, seasonal = "multiplicative", alpha =  
0.4, beta = 0.1, gamma = 0.9)
```

```
myDataV12timeseriesforecasts <- HoltWinters(myDataV2timeseries, seasonal = "multiplicative", alpha =  
0.9, beta = 0.1, gamma = 0.9)
```

```
myDataV13timeseriesforecasts <- HoltWinters(myDataV2timeseries, seasonal = "multiplicative")
```

```
plot(myDataV8timeseriesforecasts)
```

```
plot(myDataV9timeseriesforecasts)
```

```
plot(myDataV10timeseriesforecasts)
```

```
plot(myDataV11timeseriesforecasts)
```

```
plot(myDataV12timeseriesforecasts)
```

```
plot(myDataV13timeseriesforecasts)
```

```
myDataV8_SSE <- myDataV8timeseriesforecasts$$SSE
```

```
myDataV9_SSE <- myDataV9timeseriesforecasts$$SSE
```

```
myDataV10_SSE <- myDataV10timeseriesforecasts$$SSE
```

```
myDataV11_SSE <- myDataV11timeseriesforecasts$$SSE
```

```
myDataV12_SSE <- myDataV12timeseriesforecasts$$SSE
```

```
myDataV13_SSE <- myDataV13timeseriesforecasts$$SSE
```

## **## 5. Establishing low-risk intervals for long-term tariffs for tenders**

```
minvalue <- mean(myDataV3timeseriesforecasts2$mean[1:4])
```

```
maxvalue <- mean(myDataV3timeseriesforecasts2$upper[9:12])
```

```
plot(myDataV3timeseriesforecasts2)
```

```
abline(h=minvalue, col="green")
```

```
abline(h=maxvalue, col="green")
```

