Modeling of container throughput in the European port system-analysis of container liner shipping networks

Georgios Stamatopoulos

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Department of Civil and Environmental Engineering

PhD. Thesis

Modeling of Container Throughput in the European Port System- Analysis of Container Liner shipping Networks

Georgios Stamatopoulos

Phd Director:

Dr. Sergi Saurí Marchán

PhD Program in Transport Engineering and Infrastructure

Department of Civil and Environmental Engineering (DECA)

Universitat Politècnica de Catalunya (UPC)

Barcelona, March 2023

Modeling of Container Throughput in the European port system- Analysis of Container Liner shipping Network

Autor:

Georgios Stamatopoulos

Director de tesis

Dr. Sergi Saurí Marchán

Memòria presentada per optar al títolo de

Doctor

en el programa de Ingeniería e Infraestructuras del Transporte

Departmento de Ingeniería Civil y Ambiental (DECA)

Universitat Politècnica de Catalunya (UPC)

Barcelona, marzo de 2023

Abstract

This thesis records the evolution of the container throughput dynamics of the European port system from 2001-2020, analyses the dynamics of the European container port hierarchy with methodologies that have not been used in that context, applies network analysis on the liner shipping networks of individual shipping companies and proposes a model for short-term forecasts of container throughput and their accompanying prediction intervals.

Chapter 2 explores the dynamics of container throughput in the European port system from 2001-2020 at the level both of groups of ports and individual ports. The groups of ports include the maritime ranges that shape the total European port system and the multi-port gateway regions introduced by Notteboom (2010). The evolution of the container throughput is examined by the application of concentration indicators. The market share of port groups or ports in the total throughput of the system, the normalized Herfindahl–Hirschman Index and a customized form of shift-share analysis are used for this purpose. The influence of the two crises that occurred during the examined period, the global financial crisis of 2009 and the health crisis of 2020, on the container throughput dynamics is explored.

The dynamics of the European container port hierarchy are analyzed in chapter 3 by the application of 3 methodologies. First, two types of rank-size models are used in order to investigate the formation of the hierarchies and the dynamics of container throughput distribution. Second, a methodology based on the principals of Analytic Hierarchy Process (AHP) combined with an allometric growth model is introduced in order to explore the relative growth of large ports as a group upon groups of medium-sized and small ports. Third, a Markov chain modeling is applied in order to investigate the mobility of ports within the hierarchy of the European container port system. The dataset that is used in all the methodologies consists of the annual container port throughputs in TEUs of the top-100 European container ports for every year between 2001 and 2020.

Chapter 4 analyzes the container liner shipping networks of the liner operators of Maersk and COSCO Shipping in 2001, 2007 and 2010. The selected time period allows the observation of the evolution of the companies' networks prior and after the global financial crisis of 2009. First, the whole network of the carriers is analyzed through global network indicators and visualization of the networks. Second, the shipping networks are investigated for the presence of complex network properties by fitting a power law distribution to the degree distribution of the networks. Third, the positions of the ports in the companies' networks are explored by calculating their degree and betweenness centrality. Fourth, the weighted degree of the ports

and the strongest inter-port links are presented using the frequency (weekly calls) as weight of the links. Finally, the chapter focuses on the regional networks of the companies and their evolution.

A combined SARIMA-(G) ARCH model is proposed in chapter 5 in order to produce short-term forecasts of container throughput and prediction intervals of these forecasts. The uncertainty of the forecasts is expressed by the prediction intervals associated with the point forecasts. The point forecasts are generated form the SARIMA part of the model and the prediction intervals from the ARCH or GARCH part. ARCH/GARCH models are used in order to model and forecast the volatility of container throughput time series. The combined SARIMA-(G)ARCH model is applied to monthly container throughput data of the port of Barcelona in order to produce short-term point forecasts and its accompanying prediction intervals.

Keywords: container ports; port hierarchy; liner shipping networks; container throughput dynamics; point forecasts; prediction intervals.

Table of Contents

List of Figures

List of Tables

Abbreviations

NB: This section is designed to clarify abbreviations and acronyms used in the PhD thesis. Not all abbreviations in the text might be included below.

ADF : Augmented Dickey-Fuller
AHP: Analytic Hierarchy Process
AIC: Akaike information criterion
AIS: Automatic Identification System
AL: Alphaliner
APARCH: Assymetrical Power Generalized Autoregressive Conditional Heteroskedasticity
AR: Autoregressive
ARCH: Autoregressive Conditional Heteroskedasticity
ARIMA: Autoregressive Integrated Moving Average
ARMA: Autoregressive Moving Average
AX: Axsmarine
BIC: Bayesian Information Criterion
BP: Back Propagation
CIO: Containerization International Online
CIY: Containerization International Yearbook
CR: Containing Ratio
EGARCH: Exponential Generalized Autoregressive Conditional Heteroskedasticity
FANN: Fuzzy Artificial Neural Networks
GARCH: Generalized Autoregressive Conditional Heteroskedasticity
GM: Grey Model
HHI: Herfindahl-Hirschman Index
KPSS : Kwiatkowski–Phillips–Schmidt–Shin
LS-SVR: Least Squares Support Vector Regression
LMIU: Lloyds Maritime Intelligence Unit
MA: Moving Average
MAPE: Mean Absolute Percentage Error
MAIW: Mean Absolute Interval Width
MRIW: Mean Relative Interval Width
SARIMA: Seasonal Autoregressive Integrated Moving Average
SR: Searates
SVR: Support Vector Regression
TCT: Total Container Throughput
TEU: Twenty-foot Equipment Unit
TGARCH: Threshold Generalized Autoregressive Conditional Heteroskedasticity
UNCTAD: United Nations Conference on Trade and Development
YRD: Yangtze River Delta

Chapter 1

Introduction

1.1 Background and objectives

Maritime transport is responsible for over 80% of the volume of international trade in goods (UNCTAD, 2022). Containerized trade since its begging in 1956 and its first Transatlantic service in 1968, has played a very influential role in the growth of global trade and the adoption of new practices in global supply chains. The standardization of container sizes reduced the transport costs, made the transport of goods safer and enabled the multimodal transport. The value of containerized seaborn trade in 2020 was estimated at 66% of the total value of maritime trade (Rodrigue and Notteboom, 2022).

The shipping companies organize liner shipping services in order to carry the containers from the origin to the destination market. The difference of liner shipping compared to other types of maritime shipping is that it operates under a published fixed schedule. A liner shipping service is a maritime route consisted of consecutive calls at selected ports by the carrier. In order to design a service, the carriers must decide the ports of call, the frequency of the service, the speed, the number and the size of the vessels operating at the service. The combination of several services creates the maritime network of the shipping line. The decisions of individual shipping lines concerning the design of their network determine port hierarchy.

The vessels of ocean carriers sail between continents and connect markets all over the world. Europe is one of the most important markets. In 2021, the containerized trade between Asia and Europe was 26.5 million TEUs and between North America and Europe 8.2 million TEUs (UNCTAD, 2022). Consequently, the European container port system is one of the busiest in the world. The ports handle the containers arriving at their container terminals by the vessels of liner shipping companies. The number of loaded and unloaded containers determines the container throughput of ports. The container throughput determines the size of ports. The European port system is a mix of ports of different sizes (Notteboom, 2010). Based on the container throughput a port hierarchy is shaped, consisted of large, medium and small-sized ports. The changes in the hierarchy depend on the growth patterns of ports of different size.

The traffic flows arriving at ports are characterized by a high level of volatility caused by factors as seasonal fluctuations, choices of important market players and all types of crises

(Notteboom, 2022b). As a result, the distributional patterns of container flows and the emerging concentration trends present a dynamic nature and change over time. Short-term forecasts of container throughput volumes are valuable for container terminal operators and port authorities in order to make operational decisions about the allocation of port's resources.

In this context the objectives of this thesis are the following:

- Update the container throughput dynamics of the European container port system covering the period from 2001-2020.
- Compare the relative growth of size categories of European ports one upon another and investigate the mobility of ports from one size category to another.
- Propose a model for short-term forecasting of container throughput together with an assessment of the uncertainty of these forecasts in order to facilitate operational decision making.
- Analyze the container shipping networks of individual shipping lines as port hierarchies are shaped by their decisions.

1.2 Research questions

Research question 1: Which are the developments in the container throughput of the European port system?

In the first 20 years of the 21st century the European container port system evolved under alternating conditions of prosperity and recession. This thesis investigates the evolution of the container throughput dynamics of the European port system from 2001-2020 and explores the influence of the global financial crisis of 2009 and the health crisis of 2020 on the growth of the European port groups and individual ports. This study adopts a narrowing down approach by recording the evolution of the container throughput of the total European port system and continuing to port ranges, multi-port gateway regions and finally to individual ports. The market share of port groups or ports in the total throughput of the system, the normalized Herfindahl–Hirschman Index and a customized form of shift-share analysis are used as indicators of concentration or deconcentration. This study updates previous studies of Notteboom (1997, 2010) and also functions as an introduction to the rest research questions of the thesis.

Research question 2: How was hierarchy in the European container port system evolved?

Container throughput dynamics in the European container port system have been examined at the level of Individual ports or groups of ports based on their location. These groupings of ports in broader (ranges) or narrower (multi-port gateway regions) geographical units contain ports of different sizes. This thesis analyses the dynamics of port hierarchies in the European container port system by introducing an alternative grouping of ports based on their size. The impact of size categories of ports (large, medium and small-sized port groups) one upon another has not been explored yet. A methodology based on the principals of Analytic Hierarchy Process (AHP) combined with an allometric growth model is proposed in order to compare the relative growth of all size categories upon each other. Consequently, the growth dynamics of large ports as a group upon groups of medium-sized and small ports are revealed.

Another issue that has not been explored yet is the mobility of ports within the hierarchy of European container port system. A Markov chain modeling is applied in order to explore the mobility of ports from one size category to another. The Markov chain approach permits the observation of the proportion of port movements between groups that contain small, medium or large ports. Furthermore, the speed with which these movements are made is estimated by the calculation of the mean first passage time which is the expected number of years required for a port to reach a size group for the first time from his starting group.

Finally, the presence and the evolution of medium-sized ports in the European container port hierarchy is assessed by the application of a quadratic rank-size model.

Research question 3: How are the liner shipping networks of individual shipping lines structured?

Port hierarchy is determined by the decisions of individual carriers (Notteboom, 2022a). Nevertheless, the container liner shipping networks of individual shipping lines have not been explored yet. So far, the container liner shipping networks have been constructed by aggregating vessel movements or the liner services of several shipping lines. This thesis constructs and analyzes the liner shipping networks of individual liner shipping companies.

The evolution of these networks under different circumstances, firstly in a time period of prosperity and secondly in the environment of a global financial crisis, is presented. By comparing the liner shipping networks of different individual carriers and their evolution, their strategies in organizing liner shipping networks are revealed. Port hierarchies in the shipping companies' networks are formed by the calculation of network indicators. These hierarchies of ports based on their positions in the shipping lines' networks are strongly correlated with the port's throughputs volumes (Deng et al., 2009; Ducruet et al., 2011b; Kang and woo 2017). Furthermore, the whole network of the shipping line's is analyzed by the calculation of global network indicators and by visualization. It is also explored for the presence of complex network properties.

The methodology of this thesis is different in two aspects compared to other studies providing network analysis of maritime container shipping. The first is the aforementioned disaggregated approach, where the container liner shipping network is analyzed at the individual carriers' level. As a result, this study does not aggregate the liner shipping services of several carriers but uses their liner services separately in order to form their networks. The second is that the inter-ports links are weighted by the frequency of the liner services measured in weekly calls and not by the transport capacity of the vessels in TEUs. This approach permits to understand better the time proximities between ports.

Research question 4: How to produce forecasts of container throughput and their accompanying prediction intervals?

Forecasting of container throughput plays a crucial role in port planning of container ports. Due to its importance for port planning it has been a popular subject in port studies. The common element in these studies is that they produce point forecasts of the container throughput, but they do not assess the uncertainty of these forecasts. The traffic flows arriving at ports are characterized by a high level of volatility caused by factors as seasonal fluctuations, choices of important market players and all types of crises (Notteboom, 2022b). In this highly uncertain environment, forecasted values of container throughput should be accompanied by their prediction intervals which express the uncertainty of the forecasts. The generation of the accompanying prediction intervals of the point forecasts of container throughput has not been addressed yet.

In order to produce the prediction intervals of the forecasted values of container throughput, the volatility of the container throughput should be modelled and forecasted. This thesis introduces a combined Seasonal Autoregressive Integrated Moving Average (SARIMA) - Autoregressive Conditional Heteroskedasticity (ARCH) or Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model in order to produce both point forecasts and prediction intervals of container throughput. The point forecasts are generated from the SARIMA part of the model while the prediction intervals from the (G)ARCH part, which is a model created for capturing volatility in time series data. The model is applied for short-term forecasts of the monthly container throughput of the port of Barcelona.

1.3 Outline of the thesis

The reminder part of this thesis is structured as follows: Chapter 2 provides an update of the container throughput dynamics of the European container port system over the period 2001-2020. Chapter 3 analyzes the port hierarch dynamics in the European container port system by comparing the relative growth of all size categories of ports upon each other and explores the mobility of ports within the hierarchy. Chapter 4 analyzes the liner shipping networks of individual liner shipping companies. Chapter 5 proposes a model for short-term forecasts of container throughput and their prediction intervals. Finally, chapter 6 presents the main findings of this thesis and indicates directions for future research.

Chapter 2

Container throughput dynamics of the European container port system over the period 2001-2020

2.1 Introduction

The container throughput dynamics of the European container port system over the period 2001-2020 are analyzed in this chapter. This investigation functions as an introduction to the research presented in chapter 3 on the hierarchy dynamics of the European container port system.

The dynamics of container throughput in the European port system are explored at the level both of groups of ports and individual ports. The groups of ports include the maritime ranges that shape the total European port system and the multi-port gateway regions introduced by Notteboom (2010). The evolution of the container throughput is examined by the application of concentration indicators. The market share of port groups or ports in the total throughput of the system, the normalized Herfindahl–Hirschman Index and a customized form of shift-share analysis are used for this purpose. The influence of the two crises that occurred during the examined period, the global financial crisis of 2009 and the health crisis of 2020, on the container throughput dynamics is explored.

The remainder of the chapter is structured as follows: Section 2.2 presents the European container port ranges, the multi-port gateway regions and the indices used to describe container throughput dynamics. Section 2.3 describes the evolution of the total European container throughput. Section 2.4 analyzes the throughput dynamics of the European port ranges. Section 2.5 focuses on the European multi-port gateway regions. Section 2.6 records the evolution of container throughput at individual ports. The final section 2.7 presents the conclusions of this chapter.

2.2 Delineation of port ranges, multi-port gateway regions and introduction of indices used to describe container throughput dynamics.

This chapter describes the container throughput dynamics of the European port system from 2001-2020 following the studies of Notteboom (1997, 2010). The analysis is based on the annual container throughputs of 122 European ports. It includes the evolution of the container throughput not only of individual ports but also of port groups as it is useful to analyze the container traffic dynamics of groups of ports in order to have a comprehensive understanding of the distributional patterns of containers in the European port system (Notteboom, 2010). The analysis begins from the total European port system and narrows down to more disaggregated units of port ranges, multi-port gateway regions and finally to individual ports.

Figure 2-1 presents the coastline covered by each of the 6 European port ranges and the multi-port gateway regions introduced by Notteboom (2010). The 122 ports included in this study are allocated to the 6 ranges. The Hamburg-Le Havre range consists of 14 ports, the Mediterranean range of 36, the UK range of 23, the Atlantic range of 17, the Baltic/Scandinavian range of 29 and the Black Sea range of 3 ports. The multi-port gateway regions include the following ports (Notteboom, 2010):

Extended Rhine–Scheldt Delta: Rotterdam, Antwerp, Zeebrugge, Amsterdam, Ghent, Zeeland Seaports, Ostend, Dunkirk.

Helgoland Bay: Hamburg, Bremen/Bremerhaven, Cuxhaven, Wilhelmshaven.

UK South East Coast: Felixstowe, Southampton, Thamesport, Tilbury, Hull.

Spanish Med: Barcelona, Valencia, Tarragona.

Ligurian range: Genoa, Savona, Leghorn, La Spezia.

Seine Estuary: Le Havre, Rouen.

Black Sea West: Constanza, Burgas, Varna.

South Finland: Helsinki, Rauma, Hamina-Kotka, Turku.

Portuguese range: Lisbon, Leixoes, Sines.

North Adriatic: Venice, Trieste, Ravenna, Koper, Rijeka.

Gdansk Bay: Gdynia, Gdansk.

Kattegat/The Sound: Goteborg, Malmo/Copenhagen, Helsingborg, Aarhus.

The ports that were not included in any of the defined multi-port gateway regions are characterized as stand-alone gateway ports.

The description of the container throughput dynamics of the port groups or ports is based on the market share, the normalized Herfindahl–Hirschman Index (HHI^*) and the net shifts of the port groups or ports. The market share of a port group or port is defined as:

$$S_i = \frac{T_i}{\sum_{i=1}^{n} T_i} \tag{2.1}$$

where S_i is the share of port group or port i in the port system, T_i the container throughput of port group or port i and n the number of the port groups or ports in the port system. The normalized Herfindahl–Hirschman Index is given by:

$$HHI^* = \frac{\sum_{i=1}^{n} S_i^2 - \frac{1}{n}}{1 - \frac{1}{n}} \tag{2.2}$$

where HHI^* is the normalized Herfindahl–Hirschman Index $0 \leq HHI^* \leq 1$, S_i the market share of port group or port i and n the number of the port groups or ports in the port system. Higher HHI^* indicates inequality in distributions and thus a more concentrated system while lower HHI^* distributional equality and a more deconcentrated system. The net shifts are a customized form of shift-share analysis introduced by Notteboom (1997). They are calculated by the following equations:

$$\mathrm{ABSGR_i} = \mathrm{T_{it1}} - \mathrm{T_{it0}} = \mathrm{SHARE_i} + \mathrm{SHIFT_i} \tag{2.3}$$

$$\mathrm{SHARE_i} = \left(\frac{\sum_{i=1}^{n} \mathrm{T_{it1}}}{\sum_{i=1}^{n} \mathrm{T_{it0}}} - 1 \right) \times \mathrm{T_{it0}} \tag{2.4}$$

$$\mathrm{SHIFT_i} = \mathrm{T_{it1}} - \frac{\sum_{i=1}^{n} \mathrm{T_{it1}}}{\sum_{i=1}^{n} \mathrm{T_{it0}}} \times \mathrm{T_{it0}} \tag{2.5}$$

where $\mathrm{ABSGR_i}$ is the absolute growth of port group or port i in time period t0-t1, $\mathrm{SHARE_i}$ is the share effect expressing the expected change in the port group or port as if it would maintain the growth rate of the total port system and $\mathrm{SHIFT_i}$ is the shift effect expressing the difference between the expected change in the port group or port as if it would maintain the growth rate of the total port system and the real change.

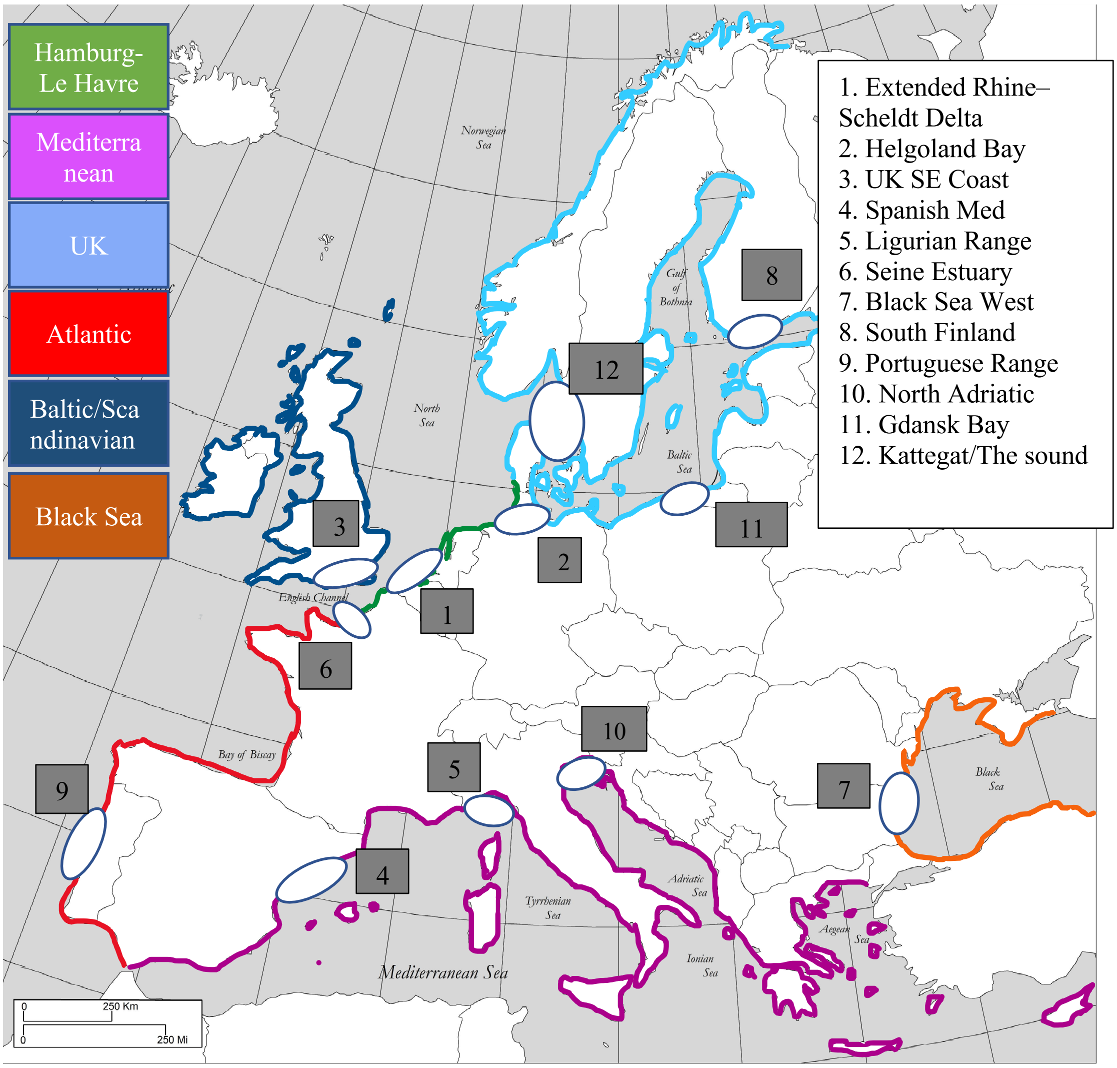

Figure 2-1: Port ranges and multi-port gateway regions of the European container port system. Source: Author's elaboration based on Notteboom (2010).

2.3 Total European container throughput

Worldwide container throughput increased from approximately 235 million TEUs in 2001 to 792 million TEUs in 2020. During the same time period, the share of the European container throughput in the worldwide throughput decreased from 21.7% to 13.57%. This decrease in the share of the European container throughput Is mainly attributed to the relocation of production centers from Europe and United States of America to China which led to the rise of Chinese container ports.

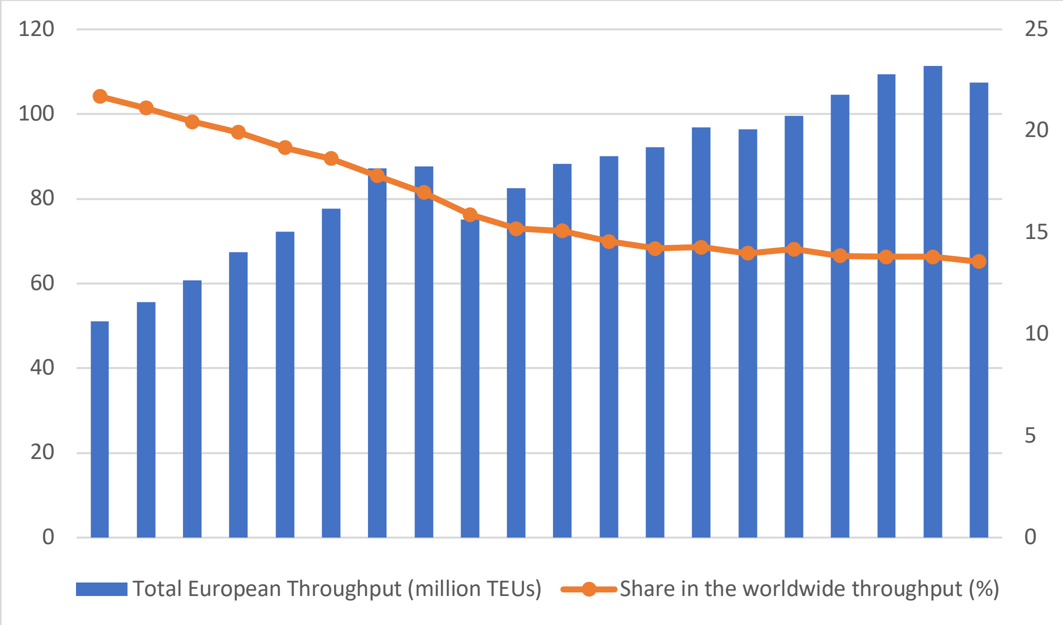

Figure 2-2: Total throughput of European port system.
Source: Author’s elaboration based on various sources.

In the first decade of the 21st century, the European container port throughput and the worldwide throughput followed parallel growth trends with the growth rate of the worldwide throughput being always higher (figure 2-2) which was also reflected in the decrease of the European share in the worldwide throughput from 21.7% in 2001 to 15.19% in 2010 (figure 2-1). This pattern didn't continue in the second decade where the growth rate of the European container throughput was at some time periods the same or even larger than the growth rate of worldwide throughput which was also depicted by a relatively stable share of the European container throughput in the worldwide throughput (figure 2-1).

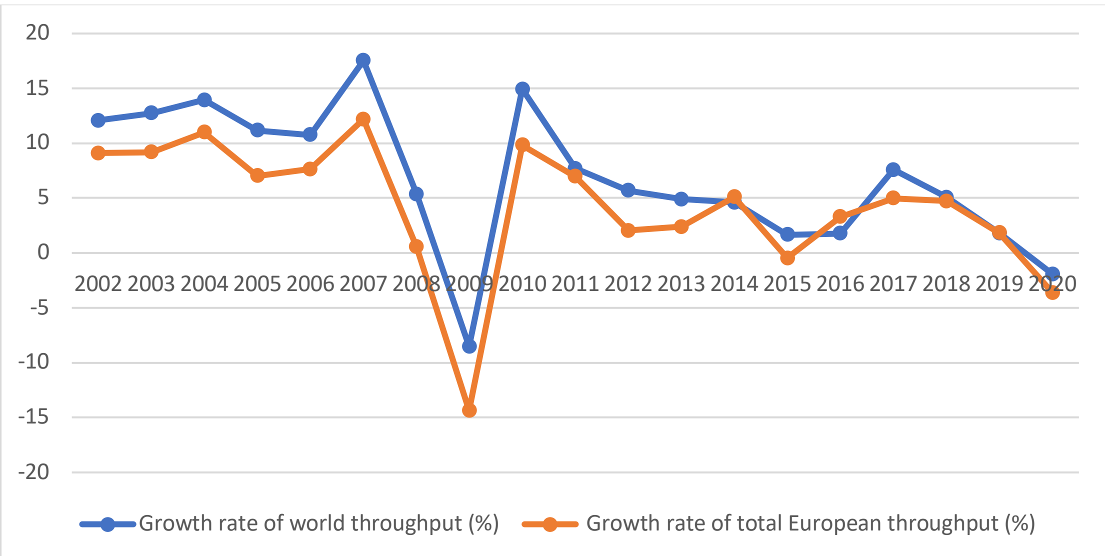

Figure 2-3: Annual growth rates of European and worldwide container throughput.
Source: Author’s elaboration based on various sources.

2.4 Analysis of European container port ranges

Hamburg-Le Havre and Mediterranean were the two ranges that concentrated the large majority of container flows, as they presented an average combined share of approximately 75% in total European throughput from 2001-2020. The market share dynamics between the two ranges were changing during the examined period, revealing changing distributional patterns of container flows between North and South European ports. From 2003 the gap between the two ranges began to widen and reached its peak in 2008 before the burst-out of the global financial crisis (figure 2-4). After the financial crisis and specifically from 2011, the Mediterranean ports started to narrow the gap which in 2020 was at the same level that it was in 2003.

The UK range started the 21st century with a market share of approximately 15% in total European container throughput but gradually was losing ground and at the second decade stabilized approximately at 10%. The Atlantic range showed stability throughout the examined period with a market share of approximately 5.5%. The Baltic/Scandinavian Began with a market share of approximately 6% in 2001 and managed to reach 8% in 2019. Finally, the Black Sea range had a very small share in total European container throughput throughout the first 20 years of 21st century with a peak in 2008, when the port of Costanza served as a regional transhipment hub and received several calls from Far East-Europe services.

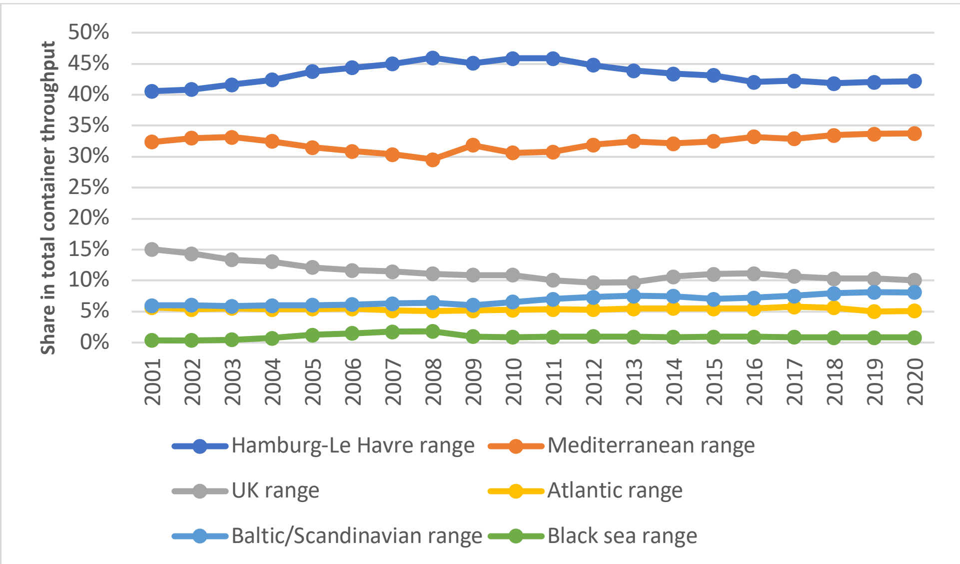

Figure 2-4: Share of European port ranges in total container throughput.
Source: Author's creation based on various sources.

An alternative way to express the dynamics of the market shares among the European container port ranges is by the calculation of annual net shifts. Figure 2-5 presents the annual net shifts of container volumes among the European container ranges. It is observed that the Hamburg-Le Havre range had larger growth rate compared to the growth rate of the total European container throughput in every year until 2008, when this pattern was interrupted by the global financial crisis of 2009. In 2016, the range counted approximately 1.1 million TEUs fewer than it would have counted if its growth rate was the same as the rate of the total European port system. On the other hand, the Mediterranean range gained in 2009 approximately 1.7 million TEUs more than it would have gained if it was growing at the same rate with the total container system.

The UK range underperformed compared to the growth of the total port system in most of the examined years contrary to the Baltic/Scandinavian range which performed above the growth rate of the total system in most years. The Atlantic range recorded its largest annual net shift in 2019 with approximately 659 thousand TEUs less than expected. Finally, the Black Sea range experienced positive net shifts until 2008. It recorded a large negative net shift in 2009 and in the rest years until 2020 grew at the same rate with the total system. All these findings are in agreement with the evolution of the market share of the ranges.

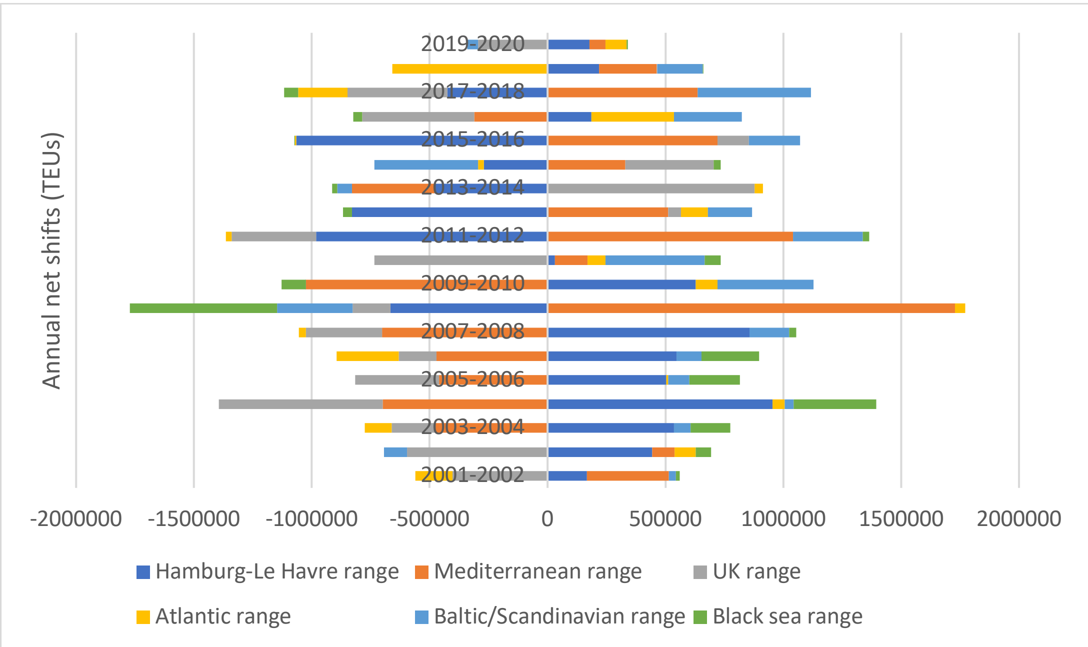

Figure 2-5: Annual net shifts between port ranges in the European container port system.
Source: Author's creation based on various sources.

Figure 2-6 presents the evolution of the *HHI** from 2001-2020 of the port ranges of the European container port system. A higher index means that the container flows are concentrated in a small number of ports, while a smaller index indicates that container traffic is more evenly distributed among the ports of the range. It is observed that the evolution of the index for all ranges, except from the Black Sea range, was rather stable throughout the examined period. This means that the distributional patterns of container flows among the ports of each range did not change much in the first 20 years of 21st century.

The Black Sea range presented a much higher index compared to the other ranges which reached a very high peak in 2007 and then gradually returned to its initial levels. This high index is explained from the huge difference in size between the port of Constanza and the other two ports of the range and the evolution of the index follows the evolution of the port throughput of Constanza. Apart from the special case of the Black Sea range, the second most concentrated range was the Hamburg-Le Havre range due to the presence of the 3 largest container ports in Europe, followed by the UK range, which is dominated by the ports of Felixstowe, Southampton and Thamesport.

The Atlantic range is a deconcatrated range which experienced some moderate concentration during the pre-crisis period 2003-2007 due to the high container volumes handled by the port of Las Palmas in Canary Islands compared to the other ports of the range. Finally, the Mediterranean and the Baltic/Scandinavian ranges, the two ranges including the highest number of ports, 36 and 29 ports respectively, presented a low index under 0.1 throughout the examined period. This means that the container flows were relatively equally distributed among the ports of those ranges.

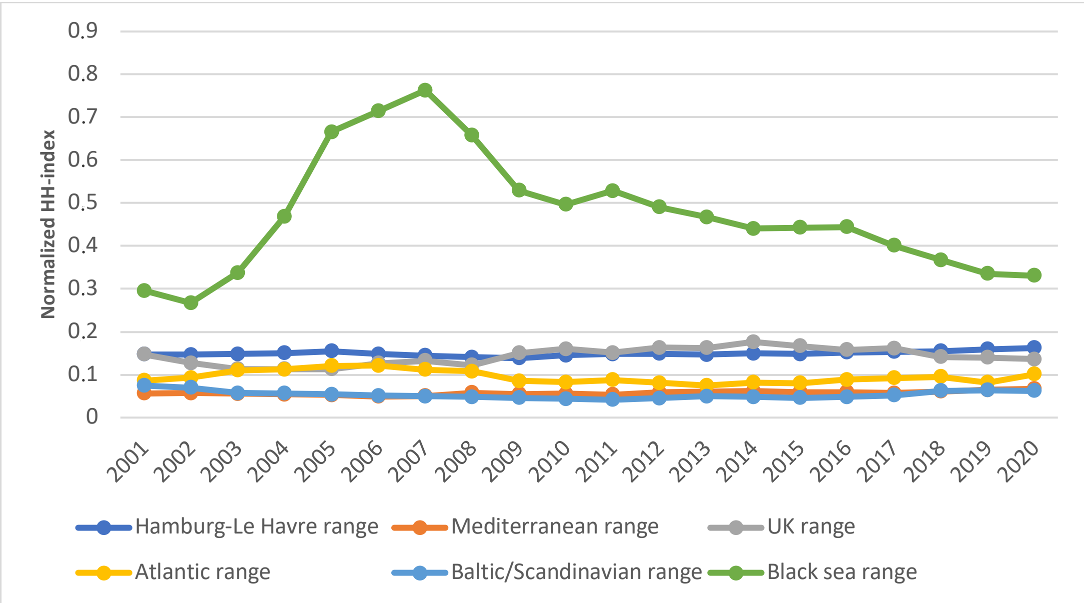

Figure 2-6: Evolution of normalized Herfindahl–Hirschman Index in European port ranges.
Source: Author's creation based on various sources.

2.5 Analysis of multi-port gateway regions

The Rhine-Scheldt multi-port gateway region, including Rotterdam and Antwerp (1st and 2nd largest container ports in Europe), is the largest region in Europe in terms of container handling volumes. Its market share was constantly growing in the first decade of the 21st century, reaching its peak in 2010 by handling approximately 27% of the total European container throughput. A period of decline and stagnation followed until 2016, while from 2017-2020 it experienced year-on-year increase reaching finally, at the closing of the second decade of the 21st century, approximately the level of 2010.

Helgoland Bay, a German multi-port gateway region, is ranked second in the hierarchy of the European container port system. Its market share was increasing until the global financial crisis of 2009. The region was hit hard from the crisis of 2009, especially Hamburg its largest port, and after a very short recovery its market share was decreasing throughout the decade 2011-2020, reaching its lowest point in 2020 by recording a share of 12.82%. UK Southeast coast began in 2001 as the 3rd largest multi-port gateway region but its share was gradually decreasing and in 2020 was surpassed by the Spanish Mediterranean region. The smaller regions followed a relatively stable course throughput the first 20 years of the 21st century (figure 2-7).

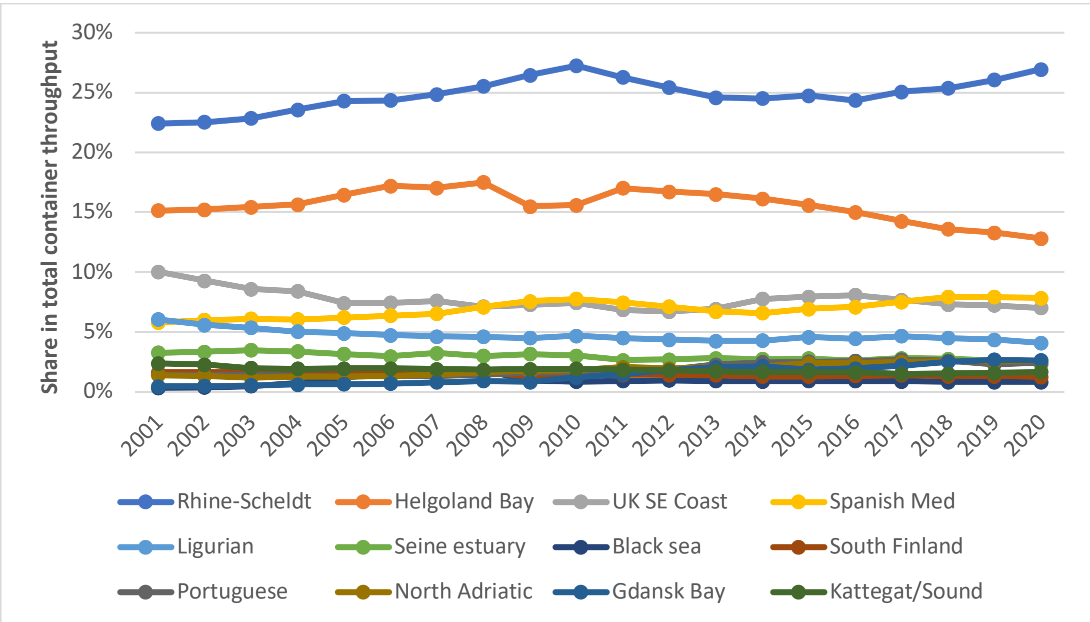

Figure 2-7: Share of European multi-port gateway regions in total container throughput. Source: Author's creation based on various sources.

Figure 2-8 presents the annual net shifts of the European multi-port gateway regions from 2001-2020. The Rhine-Scheldt region recorded positive net shifts during the first decade of the examined period, followed by some years of negative net shifts and returned to growth rates larger than the total system at the end of the second decade by recording in 2020 a positive net shift of approximately 952 thousand TEUs. For the Helgoland Bay region there was a remarkable negative net shift during the financial crisis of 2009 of approximately 1.5 million TEUs and a positive net shift of approximately 1.3 million TEUs after the crisis in 2011. But for every year after 2011 until 2020 the region was growing with smaller growth rates than the total system. The net shifts of all regions confirm the evolution of their market shares in the total European system.

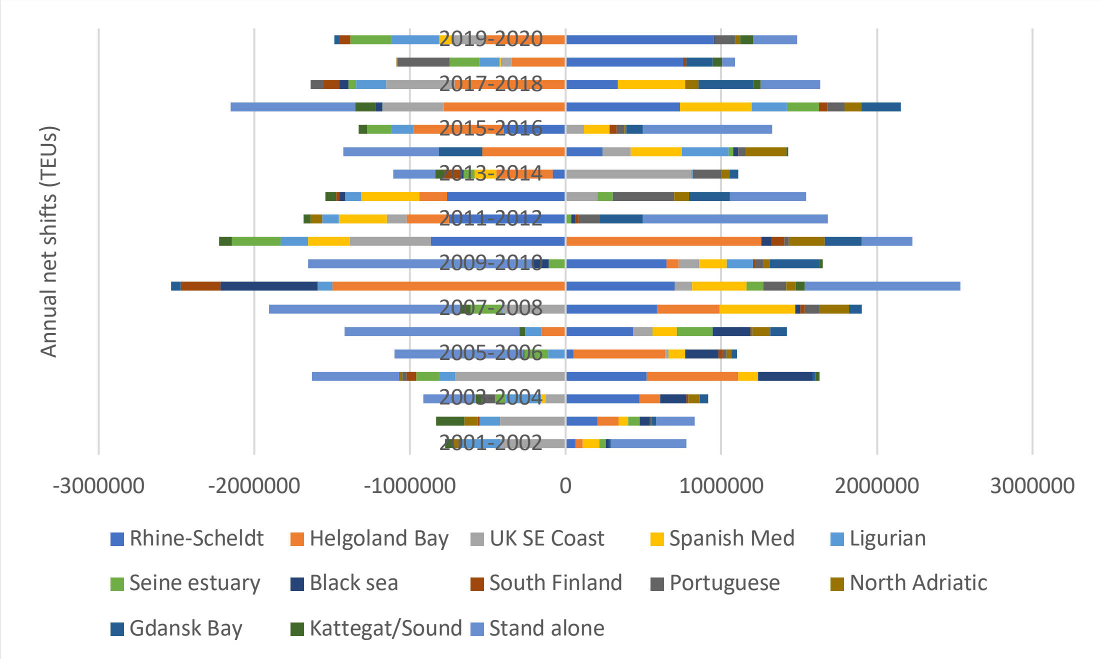

Figure 2-8: Annual net shifts between multi-port gateway regions in the European container port system.
Source: Author's creation based on various sources.

The evolution of the HHI^* of the multi-port gateway regions of the European container port system is illustrated in figure 2-9. The index of Rhine-Scheldt and Helgoland Bay was relatively stable during the examined period indicating that the concentration patterns of the 1st and 2nd largest container port regions of Europe did not change significantly.

The index of UK Southeast coast fluctuated and finally followed a deconcetration trend from 2018-2020 which was justified by the strengthening of the port of Thamesport against the port of Felixstowe and Southampton. On the other hand, the multi-port gateway region of Spanish Mediterranean experienced a concertation trend due to the high growth rate of the port of Valencia compared to that of Barcelona, mainly attributed to the fact that Valencia's port managed to capture large amounts of transhipment container traffic while the port of Barcelona retained in most years of the examined period its pure gateway function.

The Seine estuary is the region with the highest HHI^* index but this fact is totally explained by the presence of only two ports in this region, Le Havre and Rouen, which present a huge difference in their container throughput. The Portuguese region experienced a deconcentration trend as the container terminal of the port of Sines started its operations in 2004 and reached almost a complete balance among the three ports of the region in 2011. Nevertheless, the trend was reversed to concentration from 2012-2020 as the Port of Sines continued its fast rise and was receiving most of the container flows arriving at the region. A similar pattern was observed in Gdansk Bay region with the emergence and continuing growing of the port of Gdansk which matched the container throughput of the port of Gdynia in 2010 and had the double size of its adjacent port in 2020. The region of Kattegat, consisting of Swedish and Danish ports, experienced a strong deconcentration trend as the ports of Aarhus and Helsingborg presented high growth rates while the other ports of the region stagnated. The HHI^* of the rest of the multi-port gateway regions followed a rather stable course.

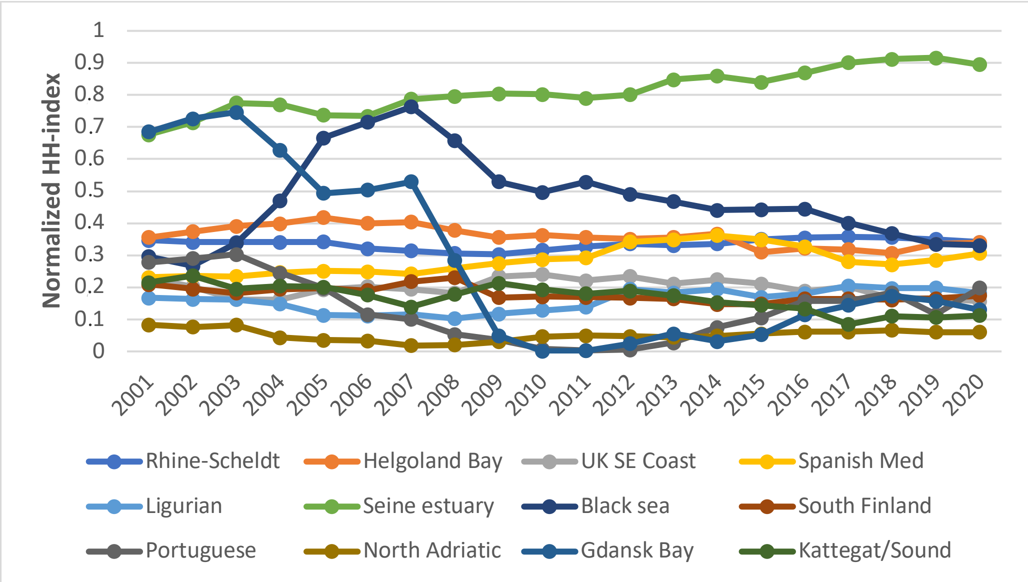

Figure 2-9: Evolution of normalized Herfindahl–Hirschman Index in European multi-port gateway regions.
Source: Author's creation based on various sources.

2.6 Individual container ports

During the examined period from 2001-2020 the European port system experienced two major crises. The global financial crisis which began in late 2008 and was fully developed in 2009 and the health crisis of COVID-19 in 2020. In order to understand how European container ports were affected from these crises the absolute and percentage growth of container ports between 2001 and the pre-crisis year 2007 and between 2007 and 2020 are presented in table 2-1. The absolute and percentage growth between the crises' years and the previous ones are also presented.

Table 2-1: Absolute (in 1,000 TEUs) and percentage growth of top-20 ports in 2020 between selected time periods.

R	Port	Absolute 2001-07	% 2001-07	% 2008-09	Absolute 2007-20	% 2007-20	% 2019-20
1	Rotterdam	4671	76.33%	-9.65%	3559	32.98%	-3.18%
2	Antwerpen	3958	93.84%	-15.63%	3855	47.14%	1.44%
3	Hamburg	5201	110.93%	-28.03%	-1363	-13.78%	-8.04%
4	Piraeus	207	17.79%	53.35%	4064	295.95%	-3.74%
5	Valencia	1536	101.93%	1.44%	2372	77.97%	-0.46%
6	Algeciras	1263	58.68%	-8.47%	1691	49.54%	-0.38%
7	Bremerhaven	1919	64.56%	-17.45%	-121	-2.48%	-1.77%
8	Felixstowe	504	17.75%	-3.53%	149	4.46%	-9.08%
9	Gioia Tauro	957	38.45%	-17.62%	-252	-7.30%	26.58%
10	Barcelona	1199	84.98%	-29.94%	348	13.32%	-11.04%
11	Le Havre	1133	74.35%	-9.96%	-211	-7.95%	-13.39%
12	Marsaxlokk	736	63.18%	-3.13%	539	28.34%	-10.39%
13	Genova	329	21.52%	-13.19%	498	26.83%	-10.04%
14	Gdansk	76	373.13%	29.60%	1827	1885.88%	-7.21%
15	Zeebrugge	1145	130.70%	5.36%	-221	-10.92%	10.30%
16	Southampton	699	59.74%	-14.58%	-107	-5.71%	-6.14%
17	Thamesport	92	12.26%	-32.79%	903	107.05%	-0.97%
18	Sines	150	n/a	8.74%	1462	974.37%	13.26%
19	Marseille	261	35.16%	2.98%	315	31.39%	-9.41%
20	La Spezia	212	21.79%	-16.06%	-13	-1.13%	-16.73%

R: Rank. Source: Author's elaboration based on various sources.

Table 2-2 presents the container throughput of top-20 European container ports in selected years from 2001-2020. The top-3 positions in the rankings were occupied every year of the examined period by the ports of Rotterdam, Hamburg and Antwerp. The port of Rotterdam was always the leading European port and experienced an increase in its share in the total European container throughput from approximately 12% in 2001 to 13.4% in 2020.

The port of Hamburg recorded double digit annual growth rates until the pre-crisis year 2007. In 2007 It had the largest absolute growth among all European ports by gaining approximately 5.2 million TEUs compared to 2001. The financial crisis of 2009 was a turning point for the port as it lost approximately 2.7 million TEUs compared to the previous year. It never really managed to recover from that heavy blow and eventually in 2015 lost the 2nd position in the ranking from the port of Antwerp and remained 3rd until 2020. Its container throughput in 2020 was reduced by 13.78% compared to the pre-crisis year 2007.

Table 2-2: Container throughput of top-20 European ports (in 1,000 TEUs).

R	Port	2001		2007		2009		2015		2019		2020
1	Rotterdam	6120	Rotterdam	10791	Rotterdam	9743	Rotterdam	12235	Rotterdam	14821	Rotterdam	14349
2	Hamburg	4689	Hamburg	9890	Antwerpen	7310	Antwerpen	9654	Antwerpen	11860	Antwerpen	12031
3	Antwerpen	4218	Antwerpen	8177	Hamburg	7008	Hamburg	8821	Hamburg	9272	Hamburg	8527
4	Bremerhaven	2973	Bremerhaven	4892	Bremerhaven	4565	Bremerhaven	5547	Piraeus	5648	Piraeus	5437
5	Felixstowe	2839	Gioia Tauro	3445	Valencia	3654	Valencia	4615	Valencia	5440	Valencia	5415
6	Gioia Tauro	2488	Algeciras	3414	Algeciras	3043	Algeciras	4516	Algeciras	5125	Algeciras	5106
7	Algeciras	2152	Felixstowe	3342	Felixstowe	3021	Felixstowe	4043	Bremerhaven	4857	Bremerhaven	4771
8	Genova	1527	Valencia	3043	Gioia Tauro	2857	Piraeus	3287	Felixstowe	3840	Felixstowe	3491
9	Le Havre	1523	Le Havre	2656	Zeebrugge	2328	Marsaxlokk	3064	Barcelona	3325	Gioia Tauro	3193
10	Valencia	1507	Barcelona	2610	Marsaxlokk	2261	Le Havre	2559	Le Havre	2823	Barcelona	2958
11	Barcelona	1411	Zeebrugge	2021	Le Havre	2241	Gioia Tauro	2547	Marsaxlokk	2723	Le Havre	2445
12	Southampton	1170	Marsaxlokk	1901	Barcelona	1800	Genova	2243	Genova	2615	Marsaxlokk	2440
13	Piraeus	1166	Southampton	1869	Genova	1534	Barcelona	1965	Gioia Tauro	2523	Genova	2353
14	Marsaxlokk	1165	Genova	1855	Southampton	1381	Southampton	1954	Gdansk	2073	Gdansk	1924
15	La Spezia	975	Las Palmas	1450	Las Palmas	1073	Zeebrugge	1569	Southampton	1878	Zeebrugge	1800
16	Zeebrugge	876	Constanta	1411	La Spezia	1046	Sines	1332	Thamesport	1764	Southampton	1762
17	London	752	Piraeus	1373	Marseille	877	La Spezia	1300	Zeebrugge	1632	Thamesport	1747
18	Marseille	742	La Spezia	1187	Goteborg	818	Marseille	1223	Marseille	1455	Sines	1612
19	Goteborg	698	Marseille	1003	Taranto	741	Thamesport	1185	Sines	1423	Marseille	1318
20	Las Palmas	672	Thamesport	844	Cagliari	737	Gdansk	1091	La Spezia	1409	La Spezia	1174

R: Rank. Source: Author's elaboration based on various sources.

The port of Antwerp experienced a growth of 185.2% between 2001 and 2020 and handled over 12 million of container movements for the first time in 2020. Not only it surpassed the port of Hamburg in 2015 but for the next 5 years their gap continued to widen reaching its peak in 2020 when the difference in the container throughput of the two ports was 3.5 million TEUs.

The Mediterranean port of Piraeus experienced the most impressive growth among the top-10 European ports in 2020. Its rise continued even during the financial crisis year 2009, and its throughput almost quadrupled in 2020 compared to 2007. It climbed from number 13 in 2001 to number 4 in 2020 at the European ports' container throughput rankings. This rise is attributed to its acquisition from the Chinese company COSCO shipping. The port of Valencia became the 5th largest container port in Europe in 2008 and kept this position until 2020. Its success is mainly explained from the opening of dedicated container terminals of the shipping line Mediterranean Shipping Company at the port.

The ports of Gdansk and Sines are relatively new players in the market, as the Baltic Hub container terminal started its operations in 2007 in Gdansk and the container terminal in Sines opened in 2004, but they developed very fast and in 2020 were the ports with the highest percentage increase compared to 2007.

The global financial crisis of 2009 affected severely almost all the European ports. Apart from its aforementioned effect on port of Hamburg, the ports of Thamesport, Barcelona and Bremerhaven recorder big losses of 32.79%, 29.94% and 17.45% compared to 2008. The pandemic of COVID-19 in 2020 had much more mild effects compared to the financial crisis. The ports of the top-20 in 2020 that recorded the highest percentage declines were the ports of La Spezia, Le Havre and Barcelona. The port of Barcelona proved to be vulnerable against external shocks as it was heavily affected by both of the crises of 21st century.

Figure (2-10) presents the evolution of the share in the total container throughput of the top-3, top-10 and top-20 container ports. The top-3, top-10 and top-20 container ports concentrated about 32%, 60% and 77.5% respectively of the total container throughput almost throughout the entire period from 2001-2020. Two conclusions are emerging from this observation: First, the largest part of the total European container throughput is concentrated in a few large ports and second the stability of their cumulative shares throughout the entire examined period reveals that the dominance of the large ports remained unchallenged by medium and small-sized ports.

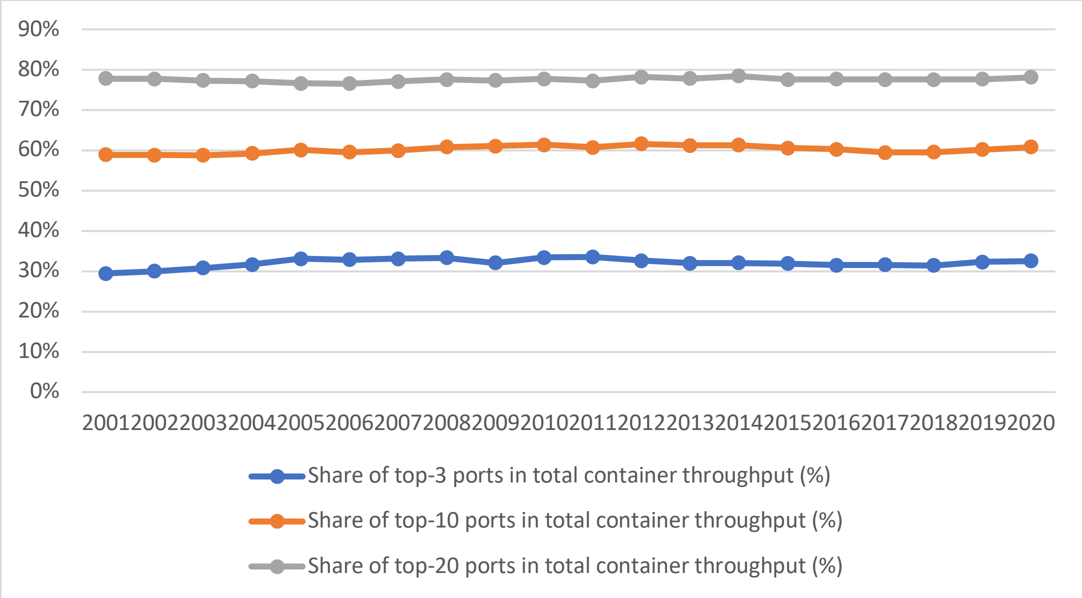

Figure 2-10: Share of top-3, top-10 and top-20 ports in total container throughput from 2001-2020.
Source: Author's creation based on various sources.

2.7 Conclusions

This chapter analyzes the container through dynamics of the European container port system from 2001-2020 updating the studies of Notteboom (1997, 2010).

The share of the total European container throughput in global throughput was constantly decreasing from 2001 to 2010 mainly due to the rise of the container port systems located in Southeast and Northeast Asia. In the following decade, the European container port system managed to stabilize its share in global throughput.

In most of the European container port ranges a concentration tendency prevailed in the first decade of the 21st century and a deconcentration trend in the following decade. The global financial crisis of 2009 worked as a turning point shifting the trend from concentration to deconcentration in the European container port system. The balance in the distribution of container flows between North and South Europe, expressed by the difference in the market shares between the Hamburg-Le Havre and the Mediterranean range, fluctuated during the 20 years of the examined period and in 2020 was at the same level as it was in 2003.

The multi-port gateway regions of the European container port system followed different growth patterns. The Rhine-Scheldt region increased its market share while the Helgoland Bay region experienced a decline mainly due to the effect of the global financial crisis of 2009 on German ports. In 2020, the Mediterranean range emerged as the 3rd largest region above the

UK Southeast coast region. All the regions, apart from the North Adriatic which is the most balanced region in the port system, are highly concentrated as one or two ports in each region receive most of the calls of ocean carriers.

The European container ports experienced large growth rates until the pre-crisis year 2007. In 2009, the year of Great Recession, their container throughput reduced by thousands to millions TEUs depending on the size of the port. In 2020, most of the leading European container ports had recovered from the blow of 2009 with some exceptions, the most characteristic being that of port of Hamburg whose container throughput remained 13.78% below the level of 2007. The health crisis of 2020 had mild effects on European container ports compared to the crisis of 2009 as it did not unfold at the same time and speed all over the world. The large cumulative market shares of the top ports throughout the examined period indicated that a few large ports dominated the European container port system throughout the period from 2001-2020.

Chapter 3

Hierarchy dynamics in the European container port system over the period 2001-2020

3.1 Introduction

The European container port system is a unique mix of ports of different sizes (Notteboom, 2010). Although its container throughput dynamics have been analyzed at the level of multi-port gateway regions (Notteboom, 2010) or at the level of port ranges (Notteboom, 1997; Notteboom, 2010; Wilmsmeier and Monios, 2013) the impact of size categories of ports (large, medium and small-sized port groups) one upon another and the mobility of ports from one size category to another have not been explored. This study aims to analyze the dynamics of port hierarchies in the European container port system by applying methodologies that have been rarely used in this context.

First, two types of rank-size models (Zipf, 1949) are used in order to study the formation of the hierarchies and the dynamics of container throughput distribution. Second, a methodology based on the principals of Analytic Hierarchy Process (AHP) combined with an allometric growth model is introduced in order to investigate the relative growth of large ports as a group upon groups of medium-sized and small ports. Third, a Markov chain modeling is applied in order to explore the mobility of ports within the hierarchy of the European container port system. The dataset that is used in all the methodologies consists of the annual container port throughputs in TEUs of the top-100 European container ports for every year between 2001 and 2020.

The remainder of the chapter is structured as follows: Section 3.2 provides a literature review of conceptual studies of port system development and empirical applications. Section 3.3 introduces the methodology that was applied in this study. Section 3.4 contains descriptive statistics of the European container port system and the top-100 container ports from 2001-2020. Section 3.5 presents the rank-size models that explore the distributional patterns and the evolution of the hierarchies in European container port system. Section 3.6 applies the AHP concept combined with an allometric growth model in order to assess the impact of the parts of the hierarchy upon each other. Section 3.7 provides the results of the Markov chain modeling in order to investigate the intra-distributional mobility of the European container ports. The final section 3.8 presents the conclusions of this study.

3.2 Literature review

The formation of hierarchies in port systems has been studied first with the introduction of theoretical models of port system development and second with empirical applications.

3.2.1 Theoretical models of port system development

Researches on the development of port systems and the formation of hierarchies began to appear in late 1950's. Lee et al. (2008) in a literature review of researches on port system development identified 16 studies on port system development in Western counties from 1957-1999 and 10 studies in developing countries from 1963-2001.

Taaffe et al. (1963) presented a port system development model based on West Africa case. As the system develops, traffic tends to concentrate in a few ports while others even disappear and inland corridors are gradually appearing and interconnecting hinterland centers with the emergence of high priority links at the last phase of the model. Rimmer (1967) proposed a five-phase model for the evolution of the Australian port system. In his model, the port system is developing from concentration to interconnection and centralization and finally to decentralization.

Barke (1986) introduced the phase of deconcentration into port system development models. He noticed that ports were facing congestion problems and had to transfer some of their activities to adjacent locations. Hayuth (1981) also proposed a final phase of deconcentration in his study of the USA port system. He presented a five-phase port system development model where after the phase of concentration of traffic to big load centers a phase of deconcentration occurs, where the smaller ports gain traffic at the expense of the larger ports. He named that phase the "challenge of the periphery".

Notteboom and Rodrigue (2005) added a sixth phase to the port development model called port regionalization. In this phase inland distribution is the most crucial factor in port competition. Ports are connected with the inland distribution centers and terminals through the inland corridors shaping this way an integrated value chain. The development of load center is interdependent with the development of the logistics platforms in its hinterland which finally leads to the creation of a "regional load center network''. Wilmsmeier and Notteboom (2011) identified four phases in the development of liner shipping networks between two differently developed regions.

3.2.2 Formation of hierarchies in empirical applications

Hierarchy in a port system has been studied mainly on the basis of two measures: throughput volumes and network indicators (Oliveira et al., 2021). The use of throughput volumes for the construction of the hierarchy in a port system is the most classical approach. The main focus of the literature, when ports in a port system are ranked according to their throughput in TEUs or tonnage, is on the identification of concentration or deconcentration patterns. The methods that are used in order to measure the level of (de)concentration in a port system are Gini coefficient, Herfindahl-Hirschman Index (HHI), Lorenz curve and Shift-Share analysis. Ducruet and Notteboom (2022) in a literature review on port system studies identified 34 studies on port systems in North Europe published from 1997 to 2018 and 28 studies on port systems in Northeast Asia published from 2000 to 2018. They observed a great diversity in the definition and the delineation of port systems in the academic literature. The geographical scales of analysis vary greatly from multi-port gateway regions to whole continents. In a literature review of port concentration studies from 1963-2008 Ducruet et al. (2009a) identified 32 studies.

Notteboom (1997) studied the (de)concentration tendencies in the European container port system from 1980-1994 and observed that port concentration levels stagnated in the 1990's. In an update of this study in 2010, (Notteboom, 2010) he extended the period of study from 1995 to 2008 and concluded that that European container port system witnessed a gradual deconcentration process. Wilmsmeier and Monios (2013) analyzed the UK container port system. They observed a potential deconcentration of container traffic within the UK container port system related to two factors: a shift in gateway region for UK trade, with container traffic transshipped through continental ports rather than the southeastern UK ports, as well as a shift in role at UK ports from gateways to transshipment hubs. Fremont and Soppe (2007) analyzed the concentration of container traffic in the North European range at the level of individual shipping companies. They observed high levels of concentration in the port networks of individual carriers as their vessels called most frequently in a few dedicated container terminals.

Wang and Ducruet (2012) examined the evolution and the concentration of container traffic at the Yangtze River Delta (YRD) from 1979 to 2010. They noticed that traffic concentration levels have been irregular from early 1980s to the mid- 1990's explained by the spread of container technology outside of the port of Shanghai which was favored at the introductory phase of containerization. Since the late 1990s, the YRD port system experienced a growing concentration at a few ports before stabilizing in the early 2000s and then passing to deconcentration until 2010. Wang and Ducruet (2013) also studied the concentration levels in the Chinese port system in a long-term period from 1868 to 2009 and identified three main traffic regimes: high hierarchy (1868-1932), medium hierarchy (1936-1978) and low hierarchy (1986-2009.) Notteboom (2006) analyzed the East Asian container port system from 1980 to 2004 and demonstrated that it witnessed a deconcentration trend as a result of the rise of the Chinese ports and the relative stagnation of the Japanese range.

Le and Ieda (2010) developed a geo-economic concentration index and measured the concentration levels of the container port systems of Japan, China and Korea from 1975 to 2005. They found diversified concentration tendencies for the three countries. China experienced a deconcentration tendency up to 1990 and then a strong concentration tendency, especially after 1995. On the other hand, Korea presented a concentration tendency up to 1995 and after that a deconcentration trend with a slight concentration pattern again after 2000. Japan experienced a steady level of concentration throughout the whole period. These differences in concentration trends were attributed to different government policies and port governance structures among the countries.

Wilmsmeier et al. (2014) in their study of the evolution of the Latin American and the Caribbean port system from 1997 to 2012 observed in most ranges of the port system an initial concentration followed by deconcentration. The rise of the secondary ports was attributed to the introduction of the private sector to port operations and the investments of shipping lines and terminal operators in competitor ports.

Port hierarchies in a port system have also been studied in researches that analyzed maritime shipping networks in global or regional scale. In these studies, the hierarchy is not based on the volume of traffic handled at the ports but on their position in the network. The most common network indicators used for the ranking of ports are degree centrality (number of connections of a port) in order to assess its connectivity and betweenness centrality (number of shortest paths between any pair of ports in the network that pass through a specific port) in order to evaluate how centrally is a port placed in the network (Cisic et al., 2007; Ducruet et al., 2009b; Ducruet et al., 2010b; Ducruet and Notteboom, 2012a,b; Gonzalez et al., 2012; Pais-Montes et al., 2012; Freire et al., 2013; Tovar et al., 2015; Ducruet, 2013a,b; Kang and Woo, 2017). Cullinane and Wang (2012) constructed a port hierarchy in a sample of 39 ports in the main East-west routes based on the number of significant connections that each port has with the other ports of the network.

The correlation between the two types of hierarchy, based on throughput volumes the first and on the network characteristics of ports the second, has been also explored (Deng et al., 2009;

Ducruet et al., 2011b; Kang and woo 2017). Ducruet et al. (2011b) studying the global container shipping network in 2006, used a factor analysis and confirmed the strong correlation between the port throughput and the centrality of a port but also identified geographic distance between ports as an important factor that influences the volume of container throughput. Kang and Woo (2017) in their study of the evolution of global container shipping network from 2006 to 2011 confirmed the findings of Ducruet et al. (2011b) about the strong influence of centrality and geographic distance between ports on port throughput. Deng et al. (2009) in their analysis of the worldwide container shipping network found strong correlation between port throughput and the connectivity of ports in the network.

3.3 Methodology

This study analyzes the dynamics of hierarchies in the European container port system with the application of methods different from the commonly used concentration indicators. The methodology consists of three parts: In the first part, two types of rank size models are used in order to identify the dynamics of the distributional patterns of traffic in the European container port system and the importance of medium-sized ports in the distribution. In the second part, a methodology based on the principals of AHP combined with an allometric growth model is introduced in order to evaluate the relative growth of size categories of ports one upon another. In the third part, the mobility of ports within the hierarchy is examined by the application of Markov chain modelling.

3.3.1 Rank-size models of port hierarchy

The rank size law was introduced by linguist George Kingsley Zipf, a Harvard linguistics professor. He ranked the frequency of the appearance of words in texts and found that their frequency is inversely proportional to their rank. The rank-size law has been used since then in many research fields such as city-size distributions and income distributions. In this study all the European container ports are ranked according to their container throughputs in TEUs (the ports of European Union-28 countries are included in this study apart from ports belonging to overseas territories of European countries and the islands of Azores and Madeira). The top-100 ports of each year from 2001-2020 are selected for the analysis. In the port context the original form of Zipf's law can be expressed as a one parameter model:

$$T(r) = \frac{T_1}{r} \tag{3.1}$$

where r refers to port rank (r =1,2,3….), T(r) to the size of the r[th] port and T_1 is the size of the largest port. This means that the 2[nd] port in the ranking has 1/2 of the size of the 1[st] port, the 3[rd] port has 1/3 of the size of the 1[st] port and so on. In this study it is used the generalized two parameters Zipf's model:

$$T(r) = T_1 r^{-b} \tag{3.2}$$

where b is a scaling exponent and the other symbols the same as in the one parameter model. Equation (3.2) can be converted to:

$$\log T(r) = a - b \log r \tag{3.3}$$

where $a = \log T_1$ and b is the scaling exponent.

The scaling exponent b is estimated with a linear fitting of equation (3.3) with the ordinary least square method. The exponent b can be used as an indicator of the concentration level of container traffic at the port system. A high exponent b means that there is strong inequality at the distribution of container throughput which is concentrated at the larger ports. On the contrary, a low exponent b indicates that medium-sized and small ports gain traffic at the expense of the larger ports and the container throughput is more equally distributed. In this study the exponent b is estimated for every year for a period of 20 years. This approach permits to observe the evolution of the scaling exponent and draw conclusions about the concentration or deconcentration tendency of the port system. When the exponent is increasing the port system tends to be more concentrated. In contrast, when the exponent is decreasing the tendency towards deconcentration prevails.

In the next step, the deviation of equation (3.3) from linearity is examined (Rosen and Resnick, 1979). Equation (3.3) is fitted with a quadratic model in the form of:

$$\log T(r) = a' + b'^{\log r} + \gamma \log r^2 \tag{3.4}$$

The γ coefficient is estimated. Coefficient γ defines the direction of the curvature of the trend line. When γ>0 an upward concavity is observed and when γ<0 a downward concavity. The meaning of the upward concavity is that the presence of medium-sized ports in the rank-size distribution is weak. On the contrary, downward concavity means that there is an abundance of medium-sized ports (Vining, 1976; Oliveira et al., 2021).

3.3.2 Analytic Hierarchy Process (AHP)-Allometric growth model

The rank-size model of the top-100 ports is a useful tool in order to monitor the concentration or deconcentration tendencies of the overall European container port system. The next step includes a more detailed approach to the growth dynamics of the port system. Typical port throughput concentration techniques do not allow us to explore the impact of one part of the distribution upon another. The methodology that is applied in this study allows us to investigate the growth of large ports as a group upon groups of medium-sized and small ports. This study intends to reveal the relationships between the growth rates of different port groups of the rank-size distribution and construct a methodology in order to compare the development of each group of the port system.

The idea behind the methodology that is introduced in this study derives from the AHP method developed by Saaty in 1971-1975. The use of this method allows us to obtain ratio scales from paired comparisons (Saaty, 1987). In this research pairs of port groups are compared and the ratio of their growth rates is estimated. The top-100 ports of every year from 2001 to 2020 were divided into five groups of 20 ports each (table 3-4). The differentiation from Saaty's methodology lies on the way that the ratios are estimated. Saaty used values from a fundamental scale introduced by him. In this study, the relationships between the relative growths of pairs of port groups are assessed with an allometric growth model and the evaluation of the growth rate of each group when all groups are compared with each other with the construction of an allometric growth matrix.

Allometry is a concept that comes from the field of biology. It describes the scaling relationship between two traits or processes (Shingleton, 2010). The term allometry was introduced by Julian Huxley and Georges Tessier in 1936 (Huxley and Tessier, 1936) when they studied the relationship between relative growths of two body traits of crabs. In port studies, as far as it is in the author's knowledge, it has been used once by Chen et al. (2019).

An allometric differential equation has the form (Broad, 1998):

$$\frac{dy}{y} = c\,\frac{dx}{x} \tag{3.5}$$

The meaning of this equation is that the ratio of the growth rates of two characteristics y and x is a constant c. Integrating both sides of equation (3.5) leads to:

$$\ln y = c \ln x + b \tag{3.6}$$

Supposing $\ln a = b$ equation (3.6) can be written as:

$$y = ax^c \tag{3.7}$$

Equation (3.7) shows that if two characteristics y and x are following an allometric growth model they are related to each other by a power function. In this study these characteristics are the container throughputs of the port groups that emerged from the discretization of the rank-size distribution (table 3-4). Equations (3.6) and (3.7) for all the pairs of port groups can be written as:

$$\ln T_i = c_{ij} \ln T_j + b \tag{3.8}$$

$$T_i = a{T_j}^{c_{ij}} \tag{3.9}$$

Where T_i, T_j are the cumulative container throughputs of the ports belonging to groups i and j, with 1≤i,j≤5 and i≠j, c_{ij} is the allometric coefficient and a, b constants. The allometric coefficient c_{ij} can be obtained from equation (3.8) with linear regression. The allometric coefficient c_{ij} as defined in equation (3.5) expresses the ratio of the growth rates of group i and group j:

$$c_{ij} = \frac{c_i}{c_j} \tag{3.10}$$

If c_{ij} >1 => $c_i > c_j$ it is the case of positive allometric growth and the growth rate of group i is higher than the growth rate of group j.

If c_{ij} <1 => $c_i < c_j$ it is the case of negative allometric growth and the growth rate of group i is lower than the growth rate of group j.

If c_{ij} =1 => $c_i = c_j$ it is the case of isometric growth and both groups are growing with the same growth rate.

In order to evaluate the relative growth of a port group against all other groups the allometric coefficients of all the pairs of port groups are estimated and an allometric growth matrix is constructed:

$$M = \left[c_{ij}\right]_{n\times n} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} = \begin{bmatrix} \frac{c_1}{c_1} & \frac{c_1}{c_2} & \cdots & \frac{c_1}{c_n} \\ \frac{c_2}{c_1} & \frac{c_2}{c_2} & \cdots & \frac{c_2}{c_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{c_n}{c_1} & \frac{c_n}{c_2} & \cdots & \frac{c_n}{c_n} \end{bmatrix} = \left[\frac{c_i}{c_j}\right]_{n\times n} \qquad (3.11)$$

In order to derive the vector of allometric coefficients from the allometric growth matrix the eigenvalue method is going to be used (Saaty, 1987). The allometric growth matrix M and its eigenvector A are following the next formula:

$$MA = \lambda A \qquad (3.12)$$

It is observed that:

$$\left[\frac{c_i}{c_j}\right]_{n\times n} [c_1 \quad c_2 \quad \dots \quad c_n]^T = n\, [c_1 \quad c_2 \quad \dots \quad c_n]^T \qquad (3.13)$$

So, A = $[c_1 \quad c_2 \quad \dots \quad c_n]^T$ is the eigenvector of the allometric matrix M and its values represent the growth rates of the port groups. Consequently, the estimation of the values of the eigenvector A provides the weight of each port group. The method that is used in this study in order to estimate the values of eigenvector A is normalizing the elements in each column of the allometric matrix and then averaging over each row (Saaty, 1987).

The final step is to check how good the estimation of the eigenvector of the allometric matrix is. This is implemented with the calculation of the Consistency Index (CI) by the following equation:

$$CI = \frac{|\lambda_{max} - n|}{n-1} \qquad (3.14)$$

where λ_{max} is the maximum eigenvalue calculated by adding the columns of M and multiply the resulting vector with the eigenvector A in normalized form (Saaty, 1987):

$$\lambda_{max} = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\ w_i \qquad (3.15)$$

Where w_i are the elements of the eigenvector A in normalized form.

The estimation of the eigenvector of the allometric matrix is acceptable when the value of CI is close to zero (Chen et al., 2019).

3.3.3 Markov chain modeling on ports mobility

The methodology that is applied at the previous part of this study reveals the relative growth of large ports as a group upon groups of medium and small ports, but it does not explore the movements of individual ports within the distribution. The aim of the third part of this study is to investigate the mobility of ports within the hierarchy of the European container port system. This intra-distributional mobility is explored with the application of Markov chain modelling. The Markov chain approach permits the observation of the proportion of port movements between groups that contain small, medium or large ports (Oliveira et al., 2021). Markov chains techniques have been used in a large extend in urban studies in order to investigate intra-distributional mobility (Quah, 1993; Eaton and Eckstein, 1997; Dobkins and Ioannides, 2000; Black and Henderson, 2003; Schaffar and Dimou, 2012) but only once in port studies (Oliveira et al., 2021).

Markov chains refer to transitions from one state to another. In order to apply the Markov chain modelling the rank-size distribution of ports must be divided into a number of discrete groups of ports. The discretization process that is used in this study in each period is based on quantiles (bottom 20%, second 20%,.., top 20%) as for the allometric growth model. A common discretization process that is used in the majority of studies (Quah, 1993; Eaton and Eckstein, 1997; Dobkins and Ioannides, 2000; Black and Henderson, 2003; Schaffar and Dimou, 2012; Oliveira et al., 2021) is based on defining the cut-off points of the groups exogenously, relative to the mean size of that period (e.g. the cut-off points of five groups of cities are defined as 0.25m, 0.5m, m, 2m in Quah (1993), m mean city size). The disadvantage of this discretization process is that a movement from one group to another doesn't necessarily mean a change in rankings (Overman and Ioannides, 2001). On the contrary, in discretization based on quantiles the movement of a port one group up is followed by a movement one group down by another port and vice versa. Consequently, all movements between groups lead to changes in rankings (Overman and Ioannides, 2001).

If we assume that the distribution follows a homogenous first order stationary Markov chain and F_t is the vector of distributional shares of each group at time t, the vector F_{t+1} after one step transition is given by:

$$F_{t+1} = F_t\,M \tag{3.16}$$

where M is the transition probability matrix from time t to t+1. The Markov property of a discrete Markov chain means that the probability of a port moving to another size group depends only on his present size group and not the previous size groups. In this way, if T_t is the container throughput of a port belonging to a size group i in time t, the probability $p_{ij,t}$ of moving to another size group j in time t+1 is given by:

$$Pr\left(T_{t+1} = j \middle| T_0 = i_0, T_1 = i_{1,\ldots,} T_t = i_t\right) = Pr(T_{t+1} = j | T_t = i_t) \qquad (3.17)$$

The transition probabilities p_{ij}, with maximum likelihood estimation, are the total number of ports moving from group i to j over 20 years divided by the total number of ports starting in i in these 20 years. The matrix of mean first passage times is also estimated. The mean first passage time is the expected number of years required for a port to reach group j for the first time when starting from group i. The transition matrix indicates the extent of movements of ports in the rank-size distribution and the matrix of the mean first passage time the speed with which these movements are made.

3.4 Descriptive statistics of the European container port system

3.4.1 Port throughput and growth rate development of the European container port system

The European container port system is one the busiest in the world. At the begging of the 21st century the total container throughput of the European ports was approximately 51 million TEUs and in 2017 it had surpassed 100 million TEUs. The container throughput of the European ports experienced an increase of 110.5% from 2001 to 2020, from 51,010,556 TEUs in 2001 to 107,388,052 TEUs in 2020 (122 ports). The port system reached its peak in 2019 with its container throughput recording 111,428,666 TEUs (figure 3-1).

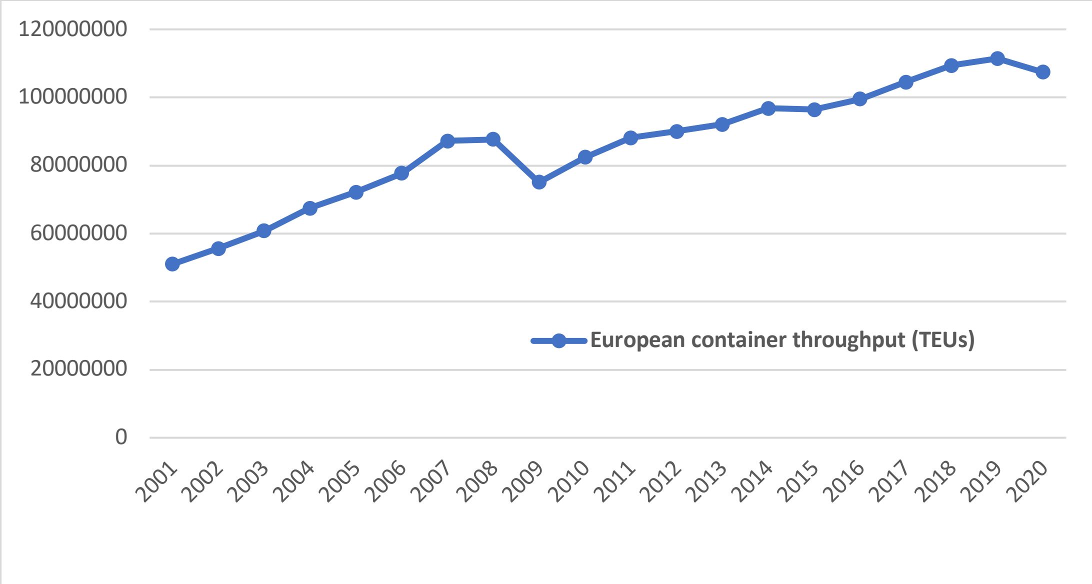

Figure 3-1: Evolution of the total container throughput of the European port system 2001-2020 (122 ports).
Source: Author's elaboration based on data from various sources.

Every year from 2001 to 2020 the European container port system had a positive growth rate with the exception of the years 2009, 2015 and 2020 (figure 3-2). The European container port system experienced a period of high growth from 2001 until the pre-crisis year 2007 with an average growth rate of its container throughput of 9.35%.

The container traffic handled at European ports stagnated in 2008 and experienced its biggest drop in 2009 due to the global financial crisis that burst out in late 2008 and had its full effect in 2009. The European container port throughput had a decrease of 14.36% in 2009 compared to 2008 which translated in a loss of approximately 12.6 million TEUs. After the crisis of 2009 the container throughput of European ports returned to growth but with a modest average growth rate of 4.07% from 2010 until 2019. In 2020 another global crisis, the Covid-19 pandemic, led to decrease of container volumes at the port system. The drop was much smaller than the financial crisis of 2009 with a decline of 3.63% compared to 2019.

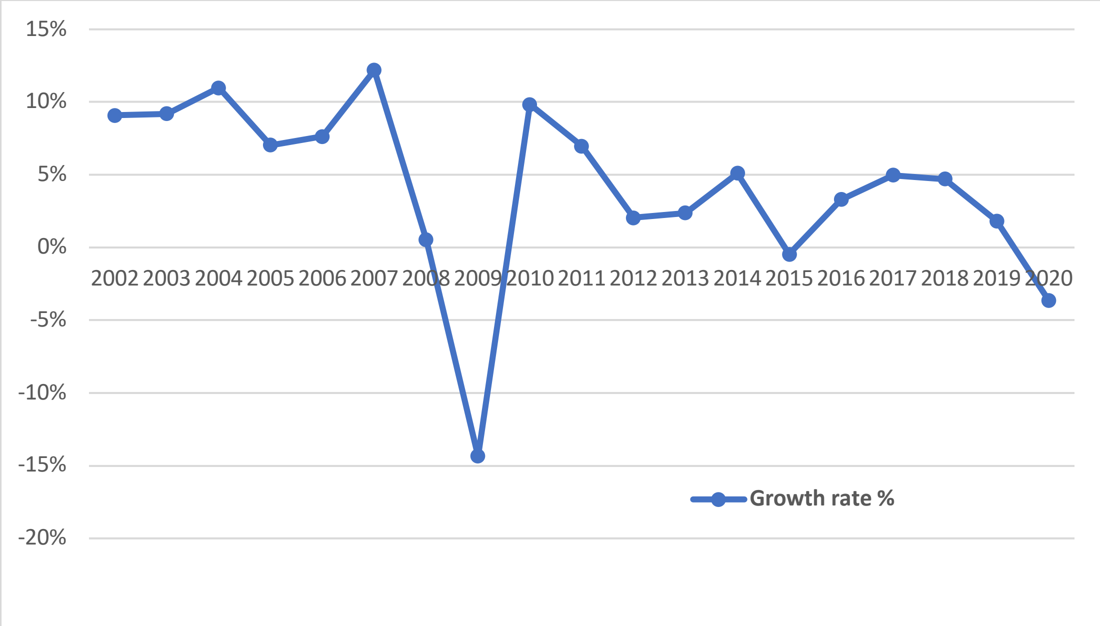

Figure 3-2: Annual growth rate of the container throughput of the European port system (122 ports).
Source: Author's elaboration based on data from various sources.

3.4.2 Descriptive statistics and market share of the top-100 European container ports

The dataset consists of the annual container port throughputs of all European ports in TEUs for the period 2001-2020. The data were collected from various sources: port authorities' websites, annual reports of port authorities, national ports associations statistics, academic and other port related publications and Eurostat. The ports are ranked every year according to their throughputs in TEUs and the top 100 ports are selected for the analysis conducted in this research. Table 3-1 provides descriptive statistics of the data for every year of the selected period.

Table 3-1: Descriptive statistics of top-100 ports from 2001-2020 (1,000 TEUs).

Year	Mean port size	Rank of mean port	Standard Deviation	Median port size (Q50)	First quartile (Q25)	Third quartile (Q75)
2001	509.9	23	993.9	131.0	48.1	438.0
2002	555.6	21	1085.5	140.5	56.8	468.2
2003	605.7	22	1195.7	148.1	54.3	512.5
2004	671.9	22	1347.9	174.5	61.0	534.2
2005	719.0	22	1483.3	180.6	66.9	632.3
2006	774.3	22	1585.0	219.6	79.3	650.4
2007	868.5	20	1787.8	227.3	88.7	727.0
2008	873.2	19	1819.7	236.0	96.4	675.4
2009	747.9	19	1535.8	207.5	103.1	578.0

2010	821.5	20	1730.8	217.2	88.3	556.3
2011	878.2	20	1856.9	230.5	92.9	615.6
2012	895.7	21	1877.7	241.4	88.7	632.2
2013	916.5	22	1884.8	251.0	83.8	652.5
2014	963.5	22	1984.6	228.0	86.2	672.6
2015	959.6	21	1957.5	244.7	90.4	733.7
2016	990.5	21	2000.6	242.9	95.3	726.0
2017	1039.8	22	2106.0	267.1	91.3	707.6
2018	1089.0	21	2203.8	261.8	102.5	752.3
2019	1108.3	21	2284.3	267.7	105.5	789.8
2020	1068.4	21	2227.2	262.7	104.2	776.0

Source: Calculated by the author based on various sources.

The mean port size every year was much higher than the third quartile. In every year around 80% of the ports were below the mean size and 20% above it. This means that the European container port system consists of a few very large container ports and a big number of small and medium-sized ports. The evolution of the median port size (Q_{50}), the first quartile (Q_{25}) and the third quartile (Q_{75}) reveals that the smaller ports were growing with higher growth rates compared to the medium-sized ports during the first decade (2001-2010) and the medium-sized ports were growing faster than the small ports during the second decade (2011-2020).

The cumulative market share of the top 100 European container ports is presented in figure 3-3. Firstly, it is observed that a few large ports dominate the market which confirms the finding of the descriptive statistics for the presence of a small group of large ports and a very large group of small and medium- sized ports. The top 10 ports concentrated 58.6% of the total throughput in 2001, 61.3% in 2010 and 60.8% in 2020. The market share of the top 20 ports was 76.1% in 2001, 76.8% in 2010 and 77% in 2020. Secondly it is remarkable that the cumulative market shares of the ports were very stable throughout the period 2001-2020.

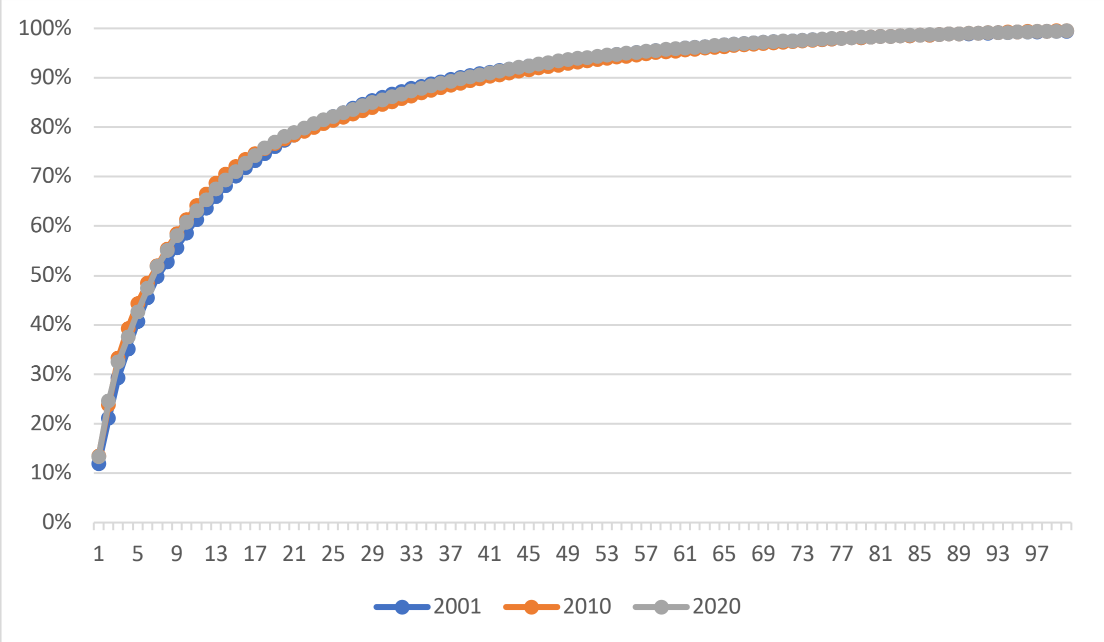

Figure 3-3: Cumulative market share of the top-100 European container ports.
Source: Visualized by the author based on data from various sources.

3.5 Results of rank-size models of port hierarchy

The results of the linear regressions of equation (3.3) are presented in table 3-2.

Table 3-2: Evolution of exponent b of the rank-size model from 2001-2020.

Year	Zipf's exponent b	Standard Deviation σ(b)	R^2
2001	1.493	0.039	0.937
2002	1.465	0.036	0.943
2003	1.493	0.040	0.935
2004	1.488	0.040	0.935
2005	1.469	0.039	0.934
2006	1.463	0.041	0.929
2007	1.464	0.040	0.932
2008	1.452	0.038	0.938
2009	1.427	0.036	0.941
2010	1.427	0.034	0.948
2011	1.429	0.035	0.944
2012	1.457	0.036	0.944
2013	1.461	0.035	0.946
2014	1.491	0.037	0.944
2015	1.475	0.038	0.938
2016	1.458	0.036	0.942

2017	1.469	0.039	0.936
2018	1.477	0.039	0.936
2019	1.475	0.038	0.939
2020	1.484	0.037	0.943

Source: Retrieved by the author based on the input data.

The evolution of the exponent b of the rank-size model reveals two different tendencies during the examined period. At first, prior the economic crisis of 2009, from 2001-2009 the exponent was decreasing which means that the port system was tending towards deconcentration. Subsequently, after the crisis of 2009, from 2010-2020 the exponent was increasing which means that the port system was directed towards concentration of container traffic (figure 3-4).

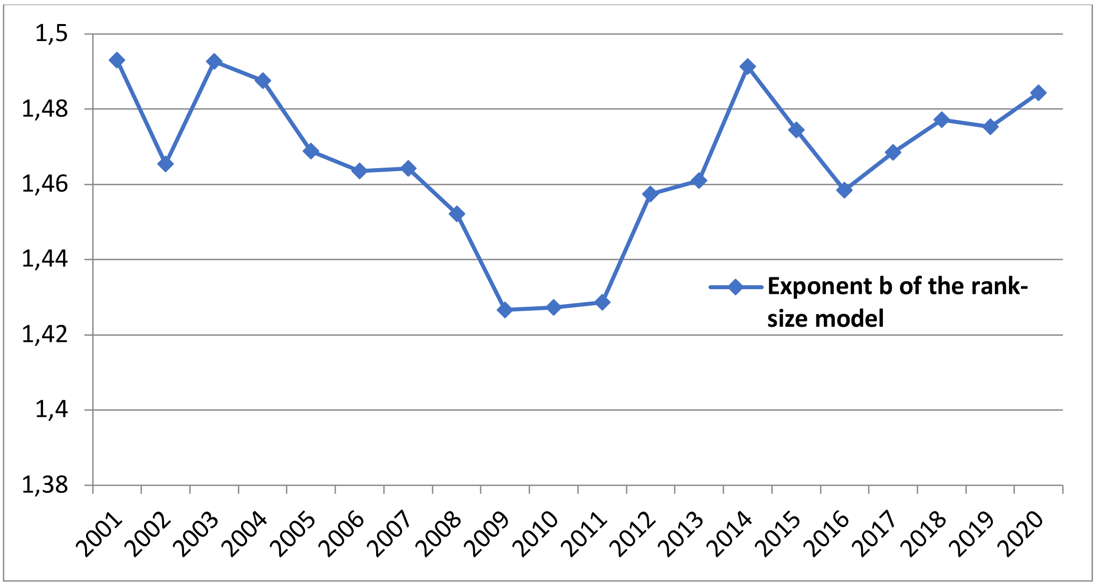

Figure 3-4: Evolution of the exponent b of the rank-size model for the European port system.
Source: Visualized by the author.

The results of the fitting of the quadratic model are presented in table 3-3.

Table 3-3: Evolution of coefficient γ of the quadratic model from 2001-2020.

Year	Quadratic coefficient γ	Standard Deviation σ(γ)	R^2
2001	-0.656	0.016	0.997
2002	-0.612	0.015	0.997
2003	-0.661	0.020	0.995
2004	-0.654	0.021	0.994
2005	-0.629	0.027	0.990
2006	-0.644	0.030	0.988
2007	-0.628	0.029	0.988
2008	-0.585	0.030	0.987

2009	-0.551	0.030	0.987
2010	-0.527	0.025	0.991
2011	-0.545	0.028	0.989
2012	-0.568	0.026	0.991
2013	-0.575	0.020	0.994
2014	-0.609	0.019	0.995
2015	-0.619	0.024	0.992
2016	-0.596	0.021	0.994
2017	-0.621	0.026	0.991
2018	-0.633	0.023	0.993
2019	-0.617	0.023	0.993
2020	-0.610	0.019	0.995

Source: Retrieved by the author based on the input data.

The quadratic coefficient γ is negative for every year of the entire period of study. This means that there is an important presence of medium-sized ports throughout the examined period. The evolution of the quadratic coefficient γ reveals two different tendencies for the growth of the medium-sized ports. At the first decade of study, from 2001-2010, the quadratic coefficient was decreasing which means that the influence of medium-sized ports in the distribution was decreasing. At the second decade of study, from 2011-2020, the quadratic coefficient was increasing which means that the influence of medium-sized ports was increasing (figure 3-5).

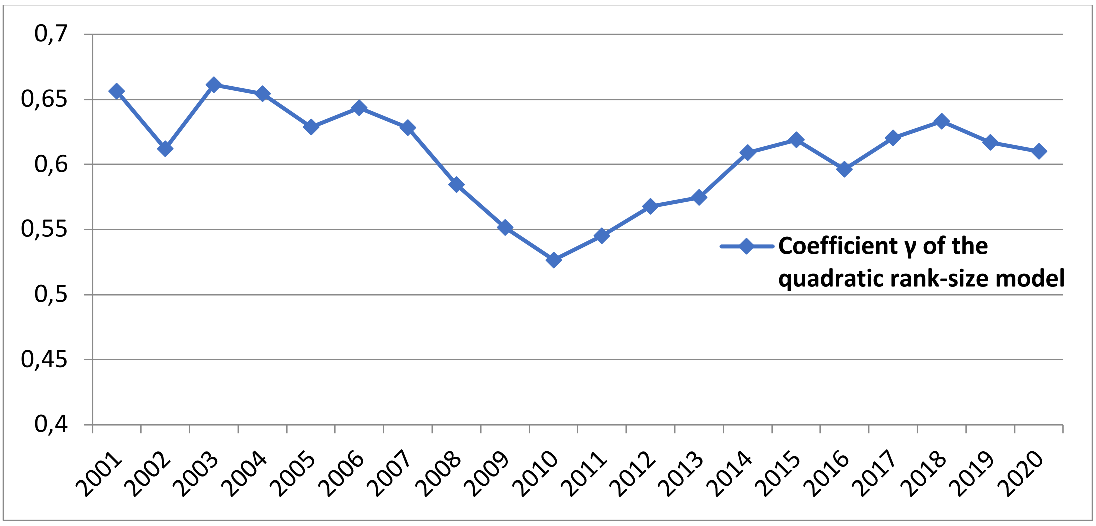

Figure 3-5: Evolution of the quadratic coefficient γ for the European port system.
Source: Visualized by the author.

Combining the findings of the evolution of the two coefficients b and γ some interesting conclusions can be drawn. During the deconcentration period (2001-2009), where container traffic was more evenly distributed, the medium-sized ports weren't those who took advantage of the situation but the small ports were growing faster. On the contrary, during the concentration period (2010-2020), where container traffic was concentrated into larger ports, medium sized ports were growing faster than the small ports and weren't overshadowed by the larger ports. It seems that deconcentration of container traffic favors the small ports while concentration favors, apart from the large ports, also the medium-sized ports.

3.6 Results of Analytic Hierarchy Process (AHP)-Allometric growth model

The top-100 ports of every year from 2001 to 2020 are divided into five groups of 20 ports each (table 3-4).

Table 3-4: Discretization of the top-100 ports into groups and descriptive statistics for selected years (1,000 TEUs).

			2001			2010			2020	
Port Groups	Ports included	Min port size	Max port size	Mean port size	Min port size	Max port size	Mean port size	Min port size	Max port size	Mean port size
G1	81-100	24,6	45,1	31,7	42,2	66,0	55,1	50,0	85,0	65,6
G2	61-80	45,4	85,6	61,2	70,0	152,6	114,7	86,3	173,1	128,0
G3	41-60	90,0	175,8	128,8	153,0	348,4	230,6	178,1	422,3	283,3
G4	21-40	198,0	531,8	344,5	357,5	661,8	503,6	453,0	951,8	672,1
G5	1-20	672,0	6119,5	1983,0	732,7	11147,6	3203,5	1173,7	14349,4	4192,7

Source: Elaborated by the author based on various sources.

The allometric growth model for any pair of port groups is constructed based on equation (3.9). The allometric coefficient for any of these pairs of groups is estimated by fitting linearly equation (3.8). Then an allometric growth matrix is created with the obtained allometric coefficients from all the group pairs (table 3-5).

Table 3-5: Allometric growth matrix of port groups.

	G1 (81-100)	G2 (61-80)	G3 (41-60)	G4 (21-40)	G5 (1-20)
G1 (81-100)	1	0.9148	0.8527	1.0734	0.9735
G2 (61-80)	0.9592	1	0.8664	1.0699	0.959
G3 (41-60)	1.124	1.0893	1	1.2593	1.1219
G4 (21-40)	0.8362	0.795	0.7442	1	0.8595
G5 (1-20)	0.9925	0.9325	0.8676	1.1247	1

Source: Elaborated by the author.

In order to estimate the eigenvector values of each port group the normalized allometric growth matrix is constructed (table 3-6).

Table 3-6: Normalized allometric growth matrix of port groups.

	G1 (81-100)	G2 (61-80)	G3 (41-60)	G4 (21-40)	G5 (1-20)
G1 (81-100)	0.203587	0.193338	0.196887	0.1942	0.198111
G2 (61-80)	0.195281	0.211345	0.200051	0.193566	0.195161
G3 (41-60)	0.228832	0.230218	0.230899	0.227833	0.228312
G4 (21-40)	0.17024	0.168019	0.171835	0.18092	0.174912
G5 (1-20)	0.20206	0.197079	0.200328	0.203481	0.203504

Source: Elaborated by the author.

The eigenvector values of the port groups are then estimated (table 3-7).

Table 3-7: Eigenvector values of the port groups.

Port group	G1 (81-100)	G2 (61-80)	G3 (41-60)	G4 (21-40)	G5 (1-20)
Eigenvalues	0.197225	0.199081	0.229219	0.173185	0.201291

Source: Calculated by the author.

In order to check how accurate, the estimation of the eigenvector values is, λ_{max} is calculated according to equation (3.15) as 4.8498 and the Consistency Index (CI) according to equation (3.14) as 0.0375. Since the value of CI is close to zero the estimation is valid.

Table 3-7 provides the scale of relative weights of the port groups. The group with the highest weight is the group that its container throughput was growing faster than the other ones. The results are depicted in figure 3-6.

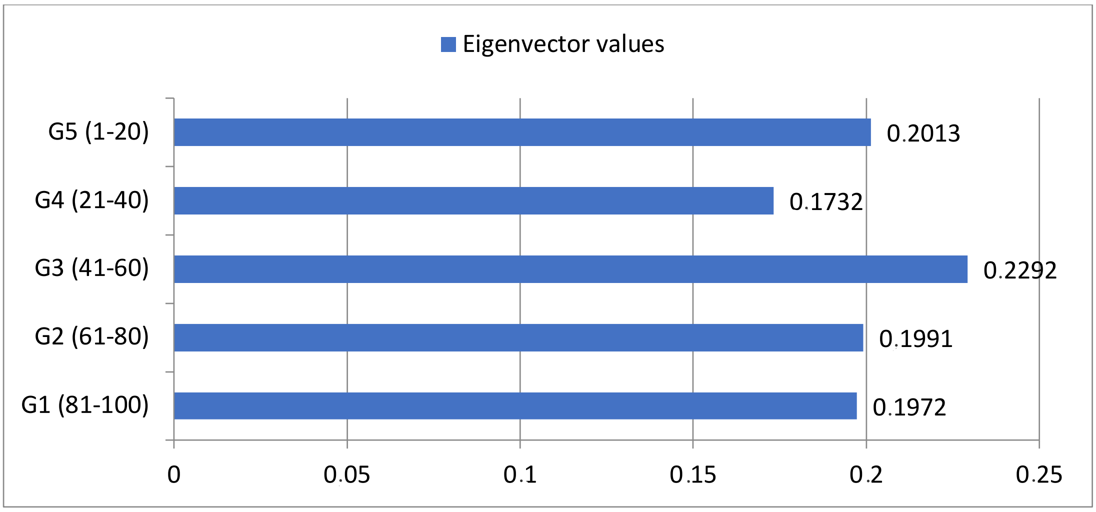

Figure 3-6: Scale of relative weights of port groups according to their growth rates.
Source: Realized by the author.

Analysis of the results

The port group with the highest relative weight according to the growth rates of port groups is G3 (ports ranked from position 41-60). The group containing the biggest ports G5 (ports ranked from position 1-20) is second and the two groups with the smaller ports, G2 and G1, are following. The group that has the lowest relative weight is G2 (ports ranked from position 21-40). These results provide us with more details about the growth potential of ports related with their positions in the hierarchy. Medium-sized ports (G4) experienced the slowest relative growth as a group compared with the groups of large and small ports. This result is in agreement with the result of the quadratic rank-size model and the finding of Oliveira et al. (2021) about a growth trap for medium-sized ports. The biggest ports (G5) experienced larger growth rates compared to the groups of medium-sized and small ports, except for the upper part of small ports (G3). This means that the top of the hierarchy remains unchallenged.

3.7 Results of Markov chain modeling on ports mobility

Table 3-8 provides the results from the Markov chain modelling. Each element of the transition matrix represents the number of ports moving from one group to another every year.

Table 3-8: Transition matrix.

P_{ij}	G1 (81-100)	G2 (61-80)	G3 (41-60)	G4 (21-40)	G5 (1-20)
G1 (81-100)	87.1%	12.4%	0.3%	0.3%	0.0%
G2 (61-80)	12.6%	80.3%	6.8%	0.3%	0.0%
G3 (41-60)	0.3%	7.4%	86.3%	5.8%	0.3%
G4 (21-40)	0.0%	0.0%	6.6%	90.3%	3.2%
G5 (1-20)	0.0%	0.0%	0.0%	3.4%	96.6%

Source: Calculated by the author.

Several remarks can be made based on the features of the transition probabilities. First, the diagonal elements of the matrix which represent the probabilities of a port remaining in the starting group are higher for the groups G4 and G5. This means that immobility is higher for large and medium-sized ports compared to small ports. The most stable group is the group of the larger ports G5 (ports ranked 1-20) where only 3.4% of the ports are moving one group down every year while the group with the highest volatility is G2 (ports ranked 61-80). Second downward mobility is much higher than upward mobility for all port groups. This means that the probability that a port will fall one group down is much higher than to climb one group up.

Table 3-9 provides the expected times needed for a port to reach a group for the first time from his starting group.

Table 3-9: Mean first passage time matrix in years.

m_{ij}	G1 (81-100)	G2 (61-80)	G3 (41-60)	G4 (21-40)	G5 (1-20)
G1 (81-100)	0.0	10.3	35.7	76.9	203.6
G2 (61-80)	30.3	0.0	28.4	70.8	196.5
G3 (41-60)	68.6	39.2	0.0	47.9	171.7
G4 (21-40)	99.5	69.6	29.9	0.0	127.2
G5 (1-20)	128.9	99.1	59.3	29.4	0.0

Source: Calculated by the author.

A port needs on average 10 years to move for the first time from G1 to G2 but 127 years to move from G4 to G5. Table3-9 shows that the transition one group up becomes slower as the size group increases. It is observed a big difference between one group up and one group down first passage times of G4 group. A port needs on average 30 years to move from G4 to G3 and on the other hand 127 years to move from G4 to G5. This finding agrees with the results of the quadratic rank-size model and the allometric growth model presented in previous parts of this study about the growth trap of medium-sized ports.

3.8 Conclusions

This study analyzes the dynamics of port hierarchies focusing on the European container port system. The evolution of container traffic distributional patterns, the relative growth of port groups of different sizes and the mobility of individual ports within the hierarchy are examined by the application of methods different than the commonly used concentration indicators. The database that is used was constructed with the annual container port throughputs of the top-100 European container ports for 20 consecutive years from 2001-2020.

The rank-size model discloses a tendency to deconcentration from 2001-2009 that confirms the finding of Notteboom (2010) for gradual deconcentation until the year 2008 that was the last year included in his study. After the financial crisis of 2009, from 2010-2020, the trend turned to concentration of traffic. It seems that the financial crisis of 2009 was a crucial turning point for the container traffic concentration patterns of the European port system, shifting from deconcentration in the decade prior the crisis to concentration in the decade after the crisis. The quadratic rank-size model reveals the existence of an abundance of medium-sized ports in the

European container port system. The influence of these medium-sized ports in the traffic distribution experienced also different trends. It decreased from 2001-2010 and then increased from 2011-2020. The combined results of the two rank-size models suggest that medium-sized ports are favored from a concentration trend in the port system and on the other hand they lose ground when deconcentration occurs.

The AHP methodology combined with an allometric model sheds light to the relative growth of groups of ports of different sizes one upon another. The results indicate that the port group containing ports ranked from position 21 to 40 in the hierarchy was the group that experienced the lowest growth rate compared to the other port groups from 2001-2020. This finding agrees with that of Oliveira et al. (2021) about a growth trap for medium-sized ports in the European region. The group of the largest ports (positioned from 1 to 20) remained unchallenged since it had the larger growth rate compared to all other port groups except from the upper part of the small ports (ports ranked from 41-60) which was the group with the highest relative growth.

The mobility of ports within the hierarchy of the European container port system is explored with the Markov chain modeling. The results reveal that the group of the largest ports presented the higher degree of immobility followed by the group of medium-sized ports. The mobility was higher at the bottom parts of the distribution and especially at the group containing the ports ranked from position 61-80. One key observation is that downward mobility was much higher than upward mobility for all port groups suggesting that it is more likely for a port to lose ground than to climb in the hierarchy. The calculation of the first passage time from one part of the distribution to another estimates the speed of the mobility within the hierarchy. The results show that the transition one group up becomes much slower as the size group increases. It is remarkable that a medium-sized port needs on average 127 years to enter for the first time the group of the largest ports compared to 30 years to move one group down.

Chapter 4

Comparative analysis of the container liner shipping networks of individual operators

4.1 Introduction

This chapter analyzes the container liner shipping networks of Maersk and COSCO Shipping, ranked 1st and 3rd at the list of the top operators (Alphaliner, 11 January 2021), in the years 2001, 2007 and 2010. The selected years permit to observe the evolution of the companies' networks prior and after the global financial crisis of 2009. Notteboom (2022a) mentions that the port hierarchy is determined by the decisions of individual carriers. Thus, it is essential to analyze the container liner shipping networks of individual container shipping companies.

Firstly, the whole network of the carriers is analyzed through global network indicators and visualization of the networks. Secondly, the complex network properties of the shipping networks are investigated with the fit of the power law distribution to the degree distribution of the networks. Thirdly, the positions of the ports in the companies' networks are explored by calculating their degree and betweenness centrality. Fourthly, the weighted degree of the ports and the strongest inter-port links are presented using the frequency (weekly calls) as weight of the links. Finally, the research focuses on the regional networks of the companies and their evolution.

The remainder of the chapter is structured as follows: Section 4.2 provides information on the network analysis of maritime transportation. Section 4.3 presents a comprehensive review of the existing studies that analyze the topological structure of maritime shipping networks with the use of graph theory. Section 4.4 introduces the methodology of the research. Section 4.5 analyses the whole networks of the selected companies by calculating global network indicators and by visualizing the networks. Section 4.6 investigates the existence of complex network properties. Section 4.7 identifies the positions of ports in the carriers' networks by calculating the degree and betweenness centrality indicators. Section 4.8 includes the weighted degree of the ports of the networks and the inter-port traffic measured by frequency (in weekly calls). Section 4.9 analyzes thoroughly the regional networks of the companies. The final section 4.10 presents the conclusions of the research.

4.2 Network analysis of maritime transportation

The main purpose of transportation is to surpass distance and move passengers, freight and information between locations. It consists of some core components that are fundamental in order to fulfil its purpose. One of these components is the network, which is a system of linked locations (Rodrigue, 2020). So, the analysis of transportation from a network perspective is a natural choice as networks are one of its core components. Maritime transportation networks, nevertheless, had been rarely studied with quantitative methods until the mid-2000's. There are two main reasons for this scarcity of network studies of maritime transport. The first reason is that the sources for the necessary data are limited and the cost to obtain them is high.

There are two ways to obtain the necessary data in order to construct a maritime transportation network. The first is through the daily movements of vessels that are recorded by the Automatic Identification System (AIS) that is installed on vessels. It is a satellite identification system of the position of vessels similar to the GPS. The main provider of AIS data is Lloyds Maritime Intelligence Unit (LMIU). The advantage of this kind of data is that they concern the actual vessel movements and the main disadvantage that the purpose of the vessel movement cannot be identified. Consequently, vessel movements with other purposes than transportation of cargos (e.g. bunkering, anchorage) cannot be excluded. The second way to construct a maritime transportation network is through the scheduled liner shipping services of carriers. The advantages of this kind of data are that only the movements that concern shipping services are included and not irrelevant movements (e.g. bunkering, anchorage) and that the origin and the destination of the shipping routes are clearly defined. The drawback is that the data concern scheduled and not actual movements.

The liner services data can be obtained from three sources: The paper annual editions called Containerization International Yearbooks (last full edition in 2012), online service databases provided by Alphaliner and Containerization International Online and through the websites of shipping companies. The websites of carriers are the only free access source but the data provided are not complete. The Containerization International Yearbooks are an affordable choice with the disadvantages that a lot of manual work is needed and the last edition with the section of liner services was in 2012. The collection of liner services data from Alphaliner or Containerization International Online, or AIS data from LMIU is a costly choice. The second reason that discouraged researchers from carrying out maritime network studies was the lack of network analysis and visualization software.

4.3 Literature review

Networks can be represented and analyzed by applying concepts from the mathematical field of graph theory. Graph theory provides the way to represent networks as systems of nodes and links and measure their characteristics through the calculation of a large number of indicators. From mid 2000`s the studies of maritime networks using graph theory began to appear on a regular basis. The main reasons that led to the emergence of these studies were the creation of software dedicated to network analysis and visualization (e.g. Gephi, Tulip, Ucinet) and the increased availability of the necessary data for the construction of networks. Nevertheless, the high cost of the data remains and important obstacle to the research of maritime networks. Table 4-1 presents a comprehensive list of the existing studies that analyze the topological structure of maritime shipping networks with the use of graph theory.

The geographical coverage of the studies is global or regional. The global maritime network has been analyzed by Deng et al. (2009), Hu and Zhu (2009), Kaluza et al. (2010), Wang C. and Wang J. (2011), Ducruet and Notteboom (2012a, b), Gonzalez et al. (2012), Pais et al. (2012) Freire et al. (2013), Kang and Woo (2017). Regional maritime network studies include the Caribbean (McCalla et al., 2005; Veenstra et al., 2005), the Mediterranean (Cisic et al., 2007), Northeast Asia (Ducruet et al., 2009b; Ducruet et al., 2010b), the Atlantic network (Ducruet et al., 2010c) and Canarian ports (Tovar et al., 2015).

The main research goals of the studies using graph theory are:

- To identify how ports are positioned in maritime networks, global or regional (Cisic et al., 2007; Ducruet et al., 2009b; Ducruet et al., 2010b; Ducruet and Notteboom, 2012a,b; Gonzalez et al., 2012; Pais-Montes et al., 2012; Freire et al., 2013; Tovar et al., 2015; Ducruet, 2013a,b; Kang and Woo, 2017). The most common indicators that are used are the degree (number of connections of a port) in order to assess its connectivity and the betweenness centrality (number of shortest paths between any pair of ports in the network that pass through a specific port) in order to evaluate how centrally is a port placed in the network. Hierarchies of ports are build based on the network characteristics of the ports.
- The correlation between network characteristics of a port and the most commonly used indicator of port performance, the throughput volume (Deng et al., 2009; Ducruet et al., 2011b; Kang and Woo 2017).

- The identification of hub-and-spoke structures and the assessment of the hub status of ports (Veenstra et al., 2005; Low et al., (2009); Ducruet et al., 2010c; Wang C. and Wang J., 2011)
- To analyze maritime transport networks as complex networks (Deng et al., 2009; Hu and Zhu, 2009; Kaluza et al., 2010). Researchers with a background different from transportation, geography or maritime studies tend to analyze the network from a complex network perspective. Their main goal is to identify and measure complex network properties of maritime networks. It is investigated if maritime transport networks can be characterized as scale-free and small world networks.
- The identification of communities and clusters in the maritime networks (Ducruet et al., 2010c; Ducruet and Notteboom, 2012a,b; Ducruet and Zaidi, 2012).

All the above studies aggregate all the vessel movements or all the liner services of carriers. There is a research gap in the analysis of more disaggregated maritime networks. This study analyzes the maritime network at the shipping company level. There are very few studies that investigate the maritime networks of individual liner shipping companies. Fremont and Soppe (2004) studied the maritime networks of the largest liner shipping companies and revealed their different strategies in the organization of their services. Fremont (2007) focused on the maritime container shipping network of Maersk as a global maritime network. These studies of the liner shipping networks of individual companies did not use network indicators based on graph theory. As far as it is in author's knowledge there are no studies that analyze the liner shipping networks of individual carriers using graph theory and complex network properties.

Another call for research that this study addresses is the need for comparative studies. The liner shipping networks of individual carriers are going to be compared in order to understand their individual strategies in organizing shipping networks.

Table 4-1: Presentation of the studies analyzing the topological structure of maritime shipping networks with the use of graph theory.

Author	Source	Main indicators	Objective
McCalla et al. (2005)	CIY	β coefficient Degree Centrality	To study the structure of the Caribbean container shipping network between 1994 and 2002
Veenstra et al. (2005)	LMIU	Degree Centrality	Analyze the container flows in the Caribbean and assess the potential for the development of a transshipment hub
Cisic et al. (2007)	N/S	Degree Centrality Betweenness Centrality Closeness Centrality	Identify the hierarchy of ports in the Mediterranean shipping network

Wang and Cullinane (2008)	SR,AX	Port accessibility Index computed by Principal Eigenvector Method	Determination of the competitiveness of a port through the measurement its accessibility
Deng et al. (2009)	Websites of shipping companies	Degree Centrality Clustering Coefficient Efficiency Complex network properties	Analyze the worldwide maritime network as a complex network, estimate its efficiency and correlate throughput with network properties
Hu and Zhu (2009)	CIO	Degree centrality Betweenness Centrality Shortest path length Node strength Clustering Coefficient Rich-club Coefficient Nearest Neighbors Degree Complex network properties	To investigate the complex network properties of the global container shipping network
Cullinane and Wang (2009)	CIY,AL	Port accessibility index	The formulation of an index of port accessibility
Low et al. (2009)	Websites of shipping companies	Connectivity Index Cooperation Index	Assessment of hub status of major Asian ports within the networks of major carriers
Ducruet et al. (2009b)	LMIU	Degree centrality Betweenness Centrality Hub Centrality Index	Evaluation of the position of South Korean ports within Northeast container shipping network in 1996 and 2006
Ducruet et al. (2010a)	LMIU	Degree centrality Betweenness Centrality Vulnerability	To explore the properties of liner shipping networks and their influence on the evolution of port hierarchies
Ducruet et al. (2010b)	LMIU	Degree centrality Betweenness Centrality Hub dependence Complex network properties Relative Diversity Index	To study how ports are positioned in the Northeast Asian container shipping network in 1996 and 2006
Ducruet et al. (2010c)	LMIU	Degree centrality Betweenness Centrality Complex network properties Hub dependence Bisecting K-means Clustering	To verify to what extent the hub-and-spoke strategies of ports and ocean carriers have modified the structure of a maritime network, based on the Atlantic case
Kaluza et al. (2010)	LMIU	Degree centrality Betweenness Centrality Clustering Coefficient Node strength Complex network properties	To understand the patterns of global trade and bio invasion through a complex network of global cargo ship movements
Ducruet et al. (2011a)	LMIU	Complex network properties	To reveal the complementarities between air and sea transport networks in shaping an urban hierarchy
Ducruet et al. (2011b)	LMIU	Degree centrality Nodal Degree	Correlation between throughput volume and network indicators

		Betweenness Centrality Eccentricity Clustering Coefficient	
Ducruet et al. (2011c)	LMIU	Degree centrality Betweenness Centrality	To explore port competition through a maritime network perspective
Wang C. and Wang J. (2011)	Websites of shipping companies	Degree centrality Hub dependence Nodal degree	To identify hub-and-spoke systems in the global shipping network
Ducruet and Notteboom (2012a)	LMIU	Degree centrality Betweenness Centrality Hub dependence Nodal degree Complex network properties	To analyze the global liner shipping network in 1996 and 2006
Ducruet and Notteboom (2012b)	LMIU	Degree centrality Betweenness Centrality	Description of worldwide liner shipping network in 1996 and 2006
Ducruet and Zaidi (2012)	LMIU	Degree centrality Complex network properties	To identify bridges and communities in maritime networks by applying the topological decomposition method
Gonzalez et al. (2012)	LMIU	Degree centrality Betweenness Centrality Hub dependence Complex network properties	To look at changes in the maritime network prior and after the financial crisis of 2009
Pais et al. (2012)	LMIU	Degree centrality Betweenness Centrality Complex network properties	To explain the evolution of containerized and general cargo networks from 2008 to 2011
Freire et al. (2013)	LMIU	Degree centrality Betweenness Centrality	To show how connectivity evolved for containership and general cargo ports from 2007 to 2011
Ducruet (2013a)	LMIU	Degree centrality Betweenness Centrality Eccentricity Clustering Coefficient Rich-club Coefficient Mean diversity of neighbors Mean diversity of links	Hierarchy of ports based on their connections with multiple commodity flows
Ducruet (2013b)	LMIU	Degree centrality Hub dependence Strength clustering	To understand the relative positions of seaports within the global maritime network by applying a strength clustering method
Tovar et al. (2015)	AL	Degree centrality Betweenness Centrality Hub dependence Port accessibility Index Complex network properties	To evaluate the accessibility and connectivity of the main Canarian ports

Tran and Haasis (2014)	CIY	Degree Centrality Complex network properties	Evolution of the East-West shipping network from 1995 to 2011
Li et al. (2015)	CIO	Degree Centrality Betweenness Centrality Closeness Centrality	To understand the dynamic changing of centrality of 25 geographical areas from 2001 to 2012
Wang and Cullinane (2014)	CIO	Multiple Linkage Analysis Hierarchical Clustering Analysis Hub dependence	To assess the impact of freight traffic consolidation in the container port industry
Xu et al. (2015)	CIO	Degree centrality Betweenness Centrality Hub dependence	To analyze the changing positions of world regions in the global shipping network from 2001 to 2012
Wang and Cullinane (2016)	CIO,SR	Degree Centrality Betweenness Centrality Closeness Centrality	To measure the centrality of ports in container shipping networks
Calatayud et al. (2017a)	CIY	Degree Centrality Betweenness Centrality Complex network properties	To assess a country's degree of connectivity to international markets
Calatayud et al. (2017b)	CIY	Degree Centrality Betweenness Centrality Complex network properties	To understand the degree of vulnerability of international freight flows in case of disruptions in the maritime transportation network
Kang and Woo (2017)	CIO	Degree Centrality Betweenness Centrality Closeness Centrality Eigenvector Centrality Complex network properties	To examine the evolution of shipping networks and the relationship between port network characteristics and cargo throughput
Zhang et al. (2018)	Websites of shipping companies	Degree Centrality Betweenness Centrality Node Strength Line Saturability Modularity Efficiency	To explore the properties of the links of the world maritime transport network
Xu et al. (2020)	AL	Degree Centrality Betweenness Centrality Closeness Centrality Complex network properties Modularity Index Density Rich-club Coefficient	To underpin the complex systems association of global liner shipping network with international trade

CIY: Containerization International Yearbook; **LMIU**: Lloyds Maritime Intelligence Unit; **SR**: Searates; **AX**: Axsmarine; **CIO**: Containerization International Online; **AL**: Alphaliner; **N/S**: Not Specified.

4.4 Research methodology

4.4.1 Objectives, time period and liner shipping companies' selection

The purpose of this study is threefold. Firstly, it aims to construct and analyze the container shipping networks of individual liner shipping companies. Secondly it intends to present the evolution of these networks during a selected time period. The third objective of the study is to compare the liner shipping networks of different individual carriers and their evolution.

The selected years for the construction of the container shipping network of the liner shipping companies are 2001, 2007 and 2010. This choice is made in order to observe how the liner shipping networks evolved prior and after the global financial crisis of 2009. In this way it is possible to explore the evolution of the maritime networks of individual carries in time periods with very different financial conditions. Firstly, between 2001 and the pre-crisis year 2007 where container liner traffic was almost doubled, from 72.3 million TEUs to 142 million TEUs (Drewry, 2007-2009). Secondly between the pre-crisis year 2007 and the after-crisis year 2010 in order to observe how the financial crisis which began in late 2008 affected the evolution of the carriers' networks (Figure 4-1).

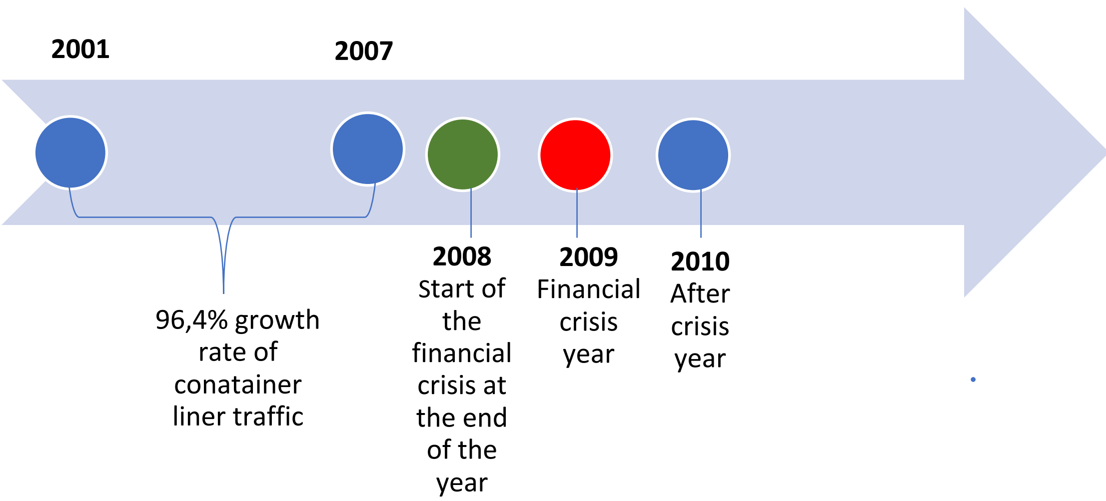

Figure 4- 1: Selected time period of the study.
Source: Visualized by the author.

The selected liner shipping companies are the Danish Maersk and the Chinese COSCO Shipping. The choice of these carriers is made for two reasons. The first one is their size. Maersk is ranked 1st at the list of the top operators (constantly from 1997) with a capacity of

4,133,988 TEUs and 17% market share and COSCO is ranked 3rd with a capacity of 3,043,046 TEUs and 12.5% market share (Alphaliner, 11 January 2021). The second reason behind their choice is that it would be interesting to compare the liner shipping networks of a European and an Asian company, with very different cultural, organizational and operational backgrounds.

4.4.2 Construction of the companies' networks

This study analyzes the liner shipping services of carriers in order to construct their maritime shipping network. A liner shipping service is a maritime route consisted of consecutive calls at selected ports by the carrier. The vessels of the liner shipping companies travel between regions all over the world and call at a number of selected ports at each region. The nature of the liner shipping services is circular and thus the vessels of the carrier travel in repeated loops beginning from a port and returning back to it. The combination of a number of services creates a maritime network. In Fig. 4-2 is presented the AE1 service of Maersk in 2007. It was a service operating between Northeast Asia (5 port calls) and North Europe (6 port calls) with an intermediate stop at both legs at Southeast Asia (2 port calls). The shipping services have been obtained from Containerization International Yearbooks.

The information that can be derived from the analysis of liner services is: the sequence of ports of call, the number of vessels used at the service and their capacity, the frequency of the service, the travelling time from a port of call to the next one.

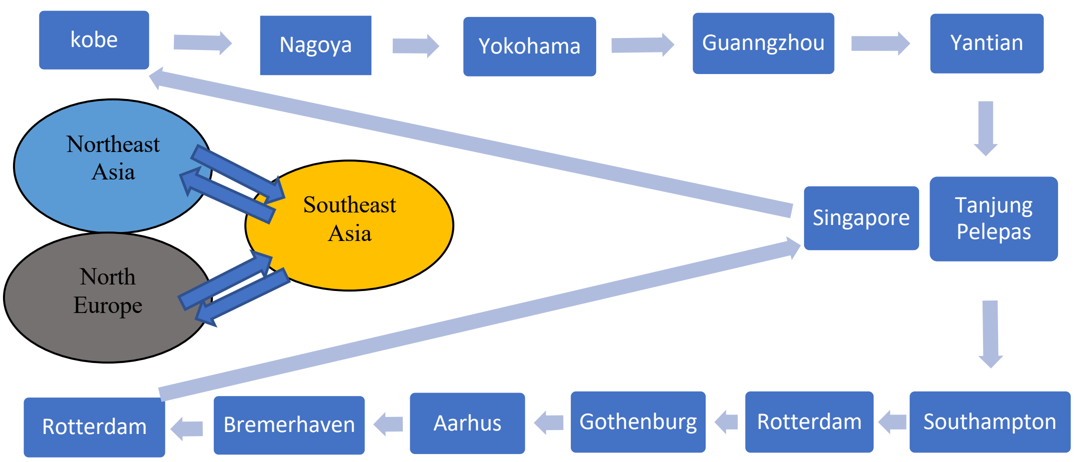

Figure 4- 2: AE1 service of Maersk in 2007.
Source: Visualized by the other based on service information published in CIY (2008).

In order to analyze the network and measure its properties it is represented as a graph. A graph is consisted of nodes and links. The ports of call are the nodes of the graph and a link is created when a vessel travels from one port to another. The basic choices that have to be made in order to construct a graph are:

- The type of graph. There are two types of graphs. The graph of direct links (GDL) (Ducruet et al., 2010c) also defined as L space (Hu and Zhu, 2009) and the graph of all links (GAL) (Ducruet et al., 2010c) also defined as P space (Hu and Zhu, 2009). In the GDL a link is created only between consecutive port calls (when a vessel calls from one port to the next of the liner service). In the GAL links are created not only between consecutive ports but between all the ports that participate in the service (Figure 4-3).
- The direction of the graph. The graph can be either directed or undirected. In a directed graph a call from port A to port B creates a link that is different from the link from port B to port A. In an undirected graph there is no difference if the call is from port A to B or from B to A (Figure 4-4).
- The weight of the links. A graph can be unweighted when all the links between the nodes considered of the same strength or weighted when there is a differentiation of the strength of each link. The weight of a link between two ports can be e.g., the capacity of vessels calling from one port to the other or the frequency with which vessels calling from one port to the other.

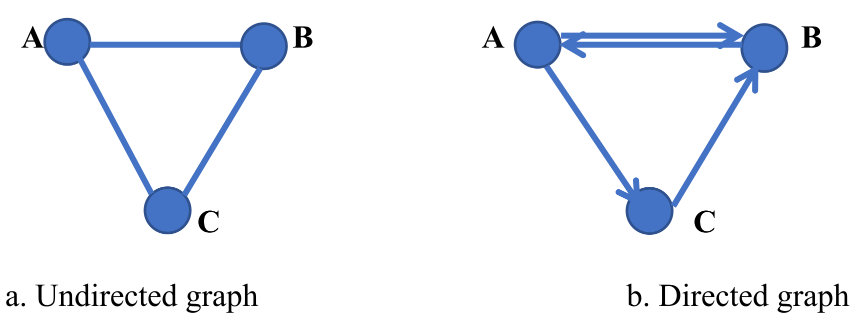

Figure 4-3: Examples of undirected (a) and directed graphs (b).
Source: Visualized by the author.

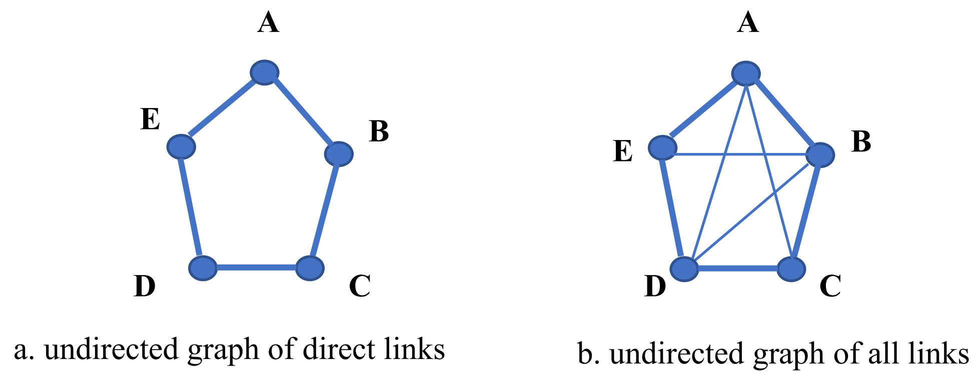

Figure 4-4: Examples of undirected graphs of direct links (a) and all links (b).
Source: Visualized by the author.

The selected graph for this study is a weighted undirected graph of direct links. The inter-port links are weighted by the frequency of the connection between the pair of ports, measured in weekly calls. The weight of the links is calculated according to the following formula:

$$W_I = \sum_{s=1}^{k} \frac{7}{F_s} \times C_s \tag{4.1}$$

Where W_I is the total weight of link I, k the number of services including link I, F_s the calling frequency of the service in days and C_s the number of times that link I is included in each service. According to this calculation if a service was calling e.g., weekly (every 7 days) at the ports, the weight of the inter-port links is 1. If it was calling every 14 days, the weight of the inter-port links is 0.5.

There are two main differences in the methodology of this study compared to others. Firstly, the container liner shipping network is analyzed at the individual carriers' level. As a result, this study does not aggregate the liner shipping services of a number of carriers but uses their liner services separately in order to form their networks. By this way, the strategies of individual carries in organizing liner shipping networks can be revealed. Secondly the inter-ports links are weighted by the frequency of the liner services measured in weekly calls and not by the transport capacity of the vessels in TEUs. This approach permits to understand better the time proximities between ports.

4.5 Analysis of liner shipping companies' networks.

4.5.1 Global indicators of the companies' networks.

Maersk has built a global maritime container shipping network. In 2001 the network of the company consisted of 241 ports and 480 inter-port links. In the pre-crisis period Maersk's network expanded significantly and in 2007 included 255 ports and 617 inter-port links. After the crisis of 2009 there was shrinkage of the company's network and in 2010 it consisted of 237 ports and 545 inter-port links. The average degree of the network and the average weighted degree followed the same course. They were increased prior the crisis and decreased after it. The density of the network increased slightly between 2001 and 2007 and remained stable after the crisis. The average clustering coefficient remained stable throughout the examined period. The average shortest path length of the network decreased in 2007 compared to 2001 and continued this tendency in the after-crisis year 2010.

COSCO had been operating a much smaller maritime network compared to Maersk's. In 2001 it consisted of 110 ports and 258 inter-port links. In the pre-crisis period, it grew significantly and in 2007 it included 136 ports and 352 inter-port links. Contrary to Maersk's, COSCO's network didn't shrink after the crisis but retained its size around the pre-crisis levels. The average degree in COSCO's network is slightly larger than in Maersk's throughout the examined period, meaning that the ports in COSCO's network had on average more connections than those of Maersk's. The average degree was increased prior the crisis and slightly decreased after it.

The average weighted degree was also larger in COSCO's network throughout the examined period, meaning that the ports in the Chinese carrier's network were receiving on average more weekly calls than those of Maersk's. The average weighted degree was growing throughout the examined period and in 2010 the ports in COSCO's network were receiving on average 1 weekly call more than in 2001. COSCO's network was less extended and denser compared to Maersk's. Its density slightly decreased in the pre-crisis period and remained at that figure after it. The average clustering coefficient was slightly lower compared to Maersk's network in 2001 and 2010 and slightly larger in 2007. The average shortest path length in COSCO's network was smaller than in Maersk's prior the crisis, meaning that lesser steps were needed to travel from one port to another, but after the crisis became slightly larger.

Table 4-2: Network indicators.

	N	L	$\overline{\mathrm{Deg}}$	$\overline{\mathrm{W}_{\mathrm{Deg}}}$	D	Den	$\overline{C}$	ASPL	γ
Maersk 2001	241	480	3.983	5.695	13	0.017	0.328	4.282	-1.642
COSCO 2001	110	258	4.691	7.896	10	0.043	0.294	3.79	-1.448
Maersk 2007	255	617	4.839	7.021	9	0.019	0.329	3.813	-1.571
COSCO 2007	136	352	5.176	8.89	9	0.038	0.34	3.305	-1.424

Maersk 2010	237	545	4.599	6.616	8	0.019	0.337	3.713	-1.562
COSCO 2010	135	341	5.052	9.821	12	0.038	0.312	3.807	-1.415

N: Nodes; **L**: Links, $\overline{\mathbf{Deg}}$: Average degree; $\overline{\mathbf{W_{Deg}}}$: Average weighted degree; **D**: Diameter, **Den**: Density; $\bar{\boldsymbol{C}}$**:** Average clustering coefficient; **ASPL**: Average Shortest Path Length; **γ**: Power law coefficient. Source: Retrieved by the author based on the input data.

Table 4-3: Definition of network indicators.

Indicator	Formula	Definition
Average degree $\overline{Deg}$	$\overline{Deg} = \frac{2l}{n}$	The average number of links connected to a node. l: Number of links, n: Number of nodes.
Average weighted degree $\overline{W_{Deg}}$	$\overline{W_{Deg}} = \frac{1}{n}\sum_{i=1}^{n} W_{Deg}(i)$ $W_{Deg}(i) = \sum_{j} w_{ij}\ , i \neq j$	The average of the weighted degree of all the nodes. n: Number of nodes. Weighted degree $W_{Deg}(i)$: The sum of weights of links connected to node i. w_{ij}: The weight of the link between node i and node j.
Diameter D	$D = max\ d_{ij}$	The longest of all the shortest paths in the network. d_{ij}: The number of links of the shortest path between node i and node j.
Density	$Density = \frac{2l}{n(n-1)}$	The ratio of the number of actual links to the number of possible links of the network. l: Number of links, n: Number of nodes.
Average clustering coefficient $\bar{C}$	$\bar{C} = \frac{1}{n}\sum_{i=1}^{n} C_i$ $C_i = \frac{2L_i}{k_i(k_i-1)}$	The average of the clustering coefficients of all the nodes. n: Number of nodes. Clustering coefficient C_i: The ratio of the number of actual links between the neighbors of node i to the maximum possible number of such links between them. L_i: The number of actual links between the neighbors of node i, k_i: The number of nodes connected to node i (neighbors).
Average shortest path length ASPL	$\text{ASPL}=\frac{1}{\text{n(n-1)}}\sum_{i\neq j} d_{ij}$	The average number of links needed to get from one node to another through the shortest path. n: Number of nodes, d_{ij}: The number of

		links of the shortest path between node i and node j.
Power law coefficient γ	$P(x \geq k) = c\chi^{-\gamma}$	The power coefficient of the complementary cumulative degree distribution function. $P(x \geq k)$: The probability that a node has degree of k or larger, c: positive scale factor.

Source: Elaborated by the author.

4.5.2 Visualization of the networks.

The visualization of the companies' networks can reveal further characteristics of the structure of the networks. The visualization was based on a force directed algorithm that brings together densely connected nodes and optimizes the detection of communities in the network. A central observation is that the port communities in the networks of both companies consisted mainly of ports that are in geographical proximity. It is a structure that emerges from graphs of direct links. The observation of the visualized networks allows the detection of changes at the clustering of ports throughout the examined period.

In Maersk's network in 2001, the South East Asian ports (Tanjung Pelepas, Singapore, Port Klang) formed an individual community. In 2007 and 2010 these ports were more closely connected with the North East Asian ports and shaped an integrated community. Another change that was detected concerned the Mediterranean cluster. In 2001 West and East Mediterranean ports formed a cluster with the exception of Algeciras that was more closely connected to the West African ports. In 2007 Port Said, Gioia Tauro and the East Mediterranean ports created a new community while the West Mediterranean ports were attached at the cluster of Algeciras.

In 2010, Port Said and the East Mediterranean ports developed more dense connections with Middle East and South Asian ports and were embedded at their community while the West Mediterranean ports were detached from Algeciras and joined with Gioia Tauro. The ports of the East coast of North America and the ports of Central America and the Caribbean Sea were involved at another change of the communities' structure of the network. In 2001 the East coast of North America ports formed a community together with the Central America and the Caribbean ports. In 2007 the East coast of North America ports were detached from the community of Central America and the Caribbean ports and joined with the transshipment hub of Balboa to a new community. In 2010 Balboa was separated from the East coast of North America cluster and was closely connected with Puerto Manzanillo in a new cluster.

In COSCO's network in 2001 Shanghai and Busan were the main ports of a North East Asian community while Singapore, Hong Kong and Port Said the main ports of another cluster. In 2007 the cluster of Shanghai was including the Japanese ports and Singapore was still at the same community with Hong Kong. In 2010 these clusters changed as Shanghai was more closely connected with Hong Kong and formed a new community while the Japanese ports were separated in an individual community. Singapore shaped a new community together with North European, Middle East and East coast of South America ports. Another observation deriving from the visualized communities is the close connection of East coast of North America ports with the Mediterranean ports in 2001 and 2007. In 2010 the East coast of North America ports formed an individual cluster.

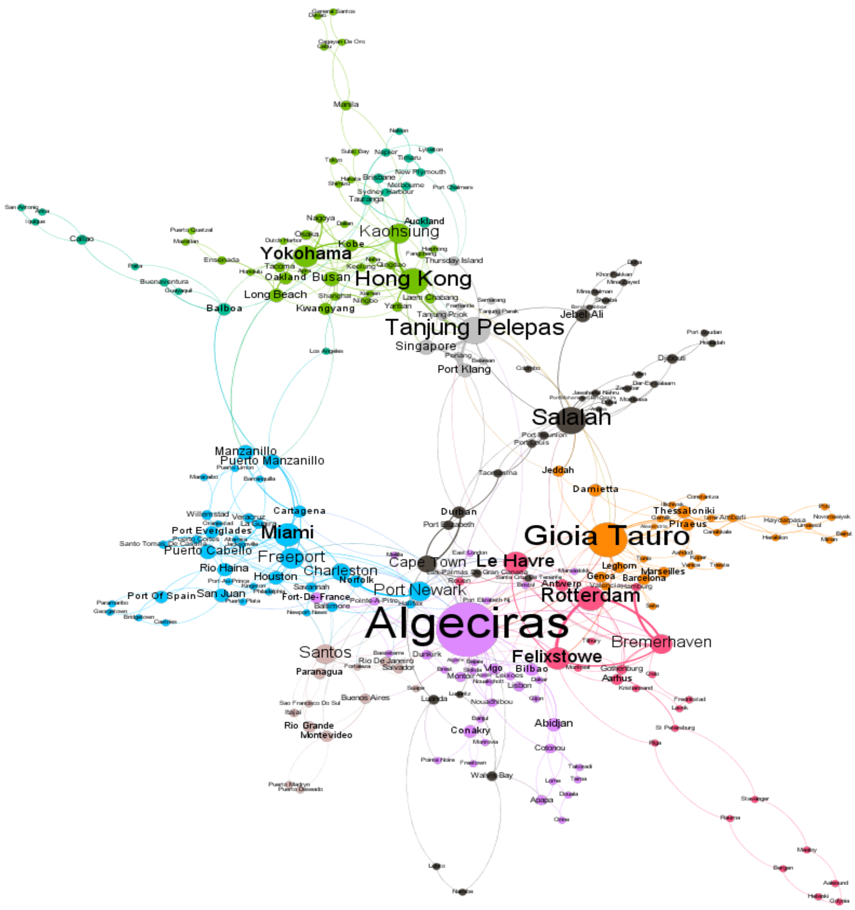

Figure 4-5: Visualization of the liner shipping network of Maersk in 2001.
Source: Visualized by the author based on CIY data, Gephi software and a force directed algorithm that optimizes the detection of communities in the network. Node's size according to degree. Node's color according to modularity (communities' detection). Link's thickness according to frequency (weekly calls).

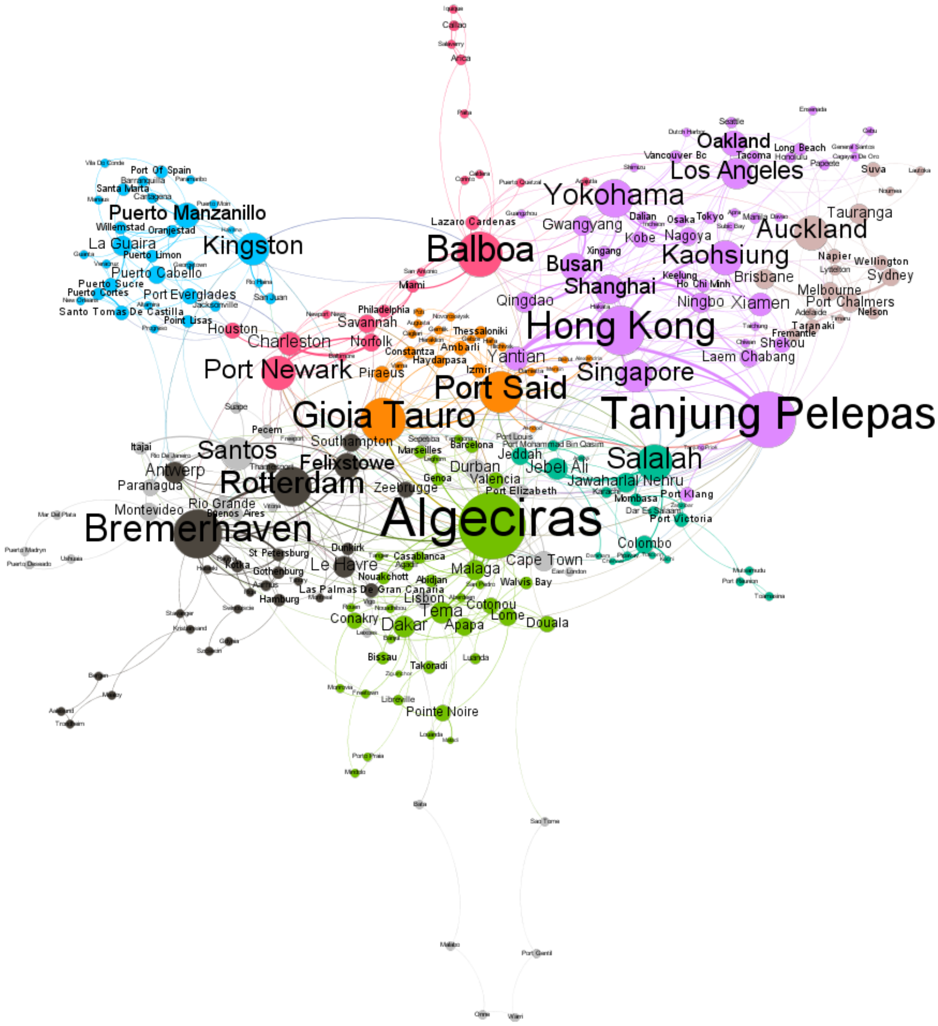

Figure 4-6: Visualization of the liner shipping network of Maersk in 2007.
Source: Visualized by the author based on CIY data, Gephi software and a force directed algorithm that optimizes the detection of communities in the network. Node's size according to degree. Node's color according to modularity (communities' detection). Link's thickness according to frequency (weekly calls).

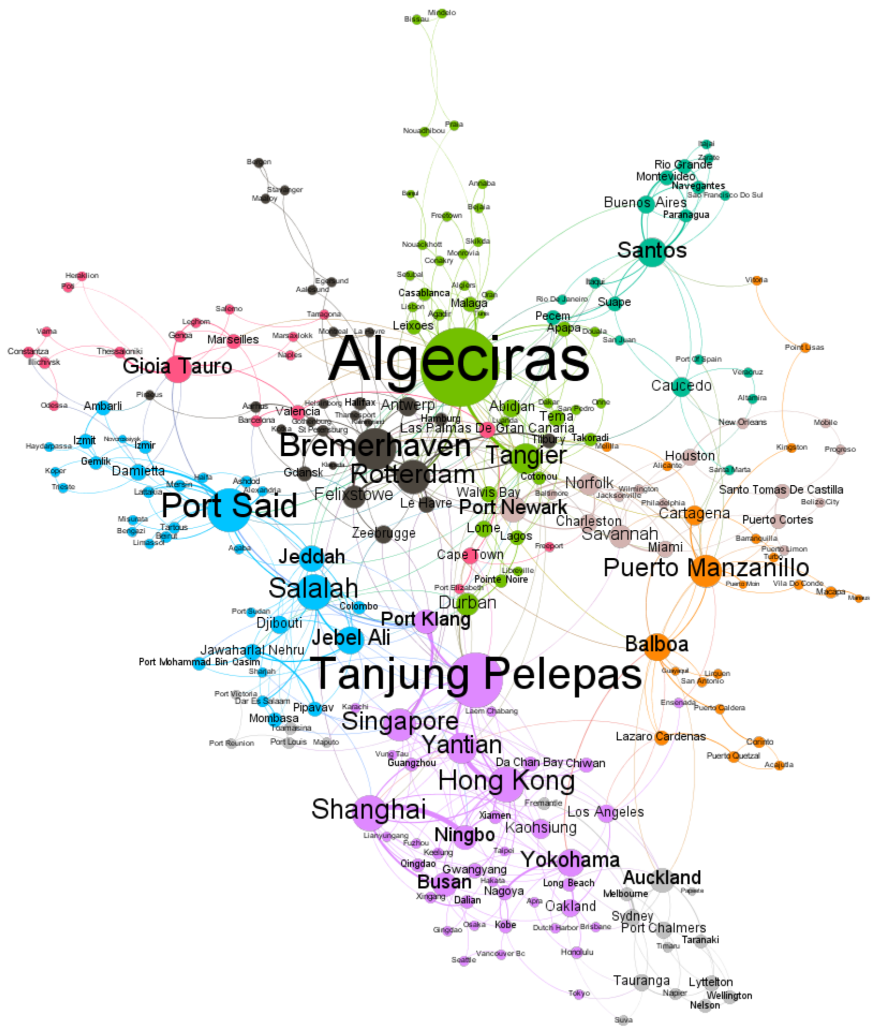

Figure 4-7: Visualization of the liner shipping network of Maersk in 2010.
Source: Visualized by the author based on CIY data, Gephi software and a force directed algorithm that optimizes the detection of communities in the network. Node's size according to degree. Node's color according to modularity (communities' detection). Link's thickness according to frequency (weekly calls).

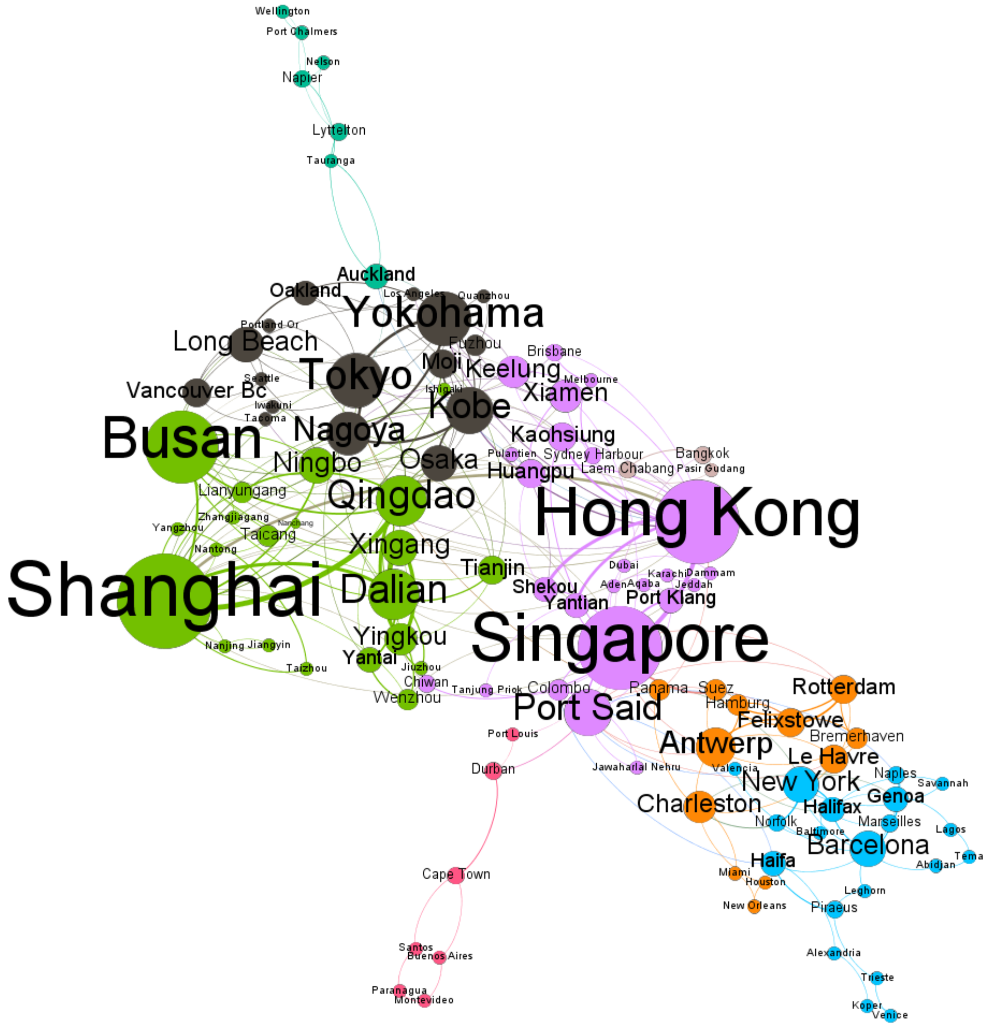

Figure 4-8: Visualization of the liner shipping network of COSCO in 2001.

Source: Visualized by the author based on CIY data, Gephi software and a force directed algorithm that optimizes the detection of communities in the network. Node's size according to degree. Node's color according to modularity (communities' detection). Link's thickness according to frequency (weekly calls).

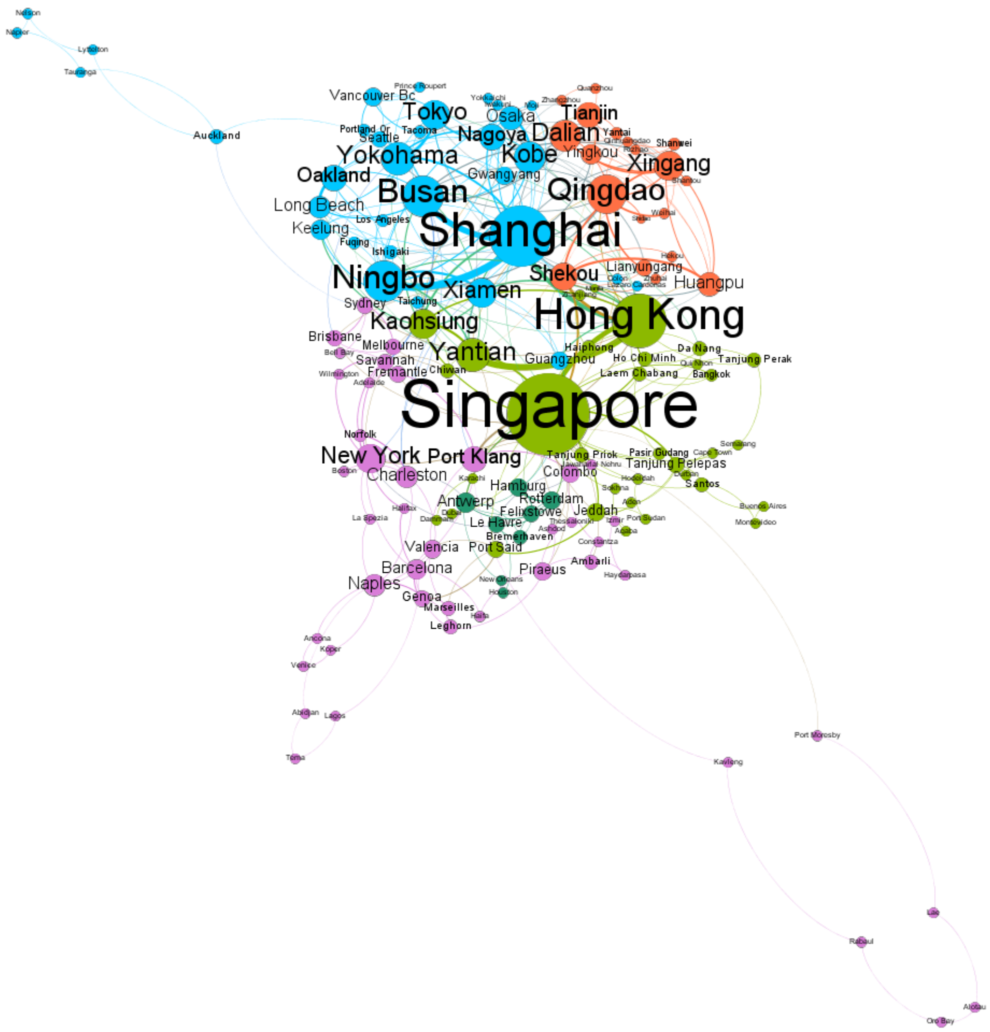

Figure 4-9: Visualization of the liner shipping network of COSCO **in 2007.**
Source: Visualized by the author based on CIY data, Gephi software and a force directed algorithm that optimizes the detection of communities in the network. Node's size according to degree. Node's color according to modularity (communities' detection). Link's thickness according to frequency (weekly calls).

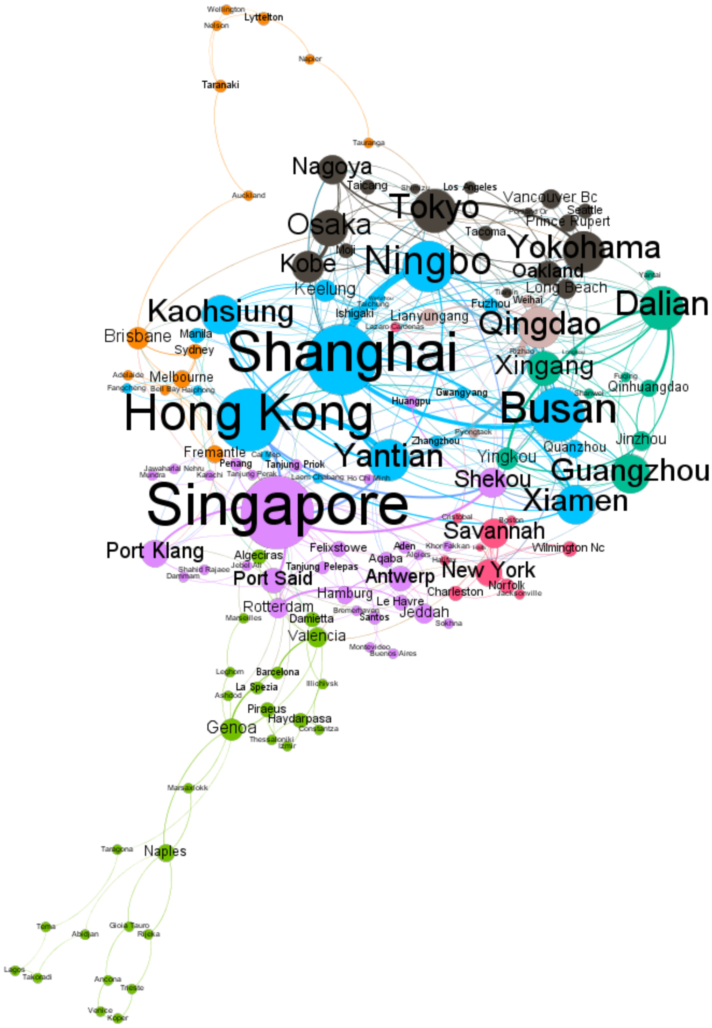

Figure 4-10: Visualization of the liner shipping network of COSCO in 2010
Source: Visualized by the author based on CIY data, Gephi software and a force directed algorithm that optimizes the detection of communities in the network. Node's size according to degree. Node's color according to modularity (communities' detection). Link's thickness according to frequency (weekly calls).

4.5.3 Complex properties of the companies' networks.

The maritime shipping network has been studied as a complex network in global or regional level. In all these studies the shipping network was constructed with aggregated data of vessel movements or shipping companies' services. This study explores the complex network properties of the shipping networks of individual carriers. In order to reveal the connection patterns of the nodes of the complex network some of its characteristics have to be investigated. One of these is the function of the degree distribution. When the degree distribution function of a network follows a power law the network is characterized as a scale free network. The power law degree distribution function is given from the following equation:

$$P(x = k) = c\chi^{-\gamma} \quad (4.2)$$

Where $P(x = k)$ is the probability of a node having degree of k, γ the power coefficient and c a positive scale factor. The term scale-free suggests that, as a result of the power function, the underlying structure of the network remains the same in any scale. The power law degree distribution implies the existence of few nodes with high degree and many nodes with small degree.

In Maersk's network in 2001, 76.8% of the nodes had 4 or less connections and 90.9% had 7 or less connections. The most frequent number of connections was 2 with 50.6% of the nodes having 2 connections. In 2007, 68.2% of the nodes had 4 or less connections and 90.2% 9 or less connections. The most frequent number of connections was again 2 with 38% of the nodes having 2 connections. In 2010 78.5% of the nodes had 5 or less connections and 89% 9 or less connections. As at the previous years the most frequent number of connections was 2 with 44.3% of the nodes having 2 connections.

In COSCO's network in 2001, 80% of the nodes had 6 or less connections and 90% had 9 or less connections. The most frequent number of connections was 2, as in Maersk's network, with 43.6% of the nodes having 2 connections. In 2007, 78.7% of the nodes had 6 or less connections and 85.3% had 9 or less connections. The most frequent number of connections was again 2 with 44.1% of the nodes having 2 connections. In 2010, 79.3% of the nodes had 6 or less connections and 85.9% had 9 or less connections. As at the previous years the most frequent number of connections was 2 with 43% of the nodes having 2 connections.

Regression analysis results indicate a very good fit of the power law distribution to the degree distribution of both companies' networks. The determination coefficients R^2 are very high (0.95 in 2001, 0.91 in 2007 and 0.94 in 2010 for Maersk and 0.90 in 2001, 0.93 in 2007 and 0.89 in 2010 for COSCO). The power law coefficient experienced a small decrease between 2001 and 2007 and remained stable between 2007 and 2010 in Maersk's network. In COSCO's network it experienced a very small decrease in both periods.

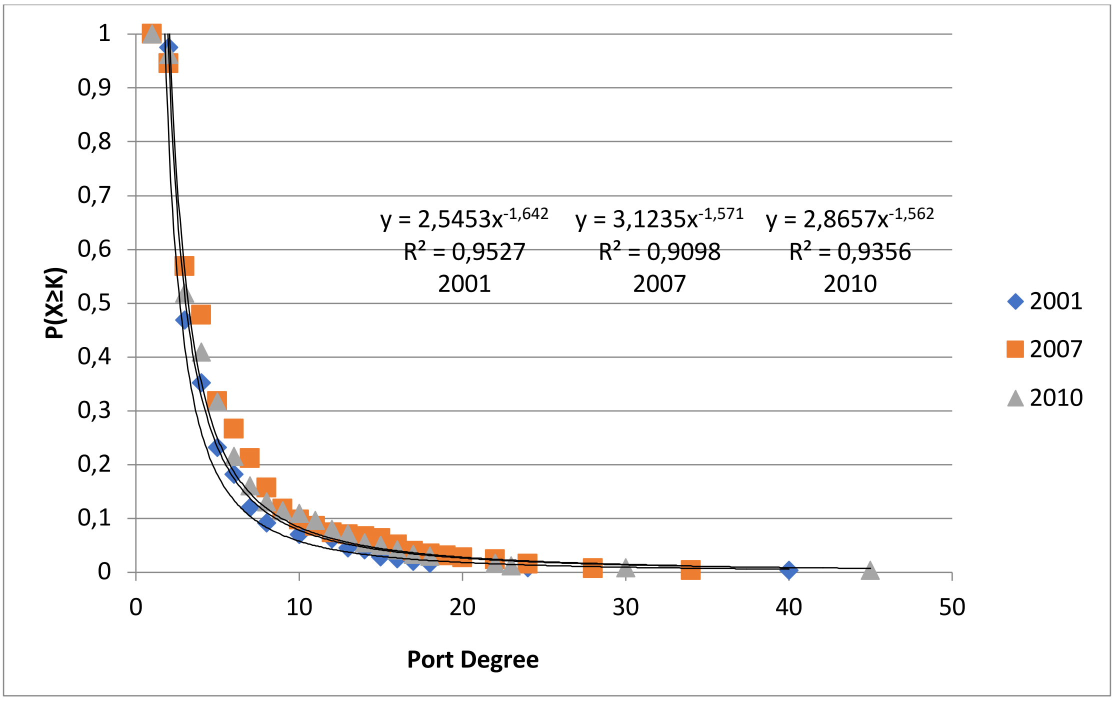

Figure 4-11: Complementary cumulative distribution functions of Maersk's port degree. Source: Visualized by the author based on input data.

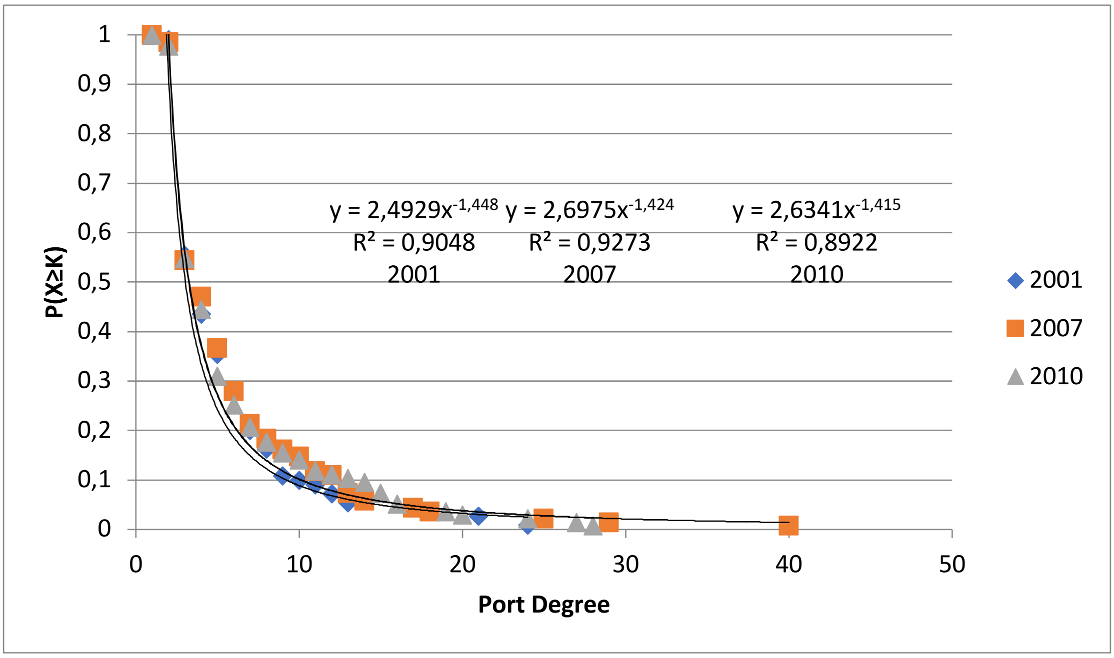

Figure 4-12: Complementary cumulative distribution functions of COSCO's port degree. Source: Visualized by the author based on input data.

4.6 Connectivity, centrality, frequency of calls and time proximities between ports in carriers' networks

4.6.1 Degree centrality

Degree centrality is the most common indicator that is calculated in network analysis. It is a measure of connectivity and it is expressed as the number of direct links that every node of the network has with the other nodes. Degree centrality for an undirected graph can be calculated from a symmetrical adjacency matrix A according to the following formula:

$$Deg_i = \sum_j \frac{a_{ij} + a_{ji}}{2} \; , i \neq j \tag{4.3}$$

Where Deg_i is the degree centrality of port i and the element a_{ij} of the adjacency matrix A is 1 if there is a link between port i and port j and 0 otherwise.

In Maersk's network the most connected port throughout the period from 2001 to 2010 was the Mediterranean transshipment hub Algeciras. In 2001 there was a big difference between Algeciras and the second port in the degree ranking. That gap was significantly decreased in 2007 and then, after the crisis, in 2010 returned at the level it was in 2001. The second port in the degree ranking of Maersk's network in 2001 was another Mediterranean transshipment hub, Gioia Tauro. The connectivity of the port in the company's network was constantly decreasing from 2001 to 2010. It lost 2 connections between 2001 and 2007 and was heavily affected from the crisis as in 2010 lost 9 more compared to 2007 and dropped at the 14th- 17th place at the company's degree ranking. It is a clear case that confirms that pure transshipment hubs are exposed to the strategic decisions of carriers.

Tanjung Pelepas was the 3rd most connected port in the company's network in 2001 and the 2nd in 2007 and 2010. As Algeciras in Mediterranean, it is the transshipment hub of Maersk in South East Asia. It experienced an important increase from 18 connections in 2001 to 28 in 2007 and remained at this level after the crisis with 30 connections in 2010. In North East Asia, Hong Kong was the port with the highest degree throughout the examined period. It had a stable place in the degree rankings of the company, with an increase from 5th to 3rd- 4th place between 2001 and 2007 and back to 5th-6th in 2010. Bremerhaven was the most connected port in the company's network in North Europe. It doubled its connections between 2001 and 2007 and retained its degree in 2010 around the same level.

Rotterdam was the second North European port with a constant presence at the top 10 of the degree rankings of the company. Its degree was stable without significant fluctuations prior and after the crisis. Another port with a stable degree in the network was the Middle East port of Salalah. Its degree remained almost the same throughout the period 2001-2010. Port Said, placed at the mouth of the Suez Canal at the Mediterranean, emerged at the network of the company after the opening of APM's terminal at East Port Said in 2004. In 2007 ranked number 7 in degree and climbed to number 3 in 2010. Balboa, another port situated strategically, at the Panama Canal increased its degree almost four times prior the crisis between 2001 and 2007. It was significantly affected from the company's choices after the crisis and lost almost half of its connections in 2010.

Table 4-4: Top ports degree hierarchy in Maersk's network.

2001		2007		2010	
Port	**Degree**	**Port**	**Degree**	**Port**	**Degree**
Algeciras	40	Algeciras	34	Algeciras	45
Gioia Tauro	24	Tanjung Pelepas	28	Tanjung Pelepas	30
Tanjung Pelepas	18	Bremerhaven	24	Port Said	23
Salalah	18	Hong Kong	24	Bremerhaven	22
Hong Kong	17	Balboa	22	Hong Kong	18
Rotterdam	16	Gioia Tauro	22	Salalah	18
Miami	15	Port Said	20	Shanghai	18
Felixstowe	14	Rotterdam	19	Rotterdam	17
Yokohama	14	Yokohama	18	Puerto Manzanillo	16
Le Havre	14	Salalah	17	Singapore	16
Freeport	13	Auckland	16	Tangier	15
Bremerhaven	12	Kaohsiung	16	Yantian	15
Kaohsiung	12	Port Newark	16	Santos	14
Port Newark	12	Kingston	15	Balboa	13
Santos	12	Santos	15	Gioia Tauro	13
Cape Town	10	Singapore	15	Jebel Ali	13
Charleston	10	Los Angeles	14	Yokohama	13
Singapore	8	Shanghai	13	Jeddah	12
Manzanillo	8	Busan	12	Port Newark	12
Puerto Manzanillo	8	Felixstowe	11	Auckland	11
Puerto Cabello	8	Oakland	11	Busan	11
Busan	8	Puerto Manzanillo	11	Ningbo	11
				Port Klang	11

Source: Retrieved by the author based on the input data.

In COSCO's network all the ports with the highest degree were coming from the region of North East Asia with the exception of Singapore from the South East Asia region. The top 3 ports at the degree rankings of the company throughout the examined period were Singapore, Shanghai and Hong Kong. Singapore experienced a very significant increase in its connections in the pre-crisis period and climbed from 2nd place in 2001 to 1st in 2007. Its degree was doubled, from 21 in 2001 to 40 in 2007. After the crisis, in 2010, Singapore remained the most connected port in the company's network but its degree decreased to 28.

Shanghai was at the 1st place of the degree ranking of the company in 2001 connected with 24 other ports. Experienced a small increase of its connections in 2007 and dropped to 2nd place behind Singapore and remained at the same place after the crisis without significant change of its connectivity. Hong Kong was equal with Singapore at the 2nd place of the degree ranking in 2001 and ranked 3rd in 2007 with a small increase of its degree from 21 to 25. After the crisis remained at the same position and the same level of connectivity. The South Korean port of Busan had a very stable course as far as its degree was concerned and was preserved at the 4th place of the company's ranking throughout the examined period with the same level of degree.

Two ports that were significantly upgraded in terms of connectivity in COSCO's network were Ningbo and Yantian. Prior the crisis Ningbo went from 8 connections in 2001 to 18 in 2007 and Yantian from 5 to 14. After the crisis they remained around the same levels of connectivity. The Japanese ports of Tokyo, Yokohama and Kobe were between the most connected ports throughout the examined period without large fluctuations. The Chinese ports of Qingdao and Dalian were constantly appearing at the top 10 of the degree ranking of the company throughout the examined period. Dalian experienced a small increase of its degree prior and after the crisis and Qingdao an increase prior the crisis and a decrease after it.

Table 4-5: Top ports degree hierarchy in COSCO's network.

2001		2007		2010	
Port	**Degree**	**Port**	**Degree**	**Port**	**Degree**
Shanghai	24	Singapore	40	Singapore	28
Hong Kong	21	Shanghai	29	Shanghai	27
Singapore	21	Hong Kong	25	Hong Kong	24
Busan	18	Busan	18	Busan	20
Tokyo	13	Ningbo	18	Ningbo	19
Yokohama	13	Qingdao	17	Dalian	16
Dalian	12	Yantian	14	Tokyo	16
Qingdao	12	Yokohama	14	Qingdao	15
Kobe	11	Dalian	13	Yantian	15

Port Said	11	Kobe	13	Yokohama	15
Nagoya	10	Kaohsiung	12	Guangzhou	14
Antwerp	9	New York	12	Kaohsiung	14
Barcelona	8	Tokyo	12	Xiamen	14
Long Beach	8	Xiamen	12	Osaka	13
New York	8	Xingang	12	Xingang	12
Ningbo	8	Shekou	11	Kobe	11
Osaka	8	Nagoya	10	Nagoya	10
Xingang	8	Oakland	10	Savannah	10
Charleston	7	Port Klang	10	Shekou	10
Keelung	7	Tianjin	10	New york	9
Xiamen	7			Port klang	9
Yingkou	7				

Source: Retrieved by the author based on the input data.

4.6.2 Betweenness centrality

Betweenness centrality is another common measure that is calculated when a network is analyzed. It expresses how centrally placed is a node of the network by calculating the frequency with which the node appears at the shortest paths of all the node pairs of the network. The betweenness centrality of a port i $C_B(i)$ can be calculated according to the following formula:

$$C_B(i) = \sum_{r \neq i \neq s \in N} \frac{\sigma_{rs}(i)}{\sigma_{rs}} \qquad (4.4)$$

Where σ_{rs} is the total number of shortest paths from node r to node s and $\sigma_{rs}(i)$ the total number of these paths that pass-through node i, N is the total number of nodes.

As with degree the port that was leading the betweenness centrality rankings of the company throughout the examined period was Algeciras. It was by far the most centrally placed port in the network with the difference from the second port being doubled after the crisis in 2010. Tanjung Pelepas was the second port in the company's betweenness centrality ranking of 2001. Its course was downward and although it kept being highly centralized, dropped to 3rd place in 2007 and 4th in 2010. Gioia Tauro's dynamics of betweenness centrality followed the same tendency as its degree. It was in third place in 2001, dropped to 7th in 2007 and to 14th in 2010.

In North East Asia, Hong Kong did not maintain a stable course as with its degree and its betweenness centrality dropped significantly in the after-crisis year 2010. In North Europe Bremerhaven as with its connectivity had a big increase in its betweenness centrality prior the crisis and ranked 4th at the company's network in 2007 and 3rd in 2010. Rotterdam's betweenness centrality was stable as it was its degree around the 10th place of the company's rankings throughout the examined period.

Balboa, the transshipment hub at the Panama Canal, was centrally placed in the company's network throughout the examined period. Contrary to its degree its betweenness centrality was already high in 2001 and reached the 2nd place in 2007. It was affected by the crisis of 2009 but in contrast to its degree remained at the top 5 positions. At the Suez Canal, the port of Port Said was very centrally placed at the network of the company since 2004 when the terminal of APM started its operations. It ranked 5th in 2007 and 2nd in 2010 which was higher than its degree rankings in the same years. Puerto manzanillo another transshipment hub at the Pacific Ocean was at central position in the network especially after the crisis when it ranked 6th in 2010.

Table 4-6: Top ports betweenness centrality hierarchy in Maersk's network.

2001		2007		2010	
Port	**BC**	**Port**	**BC**	**Port**	**BC**
Algeciras	10977.95	Algeciras	8918.529	Algeciras	10468.75
Tanjung Pelepas	6778.323	Balboa	6386.102	Port Said	5179.628
Gioia tauro	6427.128	Tanjung Pelepas	5044.38	Bremerhaven	4409.743
Hong kong	4274.425	Bremerhaven	4276.565	Tanjung Pelepas	4119.692
Salalah	4029.568	Port Said	3400.911	Balboa	3136.315
Kaohsiung	2809.182	Hong Kong	3190.496	Puerto Manzanillo	2925.555
Felixstowe	2766.867	Gioia Tauro	2760.798	Tangier	2205.587
Puerto Manzanillo	2328.313	Kingston	2466.302	Shanghai	2130.502
Balboa	2090.717	Rotterdam	2455.471	Auckland	1974.25
Rotterdam	2064.088	Auckland	2068.267	Salalah	1846.522
Le Havre	2000.081	Port Newark	1934.024	Rotterdam	1769.407
Cape Town	1977.155	Santos	1854.124	Singapore	1501.562
Yokohama	1731.692	Puerto Manzanillo	1803.479	Santos	1383.956
Santos	1727.103	Kaohsiung	1780.305	Gioia Tauro	1229.503
Freeport	1655.618	Salalah	1451.998	Yantian	1143.384
Miami	1573.604	Houston	1244.629	Yokohama	1089.818

Charleston	1567.756	Yokohama	1177.124	Cartagena	993.3677
Jebel Ali	1415.5	Antwerp	1104.71	Buenos Aires	903.3145
Cartagena	1363.158	Dakar	1081.807	Port Chalmers	885.6893
Long Beach	1159.191	Montevideo	1026.771	Hong Kong	865.5395

BC: Betweenness Centrality. Source: Retrieved by the author based on the input data.

In COSCO's network, contrary to its degree rankings where all the top ranked ports were coming from North East Asia and South East Asia, the betweenness centrality rankings include ports from different regions all over the world. The top three ports in the rankings throughout the examined period were, as happened in the degree rankings, Singapore, Shanghai and Hong Kong with the exception of 2010 where Hong Kong dropped at the 8rd place. Singapore had been ranking 1st throughout the examined period with a huge difference from the 2nd port. It experienced an increase of 71.7% between 2001 and 2007 and a decrease of 23.4% between 2007 and 2010. Shanghai was the 3rd most centrally placed port in the company's network in 2001. It climbed at the 2nd place in 2007 and remained there after the crisis in 2010 with a small decrease of its betweenness centrality. Hong Kong was ranked 2nd in 2001, moved to 3rd place in 2007 and in 2010 lost 25.5% of its betweenness centrality figure and dropped to 8th place.

Some ports from regions all over the world appeared at the top places of the betweenness centrality rankings without consistency throughout the examined period. In most cases although the betweenness centrality of these ports was high their degree was low. In 2010 e.g., the port of Genoa was ranked 3rd as far as its betweenness centrality was concerned and 25th-28th in the degree ranking. This observation is explained from the fact that occasionally ports with low degree were included in a number of services which operated between different regions and continents. Genoa in 2010 was included only in 4 services of the company but these services were operating between regions in different continents (NEA-MED, MED-WAF, ECNA-MED). As a result, Genoa had been connected with three continents (Asia, Africa, and North America) and its role as an intermediary place had been enhanced.

Table 4-7: Top ports betweenness centrality hierarchy in COSCO's network.

2001		2007		2010	
Port	**BC**	**Port**	**BC**	**Port**	**BC**
Singapore	2526.865	Singapore	4339.372	Singapore	3322.6
Hong Kong	1461.118	Shanghai	1522.029	Shanghai	1441.006
Shanghai	1058.585	Hong Kong	1333.303	Genoa	1427.042
Port Said	790.7196	Port Klang	1076.023	Busan	1326.414
Panama	760.8545	New York	834.8326	Valencia	1107.438
Auckland	620.3514	Qingdao	690.0769	Port Said	1097.046
Tokyo	583.461	Naples	679.5453	Naples	1033.194

Haifa	573.5836	Yokohama	640.5948	Hong kong	993.4069
Charleston	541.4199	Auckland	529.0081	Savannah	925.0619
Durban	520	Busan	502.152	Brisbane	584.9244
Antwerp	481.4144	Shekou	454.8871	Tokyo	574.9025
New York	468.5053	Yantian	436.148	New York	522.0833
Cape Town	421	Charleston	433.5895	Yantian	515.0089
Barcelona	346.5508	Ningbo	418.3925	Dalian	480.1896
Busan	336.7565	Tanjung Priok	375.6439	Kaohsiung	479.5465
Port klang	314.7129	Kaohsiung	361.2372	Damietta	450.3222
Dalian	292.6832	Barcelona	278.7055	Port klang	401.6344
Colombo	263.3834	Kavleng	278.3561	Auckland	382.1901
Lyttelton	261.5	Santos	266	Guangzhou	379.5333
Qingdao	239.7876	Xiamen	261.463	Ningbo	357.5868

BC: Betweenness Centrality. Source: Retrieved by the author based on the input data.

4.6.3 Weighted degree (number of weekly calls)

Almost all the existing studies that analyze maritime networks with the use of graph theory use as a weight of the inter-port links the transport capacity of the vessels (total, weekly or annualized) travelling between each pair of ports in TEU's. The novelty of this study is that the inter-port traffic is measured by frequency. As far as it is in author's knowledge there are no other studies that use frequency as a weight of the inter-port links. With this approach time proximities between ports can be better understood. The weight of each link between a pair of ports equals with the total number of weekly calls from all the calling services. Since the container shipping services are circular, the weighted in-degree of a port is similar to the weighted-out degree and equal to total weekly calls at the port.

Algeciras was the most frequently visited port in Maersk's network. In 2001 received almost the same number of weekly calls as Tanjung Pelepas and in 2007 and 2010 was in the first place. In 2001 received 25.5 weekly calls, in 2007 28.7 and in 2010 31.2. Tanjung Pelepas was the second most visited port during the examined period. In 2001 received 26 weekly calls and was in the first place of company's ranking. In 2007 and 2010 received 24.5 and 24 weekly calls respectively. Hong kong was constantly in the top 3 most visited ports during the examined period. It received almost 21 weekly calls in 2001, 26.5 in 2007 and 23 in 2010. Bremerhaven was the most frequently visited port in NEUR and ranked constantly at the fourth place of the company's most frequently visited ports.

Rotterdam was the second most frequently visited port in NEUR by Maersk. It received around 17 weekly calls in 2001 and 2007 and 21 after the crisis, in 2010. Port Said was the East Mediterranean port that the vessels of Maersk called most frequently prior the crisis, in 2007 as after it, in 2010. Another transshipment hub in Mediterranean, Gioia Tauro, received 17 weekly calls from the company in 2001 and 15 in 2007 but was seriously affected from the company's port choices after the crisis and in 2010 its weekly calls dropped to 9.5. Two ports that experienced an important increase of the calling frequency of the company were the Chinese ports of Shanghai and Yantian. In 2001 they were rarely visited by Maersk but in 2007 Shanghai received 15 weekly calls and Yantian 17. In 2010 Shanghai reached at 20 weekly calls and Yantian at 18.

Table 4-8: Most frequently visited ports in Maersk's network.

2001		2007		2010	
Ports	**Weekly Calls**	**Ports**	**Weekly Calls**	**Ports**	**Weekly Calls**
Tanjung Pelepas	26	Algeciras	28.7	Algeciras	31.2
Algeciras	25.5	Hong Kong	26.5	Tanjung Pelepas	24
Hong Kong	20.88	Tanjung Pelepas	24.5	Bremerhaven	23
Bremerhaven	17.5	Bremerhaven	23	Hong Kong	23
Gioia Tauro	17	Port Said	18	Port Said	22
Port Klang	17	Rotterdam	17	Rotterdam	21
Rotterdam	16.5	Yantian	17	Shanghai	20
Felixstowe	15	Salalah	16	Yantian	18
Singapore	14	Balboa	15	Singapore	13
Salalah	11.5	Gioia Tauro	15	Port Klang	13
Miami	11	Shanghai	15	Ningbo	13
Kaohsiung	10.88	Singapore	13.5	Salalah	12.85
Yokohama	10.88	Kaohsiung	13.5	Tangier	11
Freeport	10.5	Yokohama	13	Santos	11
Le Havre	10	Port Newark	12	Port Newark	10
Port Newark	9	Santos	11.5	Felixstowe	10
Charleston	8	Charleston	11	Jebel Ali	10
Cape Town	8	Felixstowe	10.5	Busan	10
Penang	8	Busan	10	Puerto Manzanillo	10
Norfolk	7	Colombo	10	Gioia Tauro	9.5
Kobe	7				

Source: Calculated by the author based on the input data.

In COSCO's network the ports that received the most weekly calls were all coming from the region of North East Asia with the exception of Singapore. This is in contrast to Maersk's most frequently visited ports which were coming from various regions. The most frequently visited port in the company's network throughout the examined period was Shanghai. It received 29.33 weekly calls in 2001, experienced a very significant increase in 2007 with 45.9 weekly calls and remained at this level after the crisis with 45 weekly calls in 2010. Hong Kong was the second most frequently visited port by the company in 2001, having attracted 26.5 weekly calls. Experienced also an important increase to 42.5 weekly calls in 2007 and remained at the 2nd place, but after the crisis lost many of its calls and dropped to 30 weekly calls in 2010 and 5th place.

Singapore received 21 weekly calls in 2001 and was at the 3rd place of the company's most frequently visited ports. During the pre-crisis period almost doubled its weekly calls and in 2007 remained at the 3rd place very close to Shanghai and Hong Kong. In 2010 lost around 7 weekly calls compared to 2007 and dropped to 4th place. The South Korean port of Busan had a stable presence at the top 10 of the most frequently visited ports of the company. Its weekly calls were significantly increased from 14 to 23.5 between 2001 and 2007 and remained around the same level in 2010.

As happened with their degree, Ningbo and Yantian experienced a huge increase of the frequency that the company was calling at them prior the crisis. COSCO realized 25 weekly calls at Ningbo in 2007 compared to 6 in 2001 and 21.5 at Yantian compared to 5.5 respectively. After the crisis the company kept increasing its calling frequency at the two ports with Ningbo receiving 34 weekly calls and Yantian 27.55 in 2010. Qingdao, throughout the examined period, and Dalian, mostly in 2001 and 2010, ranked high at the company's calling frequency rankings but most of their calls were coming from feeder services.

Table 4-9: Most frequently visited ports in COSCO's network.

2001		2007		2010	
Ports	**Weekly Calls**	**Ports**	**Weekly Calls**	**Ports**	**Weekly Calls**
Shanghai	29.33	Shanghai	45.9	Shanghai	45
Hong Kong	26.5	Hong Kong	42.5	Qingdao	36
Singapore	21	Singapore	39.25	Ningbo	34
Qingdao	19	Ningbo	25	Singapore	32
Dalian	18.8	Busan	23.5	Hong kong	30
Busan	14	Yantian	21.5	Yantian	27.55
Kobe	13	Qingdao	18	Busan	22
Xingang	13	Kaohsiung	16	Dalian	19.55
Yingkou	12.8	Shekou	14.94	Tokyo	16
Yokohama	12.5	Xingang	14.12	Xiamen	15

Tokyo	11.5	Yokohama	13.5	Shekou	14.75
Nagoya	11.5	Tokyo	11	Osaka	13
Port Said	7	Xiamen	11	Yokohama	13
Felixstowe	7	Kobe	10.5	Kaohsiung	13
Osaka	6	Nagoya	10	Kobe	12
Ningbo	6	Dalian	9.75	Xingang	11.95
Shekou	6	Osaka	9	Guangzhou	11.1
Antwerp	6	Oakland	9	Yingkou	10.3
Rotterdam	6	New York	9	Nagoya	10
New York	5.7	Huangpu	8.98	Weihai	10
				Rizhao	10

Source: Calculated by the author based on the input data.

4.6.4 Links weight-time proximities between ports.

The use of frequency as a weight of the links between ports provides the opportunity to understand better the time proximities between ports. It can be identified which port pairs were most frequently connected in the carriers' networks. It is observed for both companies that geographical distance is a decisive factor for the frequency of inter-port links. Throughout the examined period almost all the frequent connections were between ports of the same region. The only exception for Maersk was the connection between Tanjung Pelepas (Southeast Asia) and Hong Kong (Northeast Asia) in 2001 with 6 weekly calls and between Tanjung Pelepas (Southeast Asia) and Yantian (Northeast Asia) in 2007 and 2010 with 9 and 7 weekly calls respectively.

In COSCO's network the only exception was the link between Singapore (Southeast Asia) and Hong Kong (Northeast Asia) with 9, 12 and 7 weekly calls in 2001, 2007 and 2010 respectively. Maersk has established a global network and its most frequent inter-port links were distributed in different regions all over the world. On the other hand, COSCO's most frequent inter-port links were almost exclusively in North East Asia. In 2001 Maersk's top 10 links had larger average frequency than the top 10 links of COSCO but that fact changed in 2007 and 2010 with COSCO having a much larger average frequency of its top 10 links.

In Maersk's network in 2001 the most frequent connections were between ports in Southeast Asia and between European ports. There was only one frequent link between North East Asian ports between Hong Kong and Kaohsiung. The most frequently connected ports were Port Klang and Penang with 12 weekly calls. That frequent connection was attributed to some very frequent feeder services between the two ports.

The intra-South East Asian link between Tangung Pelepas and Singapore with 11 weekly calls was the most frequent, non-feeder, link. In 2007 the most frequent connections were stretched in more regions. The most frequent link was an intra-North East Asian link between Yantian and Hong Kong with 12 weekly calls. There were frequent intra-regional links between ports in North East Asia, South East Asia, North Europe, East Coast of North America, Australia, South Africa and South Asia. After the crisis in 2010 the most frequently connected ports in the company's network were Shanghai and Ningbo with 13 weekly calls. The most frequent links were intra-regional links between ports in North East Asia, South East Asia, North Europe and the Mediterranean Sea.

Contrary to Maersk which connected frequently ports in different regions around the world, COSCO's most frequent links were almost all in the region of North East Asia. The only exceptions were the links between Norfolk-Charleston in East Coast of North America with 7 weekly calls in 2007 and the inter-regional links of Singapore-Hong Kong in 2001, 2007 and 2010 and Singapore-Port Said, Singapore-Shekou in 2010.

In 2001 the most frequent links were between Qingdao-shanghai and Dalian-Xingang but mainly because of a very frequent feeder service. Singapore-Hong Kong was the most frequent non-feeder link with 9 weekly calls. In 2007 the link between the neighboring ports of Shanghai and Ningbo emerged as the most frequent with the large figure of 18 weekly calls and continued to be the most frequent link after the crisis in 2010 with 21 weekly calls which was the most frequent link in both companies' networks throughout the examined period.

Table 4-10: Top links frequency hierarchy in Maersk's network.

2001		2007		2010	
Links	**Freq**	**Links**	**Freq**	**Links**	**Freq**
Port klang-Penang	12	Yantian-Hong Kong	12	Shanghai-Ningbo	13
Tanjung Pelepas-Singapore	11	Yantian-Tanjung Pelepas	9	Hong Kong-Yantian	12
Felixstowe-Rotterdam	10	Shanghai-Ningbo	8	Rotterdam-Bremerhaven	11
Port klang-Tanjung Pelepas	9	Bremerhaven-Rotterdam	8	Tangier-Algeciras	8
Hong kong-Kaohsiung	8.88	Norfolk-Charleston	7	Yantian-Tanjung Pelepas	7
Singapore-Port Klang	8	Sydney-Melbourne	7	Port Klang-Singapore	7
Bremerhaven-Hamburg	8	Hong Kong-Kaohsiung	6.5	Tanjung Pelepas-Port Klang	6
Rotterdam-Bremerhaven	7.5	Port Elizabeth-Durban	6	Nagoya-Yokohama	5
Port newark-Norfolk	6	Xingang-Qingdao	6	Damietta-Port Said	5
Tanjung pelepas-Hong Kong	6	Colombo-Chennai	6	Felixstowe-Bremerhaven	5

		Bremerhaven-St Petersburg	6	Felixstowe-Rotterdam	5
				Bremerhaven-St Petersburg	5
				Busan-Shanghai	5
				Melbourne-Sydney	5

Freq: Frequency of the link (weekly calls). Source: Retrieved by the author based on the input data.

Table 4-11: Top links frequency hierarchy in COSCO's network.

2001		2007		2010	
Links	**Freq**	**Links**	**Freq**	**Links**	**Freq**
Qingdao-Shanghai	11	Shanghai-Ningbo	18	Shanghai-Ningbo	21
Dalian-Xingang	10	Hong Kong-Yantian	16	Rizhao-Qingdao	19
Hong kong-Singapore	9	Shanghai-Busan	14.5	Weihai-Qingdao	18
Xingang-Yingkou	8	Singapore-Hong Kong	12	Hong Kong-Yantian	14
Shekou-Hong Kong	7	Hong Kong-Shanghai	10.5	Busan-Shanghai	11
Shanghai-Dalian	7	Long Beach-Oakland	7	Osaka-Kobe	9
Yingkou-Qingdao	7	Hong Kong-Shekou	7	Qingdao-Yantian	9
Shanghai-Hong Kong	6	Tokyo-Yokohama	7	Wenzhou-Ningbo	8
Yantian-Hong Kong	6	Ningbo-Xiamen	6	Dalian-Qinhuangdao	8
Tokyo-Yokohama	6	Kaohsiung-Yantian	6	Yantian-Shanghai	7
		Osaka-Kobe	6	Singapore-Port Said	7
		Shekou-Huangpu	6	Hong kong-Singapore	7
				Shekou-Singapore	7
				Yingkou-Dalian	7
				Tokyo-Yokohama	7

Freq: Frequency of the link (weekly calls). Source: Retrieved by the author based on the input data.

4.7 Analysis of the regional networks of shipping lines

4.7.1 Analysis of inter-regional services and evolution of regional networks

The container liner shipping companies operate inter-regional liner services which connect various regions around the world. Xu et al. (2015) analyzed the positions of the world regions in the global maritime network based on the aggregated services of the top 100 liner carriers. Tran and Haasis (2014) investigated the development of the regional networks in the East-West axis based on the aggregated services of the top 20 liner shipping companies. This study analyses the evolution of the regional networks of individual carriers based on their inter-regional services.

There were identified 20 regions around the world which were connected by the liner services of the selected carriers. These regions were: North Europe (NEUR), Mediterranean Sea (MED), Middle East (ME), South Asia (SA), South East Asia (SEA), North East Asia (NEA), Australia (AUS), West Coast of North America (WCNA), East Coast of North America, (ECNA), United States Gulf Coast (USGC), Central America and the Caribbean (CAM), North Coast of South America (NCSA), West Coast of South America (WCSA), East Coast of South America (ECSA), West Africa (WAF), South Africa (SAF), East Africa (EAF), Indian Ocean (IND OC), North Pacific (N PAC) and South Pacific (S PAC). Both inter-regional and intra-regional calls from the inter-regional services of the carries are counted at the ports of its region.

Maersk and COSCO have adopted different strategies in the formation of their shipping networks. The Danish Maersk has established a global network with strong presence all over the world. On the other hand, the Chinese COSCO operated mainly in the East-West axis. The region of North East Asia is by far the dominant region of its network.

Both companies showed an important increase in their inter-regional services from 2001 to the pre-crisis year 2007 and then a decrease in the after-crisis year 2010. The number of services for Maersk went from 61 in 2001 to 98 in 2007 and dropped to 77 in 2010. Respectively the figures for COSCO were 31, 54 and 43 services. The average number of calls per service followed an inverse course for both companies compared to the number of services. Decreased from 2001 to 2007 and then increased in 2010. Maersk called on average at 8 to 9 ports per service in 2001, around 8 ports in 2007 and around 9 ports in 2010. The corresponding figures for COSCO are 9 to 10 ports in 2001, around 8 ports in 2007 and around 9 ports in 2010.

It is remarkable that although the two companies have different geographical coverage and complexity in their services, they have the exact same number of average calls per service in 2007 as in 2010 (table). The average number of regions served per service in 2001, 2007 and 2010 is close to 3 for Maersk while COSCO served between 2 and 3 regions per service at the same years. Both companies use the pattern of calling twice at a port in the same service. The number of double calls per service of Maersk decreased between 2001 and 2007 and increased between 2007 and 2010. COSCO limited its double calls per service in both periods. On average both companies made around one double call per service.

Table 4-12: Analysis of inter-regional services.

Maersk	2001	2007	2010
Average number of calls per service	8.70	8.22	9.27
Average number of weekly calls per service	8.49	7.76	9.03
Average number of regions called per service	2.72	2.77	2.89
Average number of double calls at ports per service	1.15	0.82	1.12
Number of Inter-regional services	61	98	77

Source: Computed by the author based on CIY data.

Table 4-13: Analysis of inter-regional services.

COSCO	2001	2007	2010
Average number of calls per service	9.53	8.26	9.28
Average number of weekly calls per service	9.11	8.09	9.20
Average number of regions called per service	2.60	2.39	2.53
Average number of double calls at ports per service	1.48	1	0.95
Number of Inter-regional services	31	54	43

Source: Computed by the author based on CIY data.

Both companies experienced an important increase in their total weekly calls between 2001 and the pre-crisis year 2007. Maersk had an increase of 46.8% and COSCO of 52.1%. After the crisis of 2009 there was shrinkage of the total weekly calls for both companies. In 2010 the total weekly calls of Maersk were reduced by 8.4% compared to 2007 and of COSCO by 7.1% respectively. Maersk has a strong weekly presence at all regions all over the world while COSCO mainly calls at Northeast Asia (NEA), Southeast Asia (SEA), Mediterranean (MED), North Europe (NEUR), East coast North America (ECNA), West coast North America (WCNA) and Australia (AUS) at the South hemisphere. Its presence at the other regions is very slim.

Table 4-14: Evolution of Maersk's regional networks.

Maersk	2001			2007			2010		
	WC	Ser	WC/Ser	WC	Ser	WC/Ser	WC	Ser	WC/Ser
NEUR	72	18	4.00	76.7	24	3.20	83	24	3.46
MED	50.5	18	2.81	81.7	35	2.33	89.9	30	3.00
ME	18.5	10	1.85	26	15	1.73	32.85	18	1.83
SA	3	3	1.00	14	9	1.56	17	10	1.70
SEA	24	12	2.00	51	25	2.04	52	22	2.36
NEA	78.25	14	5.59	159	31	5.13	143	23	6.22
ECNA	51	18	2.83	56	15	3.73	44	11	4.00
WCNA	16	7	2.29	21	12	1.75	15	9	1.67
CAM	62.5	21	2.98	56.05	21	2.67	48	13	3.69

USGC	8	5	1.60	8	5	1.60	10	4	2.50
ECSA	19.5	5	3.90	41	10	4.10	48	7	6.86
NCSA	15.5	7	2.21	20.77	8	2.60	7	4	1.75
WCSA	9	2	4.50	6	3	2.00	3	2	1.50
WAF	29.75	8	3.72	53.9	21	2.57	42.2	18	2.34
SAF	22	8	2.75	24	11	2.18	17	9	1.89
EAF	5	2	2.50	5	3	1.67	7.6	6	1.27
AUS	23.63	5	4.73	46.5	13	3.58	27.5	7	3.93
N PAC	2	1	2.00	2.25	2	1.13	2.5	2	1.25
S PAC	0	0	0.00	3.25	5	0.65	1	2	0.50
IND OC	7.5	2	3.75	8	4	2.00	6	2	3.00
Total	517.63			760.12			696.55		

WC: Number of weekly calls at regional ports; **Ser**: Number of services calling at the region; **WC/Ser**: Average number of weekly calls at regional ports per service. Source: Computed by the author based on CIY data.

Table 4-15: Evolution of COSCO's regional networks.

COSCO		**2001**			**2007**			**2010**	
	WC	**Ser**	**WC/Ser**	**WC**	**Ser**	**WC/Ser**	**WC**	**Ser**	**WC/Ser**
NEUR	29	7	4.14	25	7	3.57	24	7	3.43
MED	33.4	9	3.71	35.5	10	3.55	39	10	3.90
ME	7	3	2.33	12	5	2.40	12	5	2.40
SA	5	3	1.67	6	4	1.50	3	1	3.00
SEA	31	13	2.38	57	25	2.28	45	21	2.14
NEA	113.5	23	4.93	204.5	44	4.65	191	37	5.16
ECNA	19.4	6	3.23	28	9	3.11	26	8	3.25
WCNA	16	7	2.29	31	14	2.21	28	12	2.33
CAM	2	1	2.00	2	1	2.00	2	1	2.00
USGC	2	1	2.00	2	1	2.00	0	0	0.00
ECSA	4	1	4.00	4	1	4.00	4	1	4.00
NCSA	0	0	0.00	0	0	0.00	0	0	0.00
WCSA	0	0	0.00	0	0	0.00	0	0	0.00
WAF	3	1	3.00	1.5	1	1.50	4	1	4.00
SAF	4	1	4.00	2	1	2.00	0	0	0.00
EAF	0	0	0.00	0	0	0.00	0	0	0.00
AUS	12.5	4	3.13	17.5	5	3.50	21	5	4.20
N PAC	0	0	0.00	0	0	0.00	0	0	0.00
S PAC	0	0	0.00	1.5	1	1.50	0	0	0.00
IND OC	0.5	1	0.50	0	0	0.00	0	0	0.00
Total	282.3			429.5			399		

WC: Number of weekly calls at regional ports; **Ser**: Number of services calling at the region; **WC/Ser**: Average number of weekly calls at regional ports per service. Source: Computed by the author based on CIY data.

4.7.2 Analysis of networks of regions

In the previous sections of this study the network analysis is conducted at the port level. In this section the network analysis is carried out at the region level. The nodes of the constructed networks of the shipping lines are the regions they visit. Degree centrality which shows the number of connections that a region had with the other regions of the shipping lines networks and betweenness centrality that indicates how centrally was a region placed in the networks were calculated for both companies (Table 4-16 and Table 4-17).

Table 4-16: Degree and betweenness centrality of regions in Maersk's network.

MAERSK	2001		2007		2010	
	Degree	BC	Degree	BC	Degree	BC
AUS	4	2.3207	6	8.5844	6	15.5556
CAM	10	47.6280	12	34.6072	11	48.2806
EAF	1	0.0000	3	0.0000	3	0.0000
ECNA	7	9.5303	10	14.9035	8	13.8968
ECSA	5	3.0489	9	8.5196	4	0.8222
IND OC	2	0.0000	4	0.5945	3	0.0000
ME	7	25.8381	8	7.0549	9	8.3024
MED	9	22.2132	10	9.7625	10	12.0234
N PAC	2	0.0000	3	0.0000	3	0.0000
NCSA	3	0.6778	3	0.0000	2	0.0000
NEA	6	16.6580	11	22.9362	10	22.9619
NEUR	7	12.3330	10	5.1125	8	6.5853
S PAC	0	0.0000	3	0.0000	2	0.0000
SA	2	0.0000	5	1.2814	5	0.3095
SAF	8	23.2664	10	11.4237	9	10.4520
SEA	7	18.8429	11	22.7184	11	21.6147
USGC	2	0.0000	3	0.0000	2	0.0000
WAF	3	0.1429	6	0.2429	6	1.8512
WCNA	4	3.5000	5	2.2583	5	7.3444
WCSA	1	0.0000	2	0.0000	1	0.0000

BC: Betweenness Centrality. Source: Retrieved by the author based on the input data.

Table 4-17: Degree and betweenness centrality of regions in COSCO's network.

COSCO	2001		2007		2010	
	Degree	BC	Degree	BC	Degree	BC
AUS	1	0.0000	2	0.0000	2	0.0000
CAM	2	6.0000	1	0.0000	2	0.0000
EAF	0	0.0000	0	0.0000	0	0.0000

ECNA	4	17.0000	3	3.3333	5	3.0000
ECSA	1	0.0000	1	0.0000	1	0.0000
IND OC	2	0.0000	0	0.0000	0	0.0000
ME	3	0.3333	4	3.0000	3	0.0000
MED	5	19.8333	5	15.3333	5	10.0000
N PAC	0	0.0000	0	0.0000	0	0.0000
NCSA	0	0.0000	0	0.0000	0	0.0000
NEA	4	30.0000	7	30.0000	8	18.5000
NEUR	4	7.1667	5	15.8333	5	1.0000
S PAC	0	0.0000	1	0.0000	0	0.0000
SA	3	0.6667	4	2.0000	1	0.0000
SAF	3	13.0000	1	0.0000	0	0.0000
SEA	7	50.0000	9	45.5000	8	22.5000
USGC	1	0.0000	1	0.0000	0	0.0000
WAF	1	0.0000	1	0.0000	1	0.0000
WCNA	1	0.0000	1	0.0000	1	0.0000
WCSA	0	0.0000	0	0.0000	0	0.0000

BC: Betweenness Centrality. Source: Retrieved by the author based on the input data.

The findings of the network analysis of regions confirm that the network of Maersk is a global network which is also well connected. There are regions all over the world that are connected with a large number of other regions. In Maersk's network the region with the highest connectivity and centrality throughout the examined period was the region of Central America and the Caribbean (CAM). This observation is explained in a degree from the fact that many of the regions that are used in this study are in geographical proximity with CAM. SEA, MED, NEA, NEUR, SAF and ME were regions in different continents with a large number of connections with other regions. In 2001, after CAM the most centrally placed region in the network was ME but from 2007 the most centrally placed regions, after CAM, where in Asia (SEA AND NEA).

COSCO was operating mainly in the East-West axis. SEA was the region with the highest connectivity and centrality throughout the examined period. NEA was the second region in number of connections with other regions and betweenness centrality. Outside Asia, MED, NEUR and ECNA were the other regions that followed in degree and betweenness centrality. After the crisis, the network of the company was extremely centralized around Asia (SEA and NEA regions).

4.7.3 Regional networks of shipping lines

North Europe (NEUR)

The NEUR region had a stable presence in the network of both companies throughout the period 2001-2010. It didn't manage to gain any serious increment in the weekly calls of Maersk at the pre-crisis period but also wasn't affected from the crisis of 2009 and experienced a small increase in 2010 compared to 2007. COSCO also had a very stable calling frequency in NEUR throughout the period 2001-2010. The number of its services was the same in 2001, 2007 and 2010 with a very small decrease in weekly calls.

Mediterranean Sea (MED)

MED experienced a very important increase in the calling frequency of Maersk between 2001 and 2007. Maersk doubled its services at the region and increased its weekly calls by 61.8%. As a result, MED became the region with the most calling services and second in weekly calls at its port at the network of the company. The crisis of 2009 slowed down the increase of Maersk visits at MED's ports but still in 2010 the region kept its position in the hierarchy of the company's network. COSCO, as with NEUR, has a very stable calling frequency at MED without any significant changes throughout the period 2001-2010.

South East Asia (SEA)

Maersk doubled its services and its weekly calls at the ports of SEA between 2001 and the pre-crisis year 2007. In 2010, after the crisis, the region kept the same calling levels from Maersk as in 2007. COSCO also doubled its services and its weekly calls at the region between 2001 and 2007 and SEA became the second most important region in the company's network. The economic crisis affected importantly the visits of COSCO at the region and in 2010 the weekly calls at its ports were reduced by 21%. SEA was the most connected and most centrally placed region in COSCO's network as it was the region with the largest degree and betweenness centrality throughout the examined period.

North East Asia (NEA)

NEA was the number one region in weekly calls in both companies' networks. In Maersk's network it wasn't the region with the most calling services but it had a very large number of calling ports per service with an average between 5 and 6 port calls per service in 2001, 2007 and 2010. At Maersk's network, NEA had 1.3 times more weekly calls than the second region in 2001, 1.9 in 2007 and 1.6 in 2010. The dominance of NEA in COSCO's network was much more emphatic as the gap with the second region in weekly calls was huge. In 2001 the difference was 3.4 times, in 2007 3.6 and in 2010 4.2. In both companies' networks the region experienced a large increase in calling services and weekly calls at its ports between 2001 and 2007 and then a decrease in the after-crisis year 2010. Maersk increased its services by 121.4% and its weekly calls by 103.2%. The corresponding figures for COSCO were 80.2% and 91.3%. In 2010 compared to 2007, Maersk's services decreased by 25.8% and its weekly calls at the region's ports by 10%. The figures for COSCO were 18.1% and 6.6% respectively. The connectivity of the region with other regions was significantly increased in the pre-crisis period in both companies' networks. The region was also centrally placed in both companies' networks as its betweenness centrality was the 2nd largest in Maersk's network in 2007 and 2010 and also the 2nd largest in COSCO's network throughout the examined period.

East Coast of North America (ECNA)

ECNA experienced an increase of 9.8% in weekly calls from Maersk between 2001 and 2007 and then a decrease of 21.4% between 2010 and 2007. Maersk continued to reduce the number of its services in pre-crisis and after-crisis years at the region and to increase the number of visited ports per service. In 2001 the average number of visited ports per service was 2.8 in 2001, 3.7 in 2007 and 4 in 2010. COSCO increased its weekly calls at the region by 44.3% in 2007 compared to 2001 and kept the same level of frequency after the crisis.

West Coast of North America (WCNA) - North Pacific (N PAC)

A characteristic of the calling patterns of both companies at the region is the very limited number of visited ports per service. Maersk visited on average 2.3 ports per service in 2001, 1.75 ports in 2007 and 1.67 ports in 2010. COSCO visited 2.3 ports per service in 2001, 2.2 ports in 2007 and 2.3 ports in 2010. Maersk increased its weekly calls at the region between 2001 and 2007 but in 2010 returned at the levels of 2001. The company had also a sparse presence at the islands of North Pacific prior and after the crisis. On the other hand COSCO doubled the number of its services and its weekly calls in 2007 compared to 2001 and as with ECNA kept the level of frequency at WCNA almost intact after the crisis. The company didn't visit the region of North Pacific.

Middle East (ME)

ME is one of the two regions in the network of Maersk that had an upward course in all the examined years, both in calling services and weekly visits at its ports. Maersk increased its weekly calls at ME's ports by 77.6% between 2001 and 2010 and by 80% the number of its services at the region. COSCO increased its presence at the region between 2001 and 2007 and continued with the exact same number of services and weekly calls in the after-crisis year 2010.

South Asia (SA)

SA is the second region in Maersk's network with an upward course in all the examined years, both in calling services and weekly visits at its ports. Until the pre-crisis year 2007 the region had an important growth in weekly calls from Maersk, but this tendency slowed down from the crisis of 2009. The economic crisis affected also the calling frequency of COSCO at the region. Its slim presence until 2007 became even slimmer after the crisis, with only one service calling at the region.

Central America and the Caribbean (CAM)-United States Gulf Coast (USGC)

It was remarkable the very strong presence of Maersk at the region of CAM. CAM was the region with the largest degree and betweenness centrality throughout the examined period. In 2001 CAM had the largest number of calling services among all the regions in the network of Maersk and was second in weekly calls at its ports only behind NEA. Maersk followed a strategy of reducing its calling frequency at the region. Although reduced, still in 2010 the presence of Maersk was remarkable with 48 weekly calls.

The region of USGC is not frequently visited by Maersk and this strategy was stable throughout the examined period. COSCO's network is limited at the East-West axis. Its presence at the North-South or South-South routes is minimal. Only one service of the company was calling at CAM throughout the period. USGC was also receiving only one service until 2007. In 2010 the company didn't call at the region.

East Coast South America (ECSA)-North Coast South America (NCSA)-West Coast South America (WCSA)

ECSA is the most visited region by Maersk in South America. In 2007 Maersk doubled the number of its services and the weekly calls at the ports of the region. After the crisis, in 2010, the weekly calls at ECNA's ports continued to grow but with a much smaller growth rate. It is remarkable that the average number of port calls per service was continuously growing and in 2010 it was the biggest among all regions at the whole examined period with almost seven port calls per service. NCSA is the second most visited region by Maersk in South America. The region was seriously affected by the company's strategy after the crisis of 2009, and in 2010 lost 66.3 of its weekly calls compared to the pre-crisis year 2007. WCSA is the most neglected region by Maersk in South America. The poor presence of the company at the region followed a decreasing course, and in 2010 the region received only three weekly calls. COSCO had only one service calling at ECSA and it didn't call at all at NCSA and WCSA throughout the examined period.

West Africa (WAF)-South Africa (SAF)-East Africa (EAF)-Indian Ocean (IND OC)

Maersk as a truly global carrier has a noticeable presence also in Africa. WAF is the region of the continent with the highest number of weekly calls at its ports. In the pre-crisis period, between 2001 and 2007, the company almost tripled its services that were calling at the region. WAF experienced a very significant increase of 81.1% at the weekly calls at its ports. The reorganization of the company's network after the crisis led to a decrease of 19.8% at its weekly calls at the region. SAF is the second most visited region by Maersk in Africa. It didn't experience any significant increase at the pre-crisis period and was also affected by the crisis and lost about one third of its weekly calls in 2010 compared to 2007.

EAF is the least visited region by Maersk in the continent. It is the only region that experienced a decrease in weekly calls at the pre-crisis period 2001-2007 and recovered in the after-crisis year 2010 close to the level of frequency of 2001. Maersk was also visiting the islands of IND OC with the region receiving 8 weekly calls at its ports in 2007 which were cut to half after the crisis. COSCO's presence at the continent is extremely limited with only one service calling at WAF prior and after the crisis, one service calling at SAF only until 2007 and one service calling at IND OC only in 2001.

Australia (AUS)-South Pacific (S PAC)

The presence of Maersk in AUS followed an up and down course, prior and after the crisis, as happened with other regions in its network. Maersk doubled its weekly calls at the ports of AUS in 2007 compared to 2001 and then in 2010 returned back almost at the number of weekly calls it was realizing in 2001. Maersk was also visiting sporadically the islands of South Pacific in 2007 and limited its presence even more in 2010. AUS is the only region, apart from the regions in the East-West axis, that COSCO has a noticeable presence. Its weekly calls at the region increased prior the crisis of 2009 and continued to increase after it with a smaller growth rate. The company called only in 2007 in South Pacific with only one service.

4.8 Conclusions

This research analyses the container liner shipping networks of individual liner shipping companies at selected years prior and after the financial crisis of 2009. Network indicators are calculated, and network visualization is realized on the theoretical basis provided from graph theory, complex networks and statistical techniques. Maersk operated a much larger network than COSCO throughout the examined period. COSCO's network was smaller, denser and its ports had on average more connections and received more weekly calls than those in Maersk's network throughout the examined period. The evolution of the networks of the two carriers was different as far as their sizes were concerned. Maersk's network was shrunk after the crisis while COSCO's remained at the same level.

The visualization of the carriers' networks revealed a common structural pattern prior and after the crisis. In both companies' networks communities of ports were identified that consisted of ports in geographical proximity. The degree and betweenness centrality rankings confirmed Notteboom's (2022a) observation that liner shipping companies select different ports in order to form the backbone of their networks. Maersk visited frequently ports in various regions all over the world while all COSCO's most frequently visited ports, except Singapore, were placed in North East Asia. The use of frequency (weekly calls) as a weight of the inter-port links allowed a better understanding of the time proximities between ports. A common characteristic of both companies' networks was that their most frequent links were between ports of the same region (intra-regional links) throughout the examined period. The difference was that these frequent links of Maersk were identified in various regions all over the world while almost all of COSCO's frequent links belonged to North East Asia region.

The analysis of the regional networks of the companies confirmed the difference in the geographical coverage of their networks. Maersk had established a global network and had been visiting 20 regions all over the world. On the other hand, COSCO was operating mainly in the East-West axis at the triad Asia-Europe-North America. The ports of North East Asia, the world's biggest manufacturing region, received by far of the second region the biggest number of calls from both companies throughout the examined period. The region experienced a huge increase in the weekly calls at its ports from both companies in the pre-crisis period. After the crisis it experienced a small decrease of the weekly calls of both companies. The region with the highest number of connections with the other regions of the network was Central America and the Caribbean (CAM) in Maersk's network and South East Asia (SEA) in COSCO's. These regions were also the most centrally placed in the companies' networks.

The networks of the companies are investigated for the existence of complex network properties. The power law distribution has a very good fit to the degree distributions of both companies at all the examined years. As a result, the networks of both companies throughout the examined period can be characterized as scale-free networks. The characteristic pattern of the connections of nodes at this type of networks is the existence of very few nodes with a large number of connections and many nodes with a small number of connections. Consequently, only a very limited number of ports had a large number of connections in both companies' networks and all the other ports had very few connections. The most frequent number of connections of the ports of both companies' networks was 2.

Chapter 5

Point and interval forecasting of container throughput with combined SARIMA-(G)ARCH model

5.1 Introduction

Operational short-term planning of container ports relies heavily on short-term forecasts of container throughput volumes. These forecasts are a valuable tool in the hands of container terminal operators and port authorities in order to make decisions about the allocation of port's resources. These operational decisions aim to avoid congestion and handle the container volumes with the most efficient way (Rashed et. al, 2017). For these purposes, in peak months short-term measures can be adopted (e.g., operation for extended hours) (Farhan and Ong, 2018). Consequently, the forecasts of seasonal fluctuations of the demand for container handling services are very crucial for the port stakeholders in order to take actions that will maintain a high level at their provided services.

Due to its importance for port planning, forecasting of container throughput has been the subject of a large number of studies. These studies produce point forecasts of the container throughput, but they do not assess the uncertainty of these forecasts. The traffic flows arriving at ports are characterized by a high level of volatility caused by factors as seasonal fluctuations, choices of important market players and all types of crises (Notteboom, 2022b). Under these circumstances, the creation of accurate forecasts of the future values of cargo flows is a very challenging exercise and will be accompanied with a degree of uncertainty. The modelling and forecasting of the volatility of container throughput at the port level and the associated uncertainty of its forecasts haven't yet been addressed.

The uncertainty of the forecasts is expressed by the prediction intervals associated with the point forecasts. Consequently, it is essential to also produce prediction intervals for the point forecasts in order to estimate the level of uncertainty of each forecast (Hyndman and Athanasopoulos, 2018). In order to produce both point forecast and prediction intervals, this study introduces a combined Seasonal Autoregressive Integrated Moving Average (SARIMA) - Autoregressive Conditional Heteroskedasticity (ARCH) or Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model.

The point forecasts are generated form the SARIMA model and the prediction intervals from the ARCH or GARCH model. ARCH/GARCH models are used in order to model and forecast the volatility of container throughput values. These forecasts of the volatility of each forecast error are used for the construction of the prediction intervals.

(S)ARIMA-(G)ARCH models have been used in a very wide range of fields for forecasting purposes, indicatively from stock prices (Ena et. al, 2022), gold prices (Shetty et. al, 2018), oil prices (Xiang, 2022) to solar irradiance (David et. al, 2016), groundwater depth (Guo et.al, 2021), particulate matter (Alexis et. al, 2022) and rainfalls (Pandey et. al, 2018). To the best of the author's knowledge, this combined model hasn't been used yet for the modeling and forecasting of the container throughput of ports. The closest application of this model to the port industry is the forecasting of containers shipping freight rates for the Far East - Northern Europe trade line (Munim and Schramm, 2017). This study applies the combined SARIMA-(G)ARCH model in order to produce short-term forecasts and its accompanying prediction intervals of the total monthly container throughput of the port of Barcelona.

The remainder of the chapter is structured as follows: Section 5.2 contains a literature review of the forecasting techniques of quantitative methods that have been used for the forecasting of container throughput of ports. Section 5.3 introduces the methodology that was applied in this study, the mathematical formulation of the models, the approach for the construction of prediction intervals and the measures for the evaluation of the models' performance. Section 5.4 presents the application of the methodology and the results for the monthly container throughput of the port of Barcelona. Section 5.5 provides the evaluation of the forecasting performance of the SARIMA-(G)ARCH model. The final section 5.6 presents the conclusions of this study.

5.2 Literature review

Quantitative methods of container throughput forecasting are divided in two main approaches. The first approach relies exclusively on time series analysis. The predictions of the future values of container throughput are based on the analysis of its past values. The historical patterns of the time series of container throughput are modelled and the future values are forecasted by these univariate models. The second approach is based on the assumption that there is a causal relationship between container throughput and some external variables. In this way, multivariate models are constructed and used for container throughput forecasting (Notteboom, 2022b).

5.2.1 Univariate models

5.2.1.1 Application of ARIMA/SARIMA model in container throughput forecasting

The autoregressive integrated moving average model (ARIMA) and the seasonal autoregressive integrated moving average model (SARIMA) have been used in several cases from researchers in order to forecast the container throughput of ports. Rashed et al. (2017) compared the forecast accuracy of three models for the container throughput of the port of Antwerp. A SARIMA model, a SARIMA-intervention model and a SARIMAX model (a SARIMA model with the inclusion of the industrial confidence indicator for the Euro Area as an exogenous variable). The SARIMAX model was found to provide the most accurate forecasts of the monthly container throughput (loaded and unloaded) of the port of Antwerp. Farhan and Ong (2018) explored the use of SARIMA models in short-term forecasting at several major international container ports. They developed a SARIMA model for each of the top-20 international container ports based on monthly time series data and concluded that the models can provide reliable throughput forecasts.

Dragan et al. (2014) compared the performance of an ARIMA model, a Holt-winters exponential smoothing model and a classical decomposition model in forecasting the annual container throughput of the North Adriatic ports. The ARIMA model outperformed the other two. Schulze and Prinz (2009) constructed a SARIMA model and a Holt-winters exponential smoothing model based on quarterly data, in order to forecast the aggregated total container throughput of German ports. They also applied the models to the destinations of Asia, Europe and North America. The SARIMA model proved to be more accurate than the exponential smoothing model. Peng and Chu (2009) compared six univariate forecasting models for the container throughput of the three major ports of Taiwan. They applied monthly data to these models and concluded that the classical decomposition model provided the most accurate predictions for the ports of Kaohsiung and Keelung and had the same performance with SARIMA model for the port of Taichung.

Chan et al. (2019) also compared six univariate forecasting models for the container throughput of the port of Ningbo. Four of the six models were classical time series models and the rest two were machine learning models. They applied a small dataset of yearly data to the models and the ARIMA model had the best performance among the classical time series models and second overall behind the Support Vector Regression (SVR) machine learning model. Pang and Gebka (2017) employed four univariate time series models (SARIMA, the additive and multiplicative Seasonal Holt-Winters, and the Vector Error Correction Model) on monthly time series data in ordered to generate forecasts for each of the terminals of Tanjung Priok port and aggregate them into the total throughput forecast. The results suggested that this approach provided consistently worse results than forecasts of the total container throughput of the port based on modelling the total container throughput time series.

Chen and Chen (2010) explored a genetic programming model (GP), a decomposition approach (X-11) and a SARIMA model in order to find an optimal model for forecasting the container throughput of Taiwanese ports. They applied monthly data to the three models and all of them proved to be accurate, but the GP model had the best performance. Milenkovic et al. (2021) compared two hybrid forecasting models, based on fuzzy artificial neural networks (FANN) combined with heuristic algorithms, with ARIMA/SARIMA models. They applied the models to five monthly datasets of the port of Barcelona (total, loaded, unloaded, transit and empty container throughput). The results suggested that the hybrid models (FANN - heuristic algorithms) provided more accurate predictions for all-time series compared to ARIMA/SARIMA models.

5.2.1.2 Application of other univariate time series models in container throughput forecasting

There are also several studies that used univariate time series models, apart from ARIMA and SARIMA, in order to predict the future values of container throughput. Twrdy and Batista (2016) applied yearly time series data of container throughput to four models in order to predict the container throughput of North Adriatic ports: a Markov-chain annual growth rate model, a time-series trend model, a time-series trend model with periodical terms and a gray system model. They discovered that a quadratic trend model with a two-cycle component had the best performance with a degree of error within 5%. A statistical model based on a two-state Markov model in conjunction with Monte Carlo experiments was introduced by Grifoll (2019) in order to forecast the container throughput of the multi-port gateway region of Spanish Mediterranean. He used yearly data and concluded that the model provided reasonably good predictions.

Xie et al (2013) proposed three hybrid approaches based on least squares support vector regression (LS-SVR) model for forecasting container throughput at ports. They used monthly time series data and applied their models in order to forecast the container throughput at the ports of Shanghai and Shenzhen. They compared their hybrid models with other commonly used individual forecasting models and the results suggested that the hybrid models performed better than the individual approaches. He and Wang (2021) followed also a hybrid approach and suggested a fractional grey model (GM) combined with a back propagation (BP) neural network for forecasting the container throughput of Tianjin-Hebei Port Group. They used a small set of yearly data and concluded that their mixed model had a higher level of accuracy compared to other single models.

5.2.2 Forecasting models of container throughput based on causal relationships and time series analysis.

The second approach for the forecast of the container throughput of ports is based on both time series analysis and causal relationships. These models assume that a causal relation can be found between container throughput and one or more external variables. Chou et al. (2008) proposed a modified regression model for forecasting Taiwan's imports container throughput. They applied yearly data to the model and found that the total forecast errors were lower compared to the traditional regression model. Gosasang et al. (2011) entered yearly data of the factors that affect the container throughput of Bangkok's port into a Multilayer Perceptron (MLP) neural network and a linear regression forecasting model. The results suggested that the forecasting ability of the neural network model was superior compared to that of the linear regression model.

Neural networks' forecasting performance was also compared with that of linear regression models by Lam et al. (2004) in predicting the cargo throughput of the port of Hong Kong. They used yearly data of 37 types of freight movements and explanatory factors. As in the previous study, the predictions of the neural networks were considered more reliable and more comparable to reality. Ding et al. (2019) used a multivariate back propagation (BP) neural network-support vector machine (SVM) combined model in order to predict foreign trade container volume in Ningbo and Wenzhou. The results suggested that the relative error of the proposed mixed model was much lower than the other single models and that the model could be used effectively in logistics-demand predicting.

5.3 Methodology

This study uses a combined SARIMA-ARCH/GARCH model in order to model monthly time series data of the total container throughput of the port of Barcelona and generate point forecasts along with their accompanying prediction intervals. The point forecasts are produced from the SARIMA model. The prediction intervals of these forecasts are generated from the ARCH/GARCH model. The methodological framework for the construction of the SARIMA-ARCH/GARCH model and the production of point forecasts and their prediction intervals is presented in figure 5-1. The study uses monthly time series data of the total container throughput of the port of Barcelona from 2005-2021. The selection of the frequency of the data was based firstly on the short-term forecasting horizon of this study and secondly on the fact that the recommended number of available observations for the models used in this study (SARIMA-ARCH/GARCH) is over 50. The time series data were collected from the Spanish government agency, Puertos del Estado.

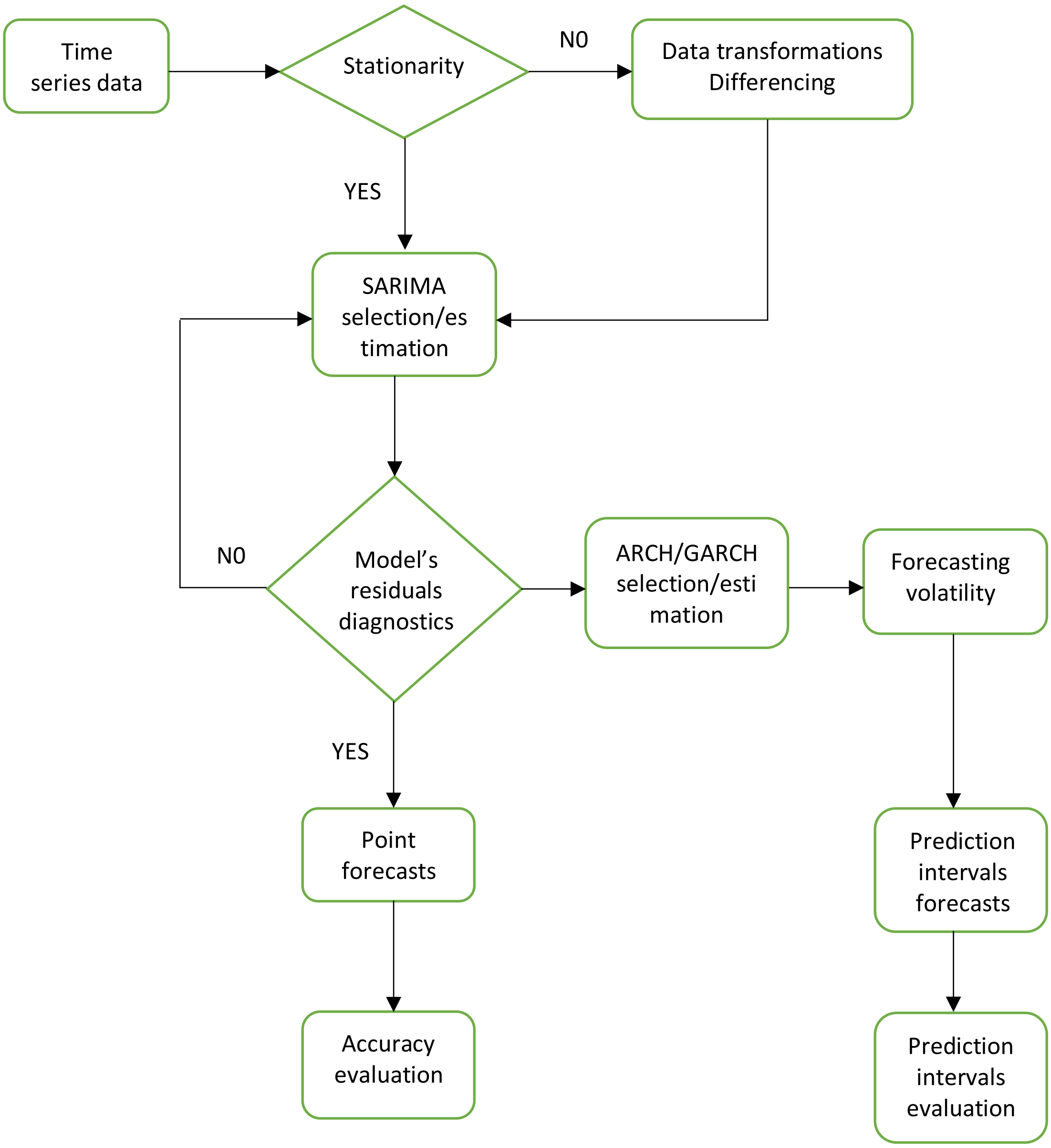

Figure 5-1: Methodological framework of SARIMA-ARCH/GARCH model for point and prediction intervals forecasts.
Source: Created by the author.

The monthly time series data of the container throughput of the two ports are split in two sets, the training set and the testing set. The training set contains 204 observations, from January 2005 until December 2021 and is used for the fitting of the models and the estimation of their parameters. The parameters of the models are estimated by using maximum likelihood estimation in R programming language. The testing set consists of 12 observations from January 2021 until December 2021 and is used for the evaluation of the out-of-sample forecasting accuracy of the models.

As in Munim and Schramm (2017), the selected SARIMA model is not necessarily the model with the lowest Akaike information criterion (AIC) (Akaike, 1974), Bayesian Information Criterion (BIC) or other information criterion. Hyndman and Athanasopoulos (2018) also note that "a model which fits the training data well will not necessarily forecast well". Different combinations of lag orders for the terms of the SARIMA model are investigated and the models with the best forecasting performance are identified. These models are submitted to diagnostic checking in order to specify if they are valid for forecasting. The model with the best forecasting performance that passes the diagnostic checking is selected.

The ARCH/GARCH models are fitted to the residuals of the selected SARIMA model in order to estimate the volatility of the errors of the model. The parameters of the models are estimated by using maximum likelihood estimation in R programming language. The ARCH/GARCH models are selected based on the lowest AIC and BIC given by the following expressions:

$$AIC = 2k - 2\ln(L) \tag{5.1}$$

$$BIC = kln(n) - 2\ln(L) \tag{5.2}$$

where k is the number of parameters estimated by the model, L the maximized value of the likelihood function of the model and n the number of observations.

5.3.1 Mathematical formulation of SARIMA model

In order to arrive at the mathematical formulation of the SARIMA model, the models that form the SARIMA model are first described.

5.3.1.1 Autoregressive Models (AR) (p)

Autoregressive models are based on the assumption that the value of a variable in one time period depends linearly on its own previous values. The mathematical formulation of the AR model is the following:

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \varepsilon_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \tag{5.3}$$

where

X_t is the container throughput at time t

X_{t-i} is the container throughput of all the previous periods until lag p

φ_i is the parameter of X_{t-i}

ε_t is the error term at time t

The back-shift operator B and the characteristic polynomials of the back-shift operator provide a useful notation and permit to express the mathematical equations in a more compact way. The back-shift operator B operates on the elements of the time series and shifts them back one period. In the general case, B^K shifts the data back k periods:

$$B^k X_t = X_{t-k} \tag{5.4}$$

Consequently, in terms of back-shift operator equation (1) can be written as:

$$\left(1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p\right) X_t = \varepsilon_t \tag{5.5}$$

Equation (3) can be written in the compact form:

$$\varphi_p(B) X_t = \varepsilon_t \tag{5.6}$$

Where $\varphi_p(B)$ is the characteristic polynomial of the back-shift operator defined as:

$$\varphi_p(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p \tag{5.7}$$

5.3.1.2 Moving Average Models (MA) (q)

Moving average models are linear regressions of a variable against white noise error terms of previous time periods. The mathematical formulation of the MA model is the following:

$$X_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t = \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \tag{5.8}$$

where

X_t is the container throughput at time t

ε_{t-i} is the error term of all the previous periods until lag q

θ_i is the parameter of ε_{t-i}

μ is the mean of the series

ε_t is the error term at time t

In terms of the back-shift operator B and the characteristic polynomial, equation (6) can be expressed as:

$$X_t - \mu = \theta_q(B)\varepsilon_t \tag{5.9}$$

Where $\theta_q(B)$ is the characteristic polynomial of the back-shift operator defined as:

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \tag{5.10}$$

5.3.1.3 Autoregressive Moving Average Models (ARMA) (p,q)

An ARMA model is a combination of an AR and MA model. The mathematical formulation of the ARMA model is the following:

$$\begin{aligned} X_t &= \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \\ &= \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \end{aligned} \tag{5.11}$$

In terms of the back-shift operator B and the characteristic polynomials, equation (9) can be expressed as:

$$\varphi_p(B) X_t = \theta_q(B)\varepsilon_t \tag{5.12}$$

5.3.1.4 Autoregressive Integrated Moving Average (ARIMA) (p,d,q) Models

When a time series is not stationary, the most common approach is to use differenced values of the variable instead of the values themselves. This procedure, which is the integrated part of the model, transforms an ARMA model to an ARIMA model. Thus, in an ARIMA (p,d,q) model, p denotes the order of the AR terms, d the order of difference and q the order of the MA terms. The back-shift operator notation is useful for describing the process of differencing. The differencing equation of first order can be written as:

$$\Delta X_t = X_t - X_{t-1} = (1 - B)X_t \tag{5.13}$$

Differences of d order can be expressed as:

$$\Delta^d X_t = (1 - B)^d X_t \tag{5.14}$$

As a consequence, the mathematical formulation of the ARIMA model is the following:

5.3.1.5 Seasonal ARIMA Models (SARIMA) (p,d,q) (P,D,Q)$_s$

A SARIMA model is an extension of the ARIMA model which can be applied to seasonal time series data. The SARIMA model is constructed by adding seasonal terms to the ARIMA model. The seasonal part of the model, (P,D,Q)s, is similar to the non-seasonal part but its terms are derived from backshifts of the seasonal period. Therefore, P denotes the order of the seasonal AR terms, D the order of the seasonal difference, Q the order of the seasonal MA terms and s is the number of observations per year. In terms of the back-shift operator B and the characteristic polynomials, the mathematical formulation of the SARIMA model is:

$$\varphi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D X_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t \qquad (5.16)$$

where

X_t is the time series at time t

B is the back-shift operator

$\varphi_p(B)$ is the nonseasonal autoregressive characteristic polynomial of order p defined as $1-\varphi_1 B-\varphi_2 B^2-\cdots-\varphi_p B^p$

$\Phi_P(B^s)$ is the seasonal autoregressive characteristic polynomial of order P defined as

$1-\Phi_1 B^s-\Phi_2 B^{2s}-\cdots-\Phi_P B^{Ps}$

$\theta_q(B)$ is the nonseasonal moving average characteristic polynomial of order q defined as $1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q$

$\Theta_Q(B^s)$ is the seasonal moving average characteristic polynomial of order Q defined as $1-\Theta_1 B^s-\Theta_2 B^{2s}-\cdots-\Theta_Q B^{Qs}$

5.3.2 Mathematical formulation of ARCH/GARCH model

5.3.2.1 Autoregressive Conditional Heteroskedasticity (ARCH) (q) models

While ARMA models are suitable for modelling and predicting future values of time series data, as the conditional mean of these model is time varying, the conditional variance is constant. Consequently, they are not capable of modelling time varying volatility. In 1982, Robert Engle in order to model volatility clustering in financial time series introduced the

ARCH model (Engle, 1982). The mathematical formulation of the ARCH model is the following:

$$\varepsilon_t = z_t \sigma_t \tag{5.17}$$

$$\sigma^2{}_t = \alpha_0 + \alpha_1 \varepsilon^2{}_{t-1} + \alpha_2 \varepsilon^2{}_{t-2} + \cdots + \alpha_q \varepsilon^2{}_{t-q} = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon^2{}_{t-i} \tag{5.18}$$

where

ε_t is the error term at time t

σ_t is the standard deviation of the error term time at time t

z_t is a sequence of iid random variables with zero mean and unit variance

α_i is the parameter of $\varepsilon^2{}_{t-i}$, $\alpha_i \geq 0$

α_0 is the constant, $\alpha_0 > 0$

5.3.2.2 Generalized Autoregressive Conditional Heteroskedasticity (GARCH) (p,q) models

A disadvantage of the ARCH model is that in order to fit adequately financial time series in practice, a large number of q parameters have to be estimated. Tim Bollerslev, a doctoral student of Robert Engle at the time, introduced in 1986 a generalization of the ARCH model, the GARCH model, as a way to fit financial times series effectively while keeping the number of parameters small (Bollerslev, 1986). The mathematical formulation of the GARCH model is the following:

$$\varepsilon_t = z_t \sigma_t \tag{5.19}$$

$$\begin{aligned}\sigma^2{}_t &= \alpha_0 + \alpha_1 \varepsilon^2{}_{t-1} + \alpha_2 \varepsilon^2{}_{t-2} + \cdots + \alpha_q \varepsilon^2{}_{t-q} + \beta_1 \sigma^2{}_{t-1} + \beta_2 \sigma^2{}_{t-2} + \cdots + \beta_p \sigma^2{}_{t-p} \\ &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon^2{}_{t-i} + \sum_{i=1}^{p} \beta_i \sigma^2{}_{t-i} \end{aligned} \tag{5.20}$$

where $\beta_i > 0$ is the parameter of $\sigma^2{}_{t-i}$, $\alpha_0 > 0$, $\alpha_i \geq 0$ and $\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i < 1$.

5.3.3 Prediction intervals

A prediction interval provides an upper and lower limit between which the forecasted value is expected to lie with a specified probability (Chatfield, 2001). Prediction intervals express the uncertainty in the predictions (Hyndman and athanasopoulos, 2018). A 100(1-α) % prediction interval for the h-step forecast expressed as:

$$\widehat{X_n}(h) \pm z_{\frac{a}{2}}\widehat{\sigma_h} \tag{5.21}$$

where $\widehat{X_n}(h)$ is the point forecast of the value at time (n+h) made using the data up to time n, $\widehat{\sigma_h}$ is the standard deviation of the forecast distribution at time (n+h) and $z_{a/2}$ denotes the appropriate (two-tailed) percentage point of a standard normal distribution (Chatfield, 2001). For a 95% prediction interval equation 5-21 can be written as:

$$\widehat{X_n}(h) \pm 1.96\widehat{\sigma_h} \tag{5.22}$$

In order to calculate the prediction interval an estimate of the $\widehat{\sigma_h}$ is needed. This estimate is provided from the ARCH/GARCH model.

5.3.4 Measures for the evaluation of the model's performance

The mean absolute percentage error (MAPE) is used in order to assess the accuracy of the point forecasts. The MAPE is defined as:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|1 - \frac{\widehat{X_t}}{X_t}\right| \times 100 \tag{5.23}$$

where, in period t, X_t is the actual value, $\widehat{X_t}$ the predicted value and n the number of observations.

The mean relative interval width (MRIW), the mean absolute interval width (MAIW) and the containing ratio (CR) are used in order to evaluate the performance of the prediction intervals. These measures are defined as:

$$MRIW = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{X_t^u - X_t^l}{X_t}\right) \tag{5.24}$$

$$MAIW = \frac{1}{n}\sum_{t=1}^{n}\left(X_t^u - X_t^l\right) \tag{5.25}$$

$$CR = \frac{1}{n}\sum_{t=1}^{n} M_t \times 100 \tag{5.26}$$

Where, X_t^u is the upper bound of the predicted value and X_t^l the lower bound. M_t is the number of actual values enveloped by its forecasted upper and lower bounds and it is calculated by:

$$M_t = \begin{cases} 1, & if\ X_t^l \leq X_t \leq X_t^u \\ 0, & else \end{cases} \tag{5.27}$$

5.4 Application of the methodology and results

The methodology described in the previous section, is applied in order to produce point forecasts and prediction intervals for the container throughput of the port of Barcelona. Figure 5-2 presents the monthly time series data of total container throughput (TCT) of the port of Barcelona from 2005-2021.

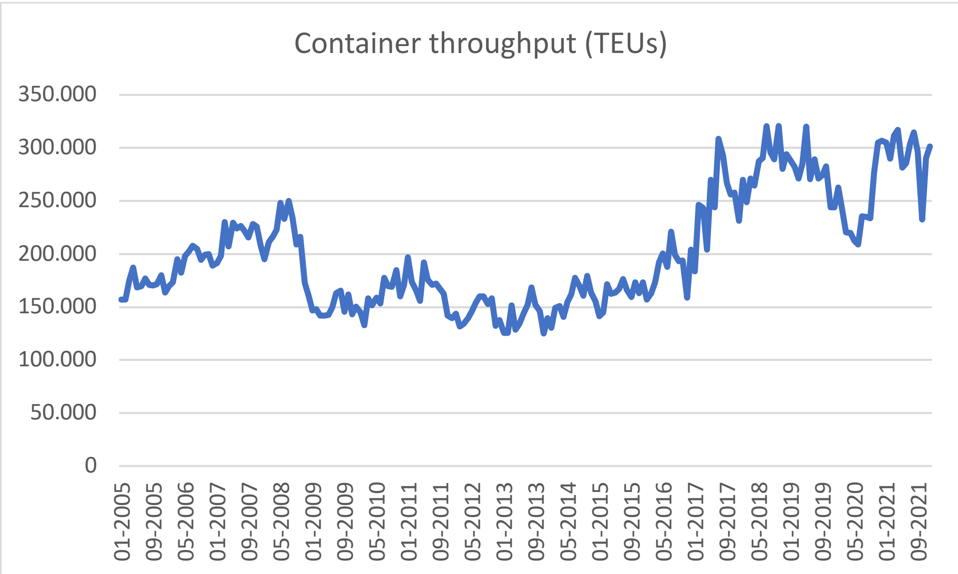

Figure 5-2: Monthly time series data of total container throughput of the port of Barcelona from 2005-2021.
Source: Created by the author based on data from Puertos del Estado.

A time series is stationary when the statistical properties of a process generating the time series do not change over time. As it is important for ARMA models to be applied at stationary time series, the series is checked for stationarity. The time series is tested with the Augmented Dickey-Fuller (ADF), the Phillips-Perron (PP) and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) unit root tests. The p-values of the ADF and PP tests are found larger than 0.05 and therefore the null hypothesis that the data is non-stationary cannot be rejected. This conclusion is also confirmed from the KPSS test as its null hypothesis that the data is stationary is rejected (p-value < 0.05). In order to achieve stationarity an appropriate transformation is applied to the time series data. First lag differencing is applied to the log time series (figure 5-3).

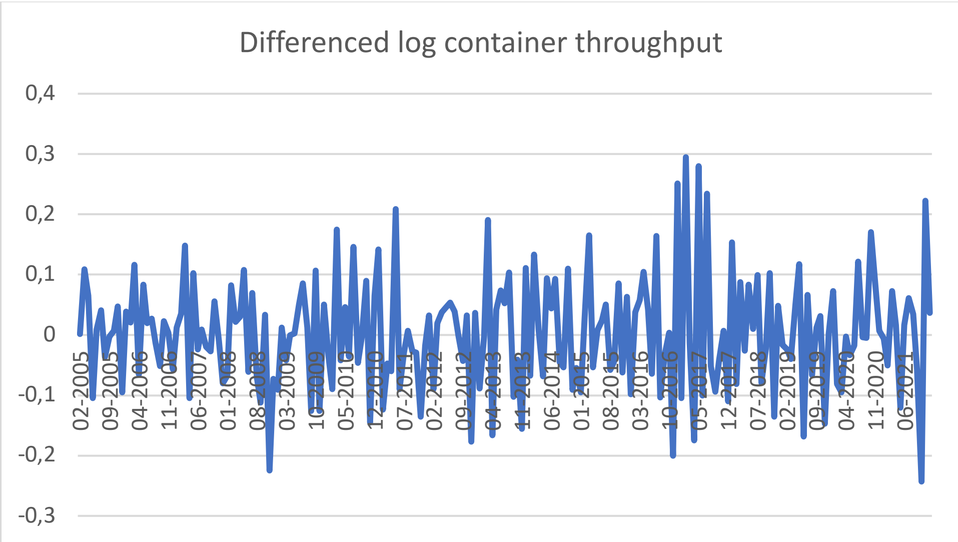

Figure 5-3: Monthly differenced log time series data of total container throughput of the port of Barcelona from 2005-2021.
Source: Created by the author based on data from Puertos del Estado.

The new time series (Δ_1LogTCT) is tested again for stationarity. The results of the ADF test (p-value < 0.05) and of the KPSS test (p-value > 0.05) prove that the transformed time series is stationary (table 5-1).

Table 5-1: Results of unit root tests for the original and the transformed time series data of Barcelona's port.

	TCT		Δ_1LogTCT	
Test	Statistic	p-value	Statitsic	p-value
ADF	-1.9946	0.5784	-6.3144	<0.01
PP	-11.575	0.4544	-262.19	<0.01
KPSS	1.4945	<0.01	0.08885	>0.1

Source: Elaborated by the author.

Subsequently, the appropriate SARIMA model is selected. The forecasting performance of all the models of the form $(p,1,q)(P,0,Q)_{12}$ combining $0 \leq p,q,P,Q \leq 2$ is evaluated based on the MAPE of the testing set. The maximum lags of the seasonal and non-seasonal AR and MA terms is 2 in order to avoid overfitting the models to the time series data. Overfitted models with large number of parameters do not necessarily perform well in forecasting (Chatfield, 2001). The model with the best forecasting performance is the SARIMA $(0,1,1)(2,0,0)_{12}$ model. The model is fitted to the training set and its parameters are estimated (table 5-2).

Table 5-2: Estimations of SARIMA (0,1,1)(2,0,0)$_{12}$ model's parameters.

Parameters	MA1	SAR1	SAR2
Estimate	-0.3232	0.2653	0.1462
Std Error	0.0643	0.0729	0.0739

Source: Elaborated by the author.

The next step includes the diagnostic checking of the selected model. The residuals of the model are inspected visually (figure 5-4) and with formal statistical tests. The plot of the autocorrelation function (ACF) of the model's residuals suggestes that there is no autocorrelation since the spikes at almost all lags are between the crucial limits.

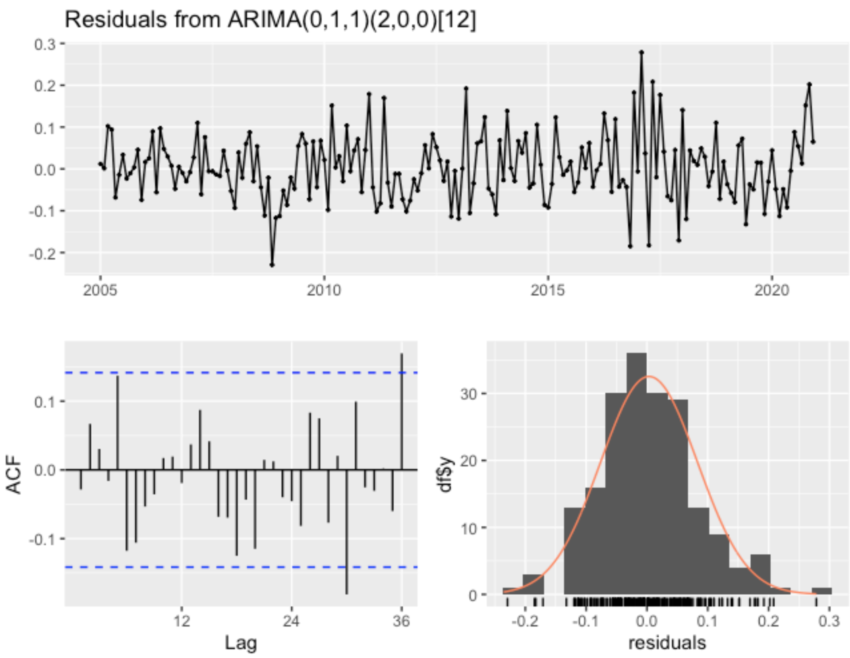

Figure 5-4: Residual diagnostics of SARIMA (0,1,1)(2,0,0)$_{12}$ model.
Source: Retrieved by the author based on the input data.

The autocorrelation of the residuals of the model is also tested with the Ljung-Box test (table 5-3). The p-value of the Ljung-Box test is found larger than 0.05. Consequently, the null hypothesis that the there is no autocorrelation in the model's residuals is accepted. Therefore, the residuals of the model are not autocorrelated and the model is appropriate for forecasting.

Table 5-3: Results of the autocorrelation and normality tests at SARIMA (0,1,1)(2,0,0)$_{12}$ residuals.

Test	Statistic	p-value
Ljung-Box	22.843	0.3524
Shapiro-Wilk	0.98811	0.1086

Source: Retrieved by the author based on the input data.

The normality of the model's residuals is tested with the Shapiro-Wilk test (table 5-3). The result of the test (p-value > 0.05) suggestes that the residuals of the model are normally distributed. The normality of the residuals is additionally checked with a Q-Q plot (figure 5-5). The plot confirmes the finding of the test as most of the points are on the theoretical line.

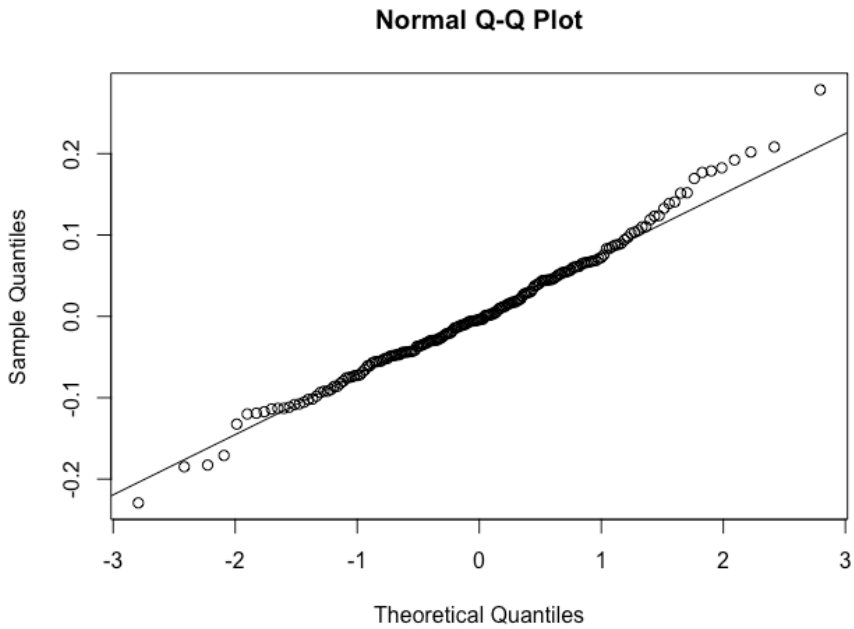

Figure 5-5: Normal Q-Q plot of the residuals of SARIMA (0,1,1)(2,0,0)$_{12}$ model.
Source: Retrieved by the author based on the input data.

Finally, the residuals of the model are checked for the presence of ARCH effects. The volatility clustering that is observed in figure 5-3 is a strong indication for the presence of ARCH effects. This assumption is also checked with a formal statistical test. Figure 5-6 presents the results of the McLeod-Li test. As the p-values for all lags (apart from the first) are < 0.05 the null hypothesis that there are no ARCH effects is rejected. Consequently, the variance of the errors of the model is time varying and can be modelled with an appropriate ARCH or GARCH model.

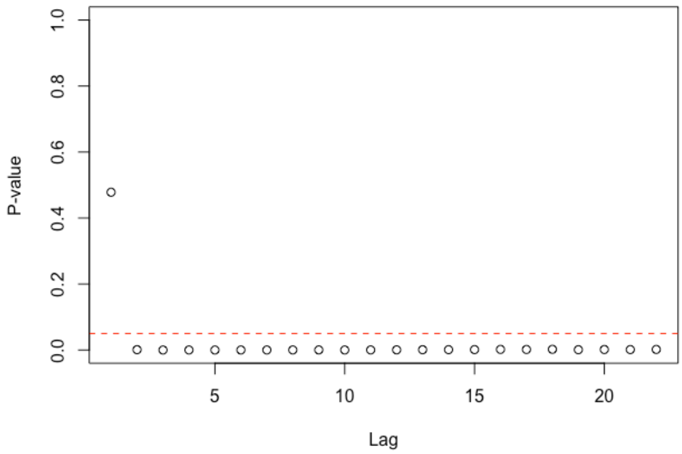

Figure 5-6: McLeod-Li test for the detection of ARCH effects at the residuals of SARIMA (0,1,1)(2,0,0)$_{12}$ model.
Source: Retrieved by the author based on the input data.

The candidate ARCH and GARCH models are fitted to the residuals of the SARIMA model in order to model their conditional variance. All the combinations of the p and q terms of the GARCH (p,q) models are tested until the maximum order of 2. The ARCH(q) models are tested until the lag q that the AIC criterion reaches a local minimum. The selected model is the ARCH (2) model, which is the model with the lowest AIC and BIC (table 5-4). The estimated parameters of the model are presented in table 5-5.

Table 5-4: AIC and BIC of the fitted ARCH/GARCH models.

Model	AIC	BIC
GARCH(1,1)	-2.2412	-2.1733
GARCH(2,1)	-2.2301	-2.1452
GARCH(2,2)	-2.2346	-2.1428
GARCH(1,2)	-2.2450	-2.1602
ARCH(1)	-2.1982	-2.1473
ARCH(2)	-2.2505	-2.1827
ARCH(3)	-2.2461	-2.1612

Source: Retrieved by the author based on the input data.

Table 5-5: Estimations of ARCH(2) model's parameters.

Parameters	ω	α_1	α_2
Estimate	0.004042	0.118919	0.259867
Std Error	0.000000	0.019941	0.006663

Source: Retrieved by the author based on the input data.

After the ARCH(2) model has been fit to the residuals of the SARIMA model the last step is to check the adequacy of the fit. If the combined model has been specified correctly the standardized residuals $\frac{\varepsilon_t}{\sigma_t}$ should not present any autocorrelation or ARCH effects and the normal q-q plot should look approximately linear. The ACF plot of the standardized residuals (figure 5-7) reveals that they do not present any autocorrelation as there is no exceedance of the boundaries at any lag. This result is also confirmed from the Ljung-Box test (table 5-6).

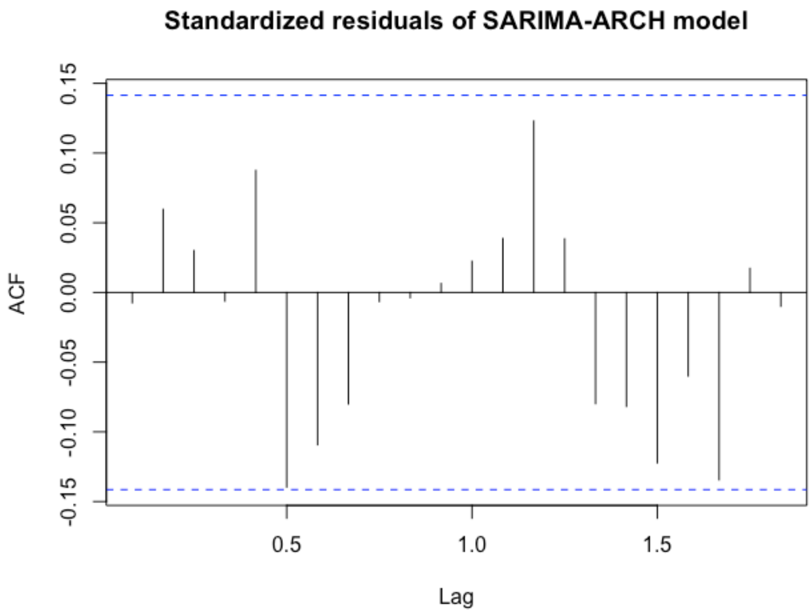

Figure 5-7: ACF plot of the standardized residuals of SARIMA (0,1,1)(2,0,0)$_{12}$ – ARCH(2) model.
Source: Retrieved by the author based on the input data.

The McLeod-Li test performed at the standardized residuals of the combined SARIMA-ARCH model certifies that the ARCH effects that are present at the residuals of the SARIMA model are removed as the p-values for all lags are > 0.05 (figure 5-8).

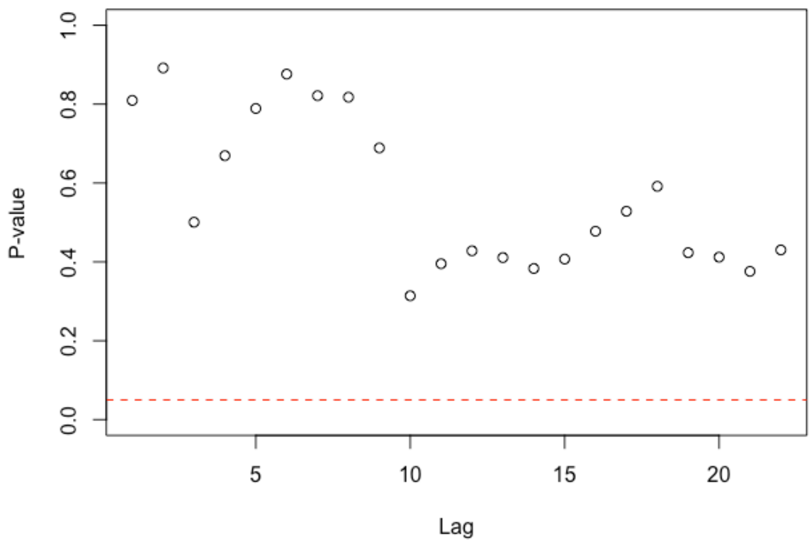

Figure 5-8: McLeod-Li test for the detection of ARCH effects at the standardized residuals of SARIMA (0,1,1)(2,0,0)$_{12}$ - ARCH(2) model.
Source: Retrieved by the author based on the input data.

As the innovations of the ARCH model have been specified to follow a normal distribution the normal q-q plot should be close to linearity which is confirmed form figure 5-9. This fact is also verified from the Shapiro-Wilk test (table 5-6).

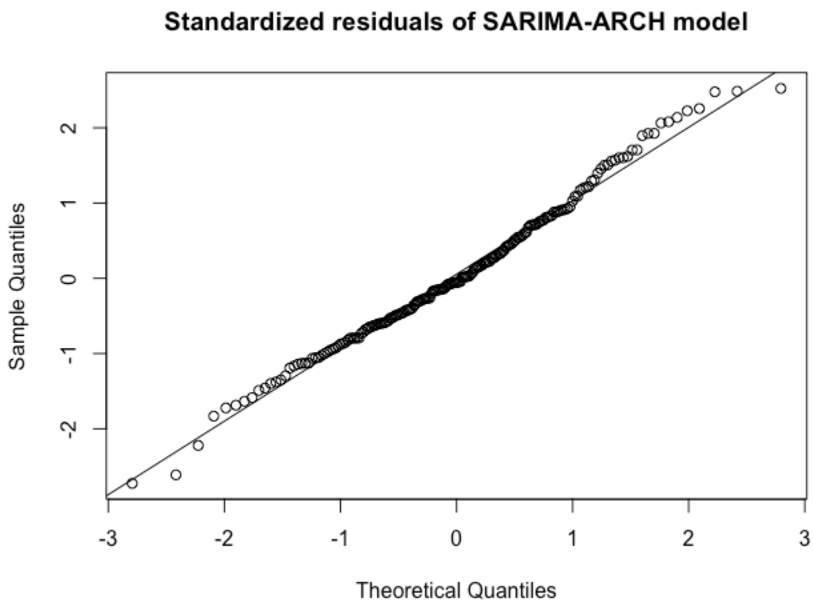

Figure 5-9: Normal Q-Q plot of the standardized residuals of SARIMA (0,1,1)(2,0,0)$_{12}$ - ARCH(2) model.
Source: Retrieved by the author based on the input data.

Table 5-6: Results of the autocorrelation and normality tests at standardized residuals of SARIMA-ARCH model.

Test	Statistic	p-value
Ljung-Box	24.893	0.4116
Shapiro-Wilk	0.98978	0.1875

Source: Retrieved by the author based on the input data.

Figure 5-10 presents the conditional standard deviation estimated with the ARCH (2) model for the port of Barcelona. It is observed that the peaks of the conditional variance correspond to the periods of large fluctuations of the container throughput, while at periods where the fluctuations of container throughput are relatively mild, the conditional standard deviation is also small and mild. These results confirm that the ARCH model has captured well the volatility of the monthly container throughput of the port of Barcelona.

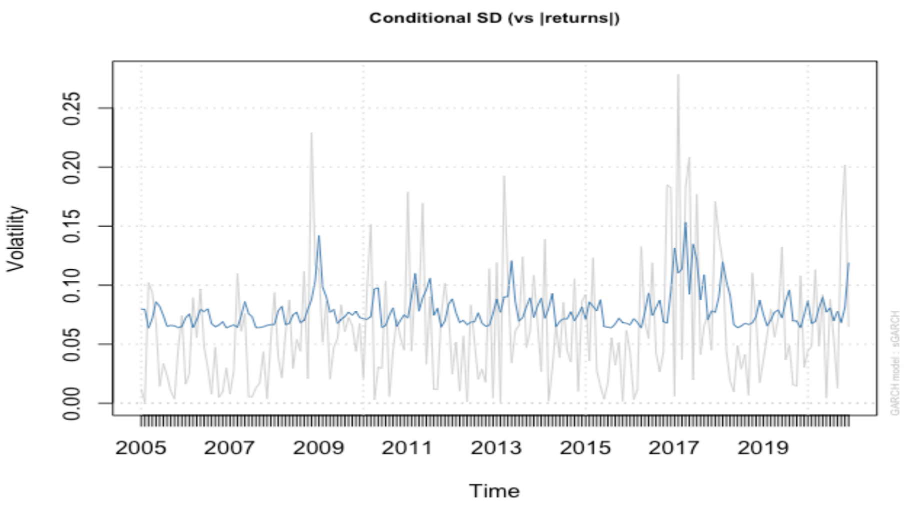

Figure 5-10: Plot of the conditional volatility of the time series against absolute values of the log differenced time series.
Source: Retrieved by the author based on the input data.

Finally, the mathematical equations of the combined SARIMA - ARCH model are the following:

$$(1 - 0.2653B^{12} - 0.1462B^{24})(1 - B)X_t = (1 + 0.3232B)\varepsilon_t \qquad (5.28)$$

$$\sigma^2{}_t = 0.004042 + 0.118919\varepsilon^2{}_{t-1} + 0.259867\varepsilon^2{}_{t-2} \qquad (5.29)$$

5.5 Forecasting evaluation of SARIMA - ARCH model

The selected SARIMA (0,1,1) $(2,0,0)_{12}$ – ARCH (2) model is used in order to produce 12 out-of-sample point forecasts and their prediction intervals. The MAPE of the point forecasts of this testing set is 6.3%. The value of MAPE is heavily influenced from the presence of an outlier at the testing set (October-2021). Such extreme values are very difficult to be predicted from SARIMA models. The MAPE of the point forecasts if the outlier is replaced by the average value of the testing set is 3.82%.

The produced prediction intervals of the SARIMA-ARCH model are compared with these generated from a simple SARIMA model. The variances of the forecast errors, needed for the construction of the prediction intervals of the SARIMA model, are generally obtained using the coefficients of the MA representation of the SARIMA model. The variances of the forecast errors of the SARIMA model are computed in R programming language by applying a Kalman filter at the fitted model and then scaled by the estimated variance of the residuals. The prediction intervals of the fitted values of the container throughput (training set) generated by the SARIMA-ARCH model are a little narrower from these of the simple SARIMA model and have a little higher containing ratio (Figure 5-11, table 5-7).

Table 5-7: Evaluation measures of the prediction intervals of SARIMA and SARIMA-ARCH models.

	Training			Testing		
Model	CR (%)	MAIW	MRIW	CR (%)	MAIW	MRIW
SARIMA	93.75	61,641	0.316	100	174,353	0.601
SARIMA-ARCH	94.27	60,485	0.31	91.66	100,672	0.344

CR: Containing Ratio; **MAIW**: Mean Absolute Interval Width (TEUs); **MRIW**: Mean Relative Interval Width. Source: Calculated by the author.

The prediction intervals of the point forecasts (testing set) present a big difference in their average width (figure 5-12). The MAIW (100,672 TEUs) and the MRIW (0.344) of the SARIMA-ARCH model are much smaller than those of the SARIMA model (174,353 TEUs and 0.601 respectively) (table 5-7). Consequently, the SARIMA-ARCH model produces much more narrow prediction intervals compared with these of the simple SARIMA. The CR of the testing set of the SARIMA model is 100%, meaning all 12 actual values are inside the predicted limits and of SARIMA-ARCH model is 91.66% meaning that 11 out of 12 actual values are enveloped by their predicted upper and lower bounds. Only the large spike in October of 2021 exceeds the 95% prediction intervals of the combined model. As a result, the combined model generates much narrower prediction intervals, which are more useful for terminal operators and port authorities, than the simple model and yet retains a high degree of containing ratio.

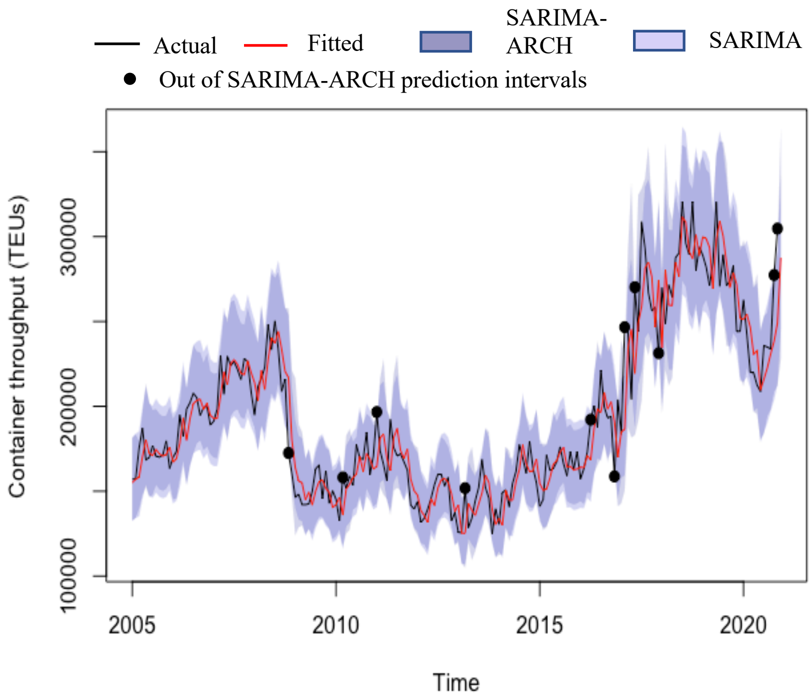

Figure 5-11: Plot of time series (training set) and fitted values with their 95% prediction intervals of SARIMA and SARIMA-ARCH model.
Source: visualized by the author.

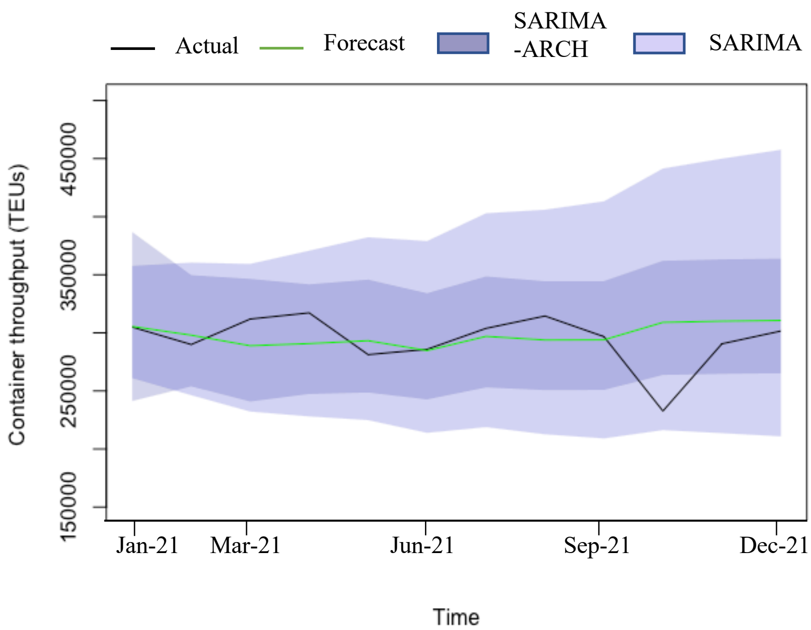

Figure 5-12: Plot of time series (testing set) and point forecasts with their 95% prediction intervals of SARIMA and SARIMA-ARCH model.
Source: visualized by the author.

5.6 Conclusions

This study proposes a combined SARIMA-ARCH/GARCH model for short-term forecasts of container throughput at ports. The SARIMA-ARCH/GARCH model is a practical choice for container throughput forecasting. As it is a univariate time series model, the time series of container throughput are the only data needed for the forecasts. This is an advantage compared to multivariate time series models that demand the collection of data for a number of exogenous variables, which is a time-consuming task and maybe impossible if the data are not available. Furthermore, this model offers reduced computational complexity compared to artificial intelligence and machine learning models. The accuracy of the model's short-term forecasts is satisfactory. Therefore, the model offers a good balance between simplicity and forecasting accuracy.

The added value of the combined SARIMA-ARCH/GARCH model is that it provides not only point forecasts of the container throughput but also their accompanying prediction intervals. In this way, the uncertainty of each point forecast is estimated. The SARIMA part of the model produces point forecasts of the container throughput. The ARCH/GARCH part, is a model specifically designed for capturing and forecasting the volatility of a time series. Volatility forecasts allow the construction of the prediction intervals of the point forecasts.

The proposed model is applied to monthly time series of container throughput data of the port of Barcelona, in order to produce short-term forecasts and their prediction intervals. The MAPE of a 12-month forecasting horizon is 6.3% mainly enlarged by the presence of a large spike at the testing set. The proposed combined model produces much narrower out-of-sample prediction intervals compared to these of the SARIMA model without the inclusion of the ARCH/GARCH part. The prediction intervals of the simple SARIMA model tend to be very large as the forecasting horizon grows that almost lose any practical value. These narrower prediction intervals of the combined model contain all the actual container throughput values of the testing set except from one very large spike. Therefore, these prediction intervals are narrow enough in order to have practical value for terminal operators and port authorities but also maintain a high containing ratio of the actual values.

Some recommendations for future research on the application of the SARIMA-ARCH/GARCH model for forecasting container throughput volumes and their prediction intervals can be proposed. Firstly, the methodology can be tested to other ports or port regions or ranges.

Secondly, the model at this study is applied to the total monthly container throughput of the port of Barcelona. A disaggregated approach could be adopted by specifying the model independently for import, export and transshipment container flows as these types of flows present different characteristics. Thirdly, at the specifications of the ARCH/GARCH model other distributions of the innovations of the model could be used apart from the normal (e.g., student-t). Fourthly, extensions of the GARCH model that are designed for capturing asymmetric effects on conditional volatility like EGARCH, TGARCH or APARCH could be applied. Finally, it would be an improvement if the candidate models were evaluated through a process of cross validation, by using rolling training and testing sets throughout the available data rather than using one static training and testing set.

Chapter 6

Conclusions and further research

This chapter presents the main findings on the research questions that had been set and proposes some directions for future research.

6.1 Main findings and conclusions

Research question 1: Which are the developments in the container throughput of the European port system?

- Most of the European container port ranges experienced a concentration trend until the global financial crisis of 2009 which converted to deconcentration in the following decade. This fact indicates that shipping lines reorganized their shipping networks after the crisis and selected to call in a smaller number of ports.

- The multi-port gateway regions of the European container port system followed different growth patterns.

- All the multi-port gateway regions, apart from the North Adriatic, were highly concentrated as one or two ports in each region collected the largest amount of the container flows.

- European container ports experienced a period of high growth until the pre-crisis year 2007. During the global financial crisis of 2009, they recorded heavy losses but managed to recover and in 2020 most of the leading ports showed very large growth rates compared to the pre- crisis year 2007. The health crisis of 2020 had mild effects on European container ports compared to the crisis of 2009 as it did not unfold at the same time and speed all over the world. The large cumulative market shares of the top ports throughout the examined period indicate that container traffic was concentrated in a few large ports throughout the period 2001-2020.

Research question 2: How was hierarchy in the European container port system evolved?

In order to answer this question 3 methods different than the commonly used concentration indicators were applied. Firstly, two types of rank-size models, secondly, a methodology based on the principals of Analytic Hierarchy Process (AHP) combined with an allometric growth model and finally a Markov chain modeling. The main findings of the application of these methodologies are the following:

- The global financial crisis of 2009 was a crucial turning point for the container traffic concentration patterns of the European port system, shifting from deconcentration in the decade prior the crisis to concentration in the decade after the crisis.

- There is an abundance of medium-sized ports in the European container port system. These medium-sized ports are favored when a concentration trend prevails in the port system and on the other hand, they lose ground when deconcentration occurs.

- The group of ports ranked from position 21 to 40 in the hierarchy of the European container port system experienced the lowest growth rate compared to the other port groups from 2001-2020. This finding suggests that there is a growth trap for medium-sized ports in the European port system.

- The upper part of the small ports, ranked from position 41 to 60 was the group with the highest relative growth.

- The group of the largest ports, positioned from 1 to 20, had the second largest growth rate related to the growth of all other port groups, meaning that the dominance of the largest ports remained unchallenged.

- The group of the largest ports presented the higher degree of immobility followed by the group of medium-sized ports. This finding agrees with the aforementioned findings which suggest that the positions of the largest ports in the European container port system are not challenged by the medium-sized ports.

- The mobility was higher at the bottom parts of the distribution and especially at the group containing the ports ranked from position 61-80.

- Downward mobility was much higher than upward mobility for all port groups suggesting that it is more likely for a port to lose ground than to climb in the hierarchy.

- The transition one group up becomes much slower as the size group increases.

- A medium-sized port needs on average 127 years to enter for the first time the group of the largest ports compared to 30 years to move one group down.

Research question 3: How are the liner shipping networks of individual shipping lines structured?

This thesis attempts to answer this research question by analyzing and visualizing the container shipping networks of Maersk and COSCO shipping prior and after the global financial crisis of 2009. Network indicators are calculated on the theoretical basis provided from graph theory, complex networks and statistical techniques. The main findings of this analysis are the following:

- Maersk operated an extended global network which was shrunk after the global financial crisis of 2009. On the other hand, COSCO's network was much smaller and denser and retained its size after the crisis. Its ports had on average more connections and received more weekly calls than those in Maersk's network throughout the examined period.

- In the visualized networks of both shipping lines', before and after the crisis, communities of ports are detected, formed by ports in geographical proximity.

- Different ports emerged at the top of the hierarchy of each company, based on network indicators, revealing that shipping lines follow different strategies in port selection and the formation of their networks.

- COSCO’s most frequently visited ports, except Singapore, were placed in North East Asia while Maersk visited frequently ports in various regions all over the world.

- The most frequent inter-port links, in both companies’ networks, were between ports of the same region (intra-regional links) throughout the examined period. But while Maersk’s most frequent links were spread all over the world, almost all of COSCO’s frequent links belonged to North East Asia region.

- The global character of Maersk’s network is confirmed by the analysis of the regional networks of the shipping lines, while COSCO was mainly operating in the East-West axis. The ports of North East Asia received the largest number of weekly calls from both shipping lines before and after the financial crisis. The region experienced a huge increase in the weekly calls at its ports from both companies in the pre-crisis period and a small decrease after the crisis. Central America and the Caribbean was the region that was most centrally placed and had the largest connectivity with the other regions of the network, in Maersk’s network while the same happened with South East Asia in COSCO’s network.

- The networks of both companies throughout the examined period can be characterized as scale-free networks indicating the existence of very few ports with a large number of connections and many ports with a small number of connections.

Research question 4: How to produce forecasts of container throughput and their accompanying prediction intervals?

This thesis addresses this question by proposing a combined SARIMA-ARCH/GARCH model and applies it in order to produce short-term forecasts for the container throughput of the port of Barcelona. The main findings of this application are the following:

- The accuracy of the model’s point forecasts produced by the SARIMA part of the model is satisfactory.

- The ARCH/GARCH part of the model is used in order to model and forecast the volatility of the container throughput volumes. The ARCH/GARCH model captures very well the volatility of the container throughput time series data.

- The forecasts of volatility allow the construction of prediction intervals for the point forecasts. The prediction intervals of the combined model are much narrower compared to these of the simple SARIMA model and yet retain a high containing ratio of the actual values of container throughput.

- The model offers a good balance between simplicity and forecasting accuracy, as the time series of container throughput are the only data needed for the forecasts.

- The prediction intervals produced by the model are narrow enough in order to have practical value for terminal operators and port authorities maintaining at the same time a high level of accuracy as indicated by their containing ratio.

6.2 Future research

Concerning the research on the hierarchy dynamics of the European container port system the proposed methodologies could be applied to other port systems in order to obtain some comparative results. The methodology based on the principals of Analytic Hierarchy Process combined with an allometric growth model was applied to groups of ports based on their size in order to assess the impact of size categories of ports one upon another. A direction for future research could be the application of the methodology to individual ports of a multi-port gateway region in order to compare the relative growth of all the neighboring ports against each other. This research focuses on the container throughput of European ports. The aforementioned methodology could be applied to other variables as it provides a general framework in order to weight the relative growth of a variable of a port or port group against all the other concerned ports or port groups.

As far as the analysis of the container shipping networks of individual shipping lines is concerned, this thesis analyzes the networks of Maersk and COSCO Shipping. The research could be expanded in order to include other shipping lines. Another direction of possible future research could be the analysis of the networks of the strategic alliances of shipping lines in order to reveal their strategies in the formation of container shipping networks. This research examines the evolution of shipping networks prior and after a global financial crisis. The time frame of the research could be extended in order to include the health crisis of 2020 in order to compare the results of different types of global crises on the configuration of shipping lines' networks. In this thesis, the inter-ports links are weighted by the frequency of the liner services measured in weekly calls in order to understand better the time proximities between ports. Other weights could be used as the transport capacity of the vessels in TEUs, which has been used in most studies of network analysis.

Some recommendations for future research on the application of the SARIMA-ARCH/GARCH model for forecasting container throughput volumes and their prediction intervals can be proposed. One possible direction is to the apply the methodology to other ports or port groups. Another approach could be to specify the model independently for import, export and transshipment container flows as these types of flows present different characteristics. In addition, extensions of the GARCH model that are designed for capturing asymmetric effects on conditional volatility like EGARCH, TGARCH or APARCH could be tested. A further improvement in the methodology would be the evaluation of the candidates models through a process of cross validation, by using rolling training and testing sets throughout the available data rather than using one static training and testing set. Finally, at the specifications of the ARCH/GARCH model other distributions of the innovations of the model could be used apart from the normal (e.g., student-t).

References

Akaike, H. (*1974*). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control,* 19, 716-723.

Alexis, E., Plocoste, T., Nuiro, S.P. (2022). Analysis of Particulate Matter (PM10) Behavior in the Caribbean Area Using a Coupled *SARIMA-GARCH* Model. *Atmosphere*, 13 (6), 862. https://doi.org/10.3390/atmos13060862

Alphaliner. (2021). Top-100 operated fleets as per 11 January 2021. Retrieved January 21, 2021, from https://alphaliner.axsmarine.com/PublicTop100/.

Association of Italian ports. (2021). Annual Statistics. Retrieved May 15, 2021, from https://www.assoporti.it/en/autoritasistemaportuale/statistiche/statistiche-annuali-complessive/

Barke, M. (1986). *Transportation and Trade: Conceptual Frameworks in Geography*. Edimburgh.

Black, D., V. Henderson. (2003). Urban evolution in the USA. *Journal of Economic Geography,* 3, 343-372. https://doi.org/10.1093/jeg/lbg017

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307-327.

Broad, L. R., (1998). Allometry and Growth. Forest Science, 44(3), 458 - 464. https://doi.org/10.1093/forestscience/44.3.458

Calatayud, A., Mangan, J. Palacin, R. (2017a). Connectivity to international markets: a multilayered network approach. *Journal of Transport Geography*, 61, 61-71. https://doi.org/10.1016/j.jtrangeo.2017.04.006

Calatayud, A., Mangan, J., Palacin, R. (2017b). Vulnerability of international freight flows to shipping network disruptions: A multiplex network perspective. *Transportation Research Part E: Logistics and Transportation Review,* 108, 195-208. https://doi.org/10.1016/j.tre.2017.10.015

Chan, H. K., Xu, S., Qi, X., (2019). A comparison of time series methods for forecasting container throughput. *International Journal of Logistics Research and Applications*, 22 (3), 294-303. https://doi.org/10.1080/13675567.2018.1525342

Chatfield, C. (2001). Prediction Intervals for Time-Series Forecasting. In: Armstrong, J.S. (Ed.), *Principles of Forecasting*. International Series in Operations Research & Management Science, vol 30. Springer, Boston, MA. https://doi.org/10.1007/978-0-306-47630-3_21

Chen, J, Fei, Y, Wan, Z, Yang, Z, Li, H, Choi, K-S and Xie, X. (2019). Allometric relationship and development potential comparison of ports in a regional cluster: A case study of ports in the Pearl River Delta in China. *Transport Policy,* 85, 80-90. https://doi.org/10.1016/j.tranpol.2019.10.018

Chen, S.H., Chen, J.N. (2010). Forecasting container throughputs at ports using genetic programming**.** *Expect Systems with Applications,* 37(3), 2054-2058. https://doi.org/10.1016/j.eswa.2009.06.054

Chou, C. C., Chu, C. W., Liang, G. S. (2008). A modified regression model for forecasting the volumes of Taiwan's import containers. *Mathematical and Computer Modelling*, 47 (9), 797-807. https://doi.org/10.1016/j.mcm.2007.05.005

Cisic, D., Komadina, P., Hlača B. (2007). Network Analysis of the Mediterranean port supply chain structures. *Pomorstvo* 21(1), 211-220.

CI - Containerisation International (2002, 2008, 2011). Containerisation International Yearbooks, Informa, London.

Cullinane, K., Wang, Y. (2009). A capacity-based measure of container port accessibility. *International Journal of Logistics: Research and Applications*, 12 (2), 103-117. https://doi.org/10.1080/13675560902749340

Cullinane, K., Wang, Y.-H. (2012). The hierarchical configuration of the container port industry: an application of multiple linkage analysis. *Maritime Policy and Management*, 39 (2), 169-187. https://doi.org/10.1080/03088839.2011.650719

David, M., Ramahatana, F., Trombe, P.J., Lauret, P. (2016). Probabilistic forecasting of the solar irradiance with recursive ARMA and GARCH models. *Solar Energy* 113, 55-72. https://doi.org/10.1016/j.solener.2016.03.064

Deng, W.B., Long, G., Wei, L., Xu, C., (2009). Worldwide marine transportation network efficiency and container throughput. *Chinese Physics Letters*, 26 (11), 118901. https://doi.org/10.1088/0256-307X/26/11/118901

Ding, M., Zhang, S., Zhong, H., Wu, Y., Zhang, L. (2019). A prediction model of the sum container based on combined BP Neural Network and SVM. *Journal of Information Processing Systems*, 15 (2), 305-319. https://doi.org/10.3745/JIPS.04.0107

Dobkins, L. and Ioannides, Y. M. (2000). Dynamic evolution of the US city size distribution. In: Huriot, J. M., J. F. Thisse (Eds.), *The Economics of Cities*. Cambridge, 217-260.

Dragan, D., Kramberger, T., Intihar, M. (2014). A comparison of Methods for Forecasting the Container Throughput in North Adriatic Ports. In *IAME 2014 Conference*, Norfolk VA USA

Drewry. (2007-2009). Container market-Annual review and forecast. London: Drewry Shipping Consultants.

Ducruet, C. (2013a). Network diversity and maritime flows. *Journal of Transport Geography*, 30, 77-88. https://doi.org/10.1016/j.jtrangeo.2013.03.004

Ducruet, C. (2013b). Ports in a World Maritime System: A Multilevel Analysis. In: Rozenblat, C., Melançon, G. (Eds.), *Methods for Multilevel Analysis and Visualisation of Geographical Networks*. Methodos Series, vol 11. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-6677-8_8

Ducruet, C., Ietri, D., Rozenblat, C. (2011a). Cities in worldwide air and sea flows: A multiple networks analysis. *Cybergeo: European Journal of Geography* 528. http://cybergeo.revues.org/23603

Ducruet, C., Lee, S.W., Ng, A.K. (2010b). Centrality and vulnerability in liner shipping networks: Revisiting the Northeast Asian port hierarchy. *Maritime Policy and Management* 37 (1), 17-36. https://doi.org/10.1080/03088830903461175

Ducruet, C., Lee, S.W., Roussin, S. (2009b). Local strength and global weakness: A maritime network perspective on South Korea as Northeast Asia's logistics hub. *International Journal of Maritime Affairs and Fisheries*, 32-50.

Ducruet, C., Lee, S.W., Song, J.M. (2011b). Network Position and Throughput Performance of Seaports. In: Notteboom, T.E. *Current Issues in Shipping, Ports and Logistics*. Brussels.

Ducruet, C., Lee, S.-W., Ng, A. K. (2011c). Port competition and network polarization in the East Asian maritime corridor. *Territoire en mouvement Revue de Géographie et Aménagement*, 10, 60-74. https://doi.org/10.4000/tem.1327

Ducruet C., Notteboom T.E. (2012a). Developing liner service networks in container shipping. In: Song, D.W, Panayides, P., (eds). *Maritime logistics: A guide to contemporary shipping and port management.* London: Kogan, 77–100.

Ducruet, C., y Notteboom, T.E. (2012b). The worldwide maritime network of container shipping: spatial structure and regional dynamics, *Global Networks,* 12 (3), 395-423

Ducruet, C., Notteboom, T.E. (2022). Revisiting port system delineation through and analysis of maritime interdependencies among seaport. *GeoJournal*, 87 (22), 1-29. https://doi.org/ 10.1007/s10708-020-10341-x

Ducruet, C., Notteboom T., Banos A., Ietri D., Rozenblat C., (2010a). Structure and dynamics of liner shipping networks. In: Proceedings of the 2010 Annual Conference of the International Association of Maritime Economists (IAME2010), Lisbon.

Ducruet, C., Roussin, S., Jo, J-C. (2009a). Going west? Spatial polarization of the North Korean port system. *Journal of Transport Geography*, 17 (5), 357-368. https://doi.org/10.1016/j.jtrangeo.2008.10.011

Ducruet C., Rozenblat C., Zaidi F. (2010c). Ports in Multi-level Maritime Networks: Evidence from the Atlantic (1996-2006). Journal of Transport Geography, 18 (4), 508-518. https://doi.org/10.1016/j.jtrangeo.2010.03.005

Ducruet C., Zaidi F. (2012). Maritime constellations: A complex network approach to shipping and ports. *Maritime Policy and Management*, 39 (2), 151-168. https://doi.org/10.1080/03088839.2011.650718

Eaton, J., Eckstein, Z. (*1997*). Cities and Growth: Theory and Evidence from France and Japan. *Regional Science and Urban Economics*, 27, 443-474. https://doi.org/10.1016/S0166-0462(97)80005-1

Ena, M., Tri Herdiani, E., Tinungki, G. M. (2022). The application of the Arima – Garch method in predicting the Indosat stock price. *International Research Journal of Advanced Engineering and Science*, 7 (3), 262-265.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 987-1007.

Eurostat. (2021). Volume of containers transported to/from main ports - quarterly data. Retrieved May 15, 2021, from https://ec.europa.eu/eurostat/databrowser/explore/all/transp?lang=en&subtheme=mar.mar_go.mar_go_qm_cont&display=list&sort=category

Farhan, J., Ong, G. P. (*2018)*. Forecasting seasonal container throughput at international ports using SARIMA models. *Maritime Economics and Logistics*, 20 (1), 131-148. https://doi.org/10.1057/mel.2016.13

Finnish port association. (2021). Monthly statistics. Retrieved May 15, 2021, from https://www.finnishports.fi/eng/statistics/monthly-statistics/?stats=monthly&T=0

Freire Seoane, M.J., González- Laxe, F., Pais- Montes, C. (2013). Foreland determination for containership and general cargo ports in Europe (2007–2011). *Journal of Transport Geography*, 30, 56-67. https://doi.org/10.1016/j.jtrangeo.2013.03.003

Fremont., A. (2007). Global Maritime *Networks*. The case of Maersk. *Journal of Transport Geography* 15 (6), *431-442.* https://doi.org/10.1016/j.jtrangeo.2007.01.005

Fremont, A., Soppe, M. (*2004*). Les stratégies des armateurs de lignes régulières le desserte maritime et la domination asiatique et le renouveau. *Belgeo*, 4, 391-406. https://doi.org/10.4000/belgeo.13429

Frémont, A., Soppé, M. (2007). Northern European range: shipping line concentration and port hierarchy. In: Wang J., Olivier D., Notteboom T. E. and Slack B., *Ports, Cities, and Global Supply Chains.* Aldershot: Ashgate, 105-120.

Gonzalez, F., Freire, M., Pais, C. (2012). Maritime degree, centrality and valnerability: port hierarchies and emerging areas in containerized transport (2008-2010). *Journal of Transport Geography*, 24, 33-44. https://doi.org/10.1016/j.jtrangeo.2012.06.005

Gosasang V., Chandraprakaikul W., Kiattisin S. (2011). A comparison of traditional and neural networks forecasting techniques for container throughput at Bangkok port. The *Asian Journal of Shipping and* Logistics, **27** (3), 463–482. https://doi.org/10.1016/S2092-5212(11)80022-2

Grifoll Colls, M. (2019). A statistical forecasting model applied to container throughput in a multi-port gateway system: the Barcelona-Tarragona-Valencia case. *International journal of shipping and transport logistics*, 11 (4), https://doi.org/316-333. 10.1504/IJSTL.2019.10013465

Guo, T., Song, S., & Ma, W. (2021). Point and interval forecasting of groundwater depth using nonlinear models. *Water Resources Research*, 57, e2021WR030209. https://doi.org/10.1029/2021WR030209

Hayuth, Y. (1981). Containerization and the load centre concept. *Economic Geography,* 57 (2*),* 160-176. https://doi.org/10.1080/03088839400000041

He, C., Wang, H., (2021). Container Throughput Forecasting of Tianjin-Hebei Port Group Based on Grey Combination Model. *Journal of Mathematics*, vol. 2021, Article ID 8877865, 9 pages. https://doi.org/10.1155/2021/8877865

Hu, Y., Zhu, D. (2009). Empirical analysis of the worldwide maritime transportation network. *Physica A: Statistical Mechanics and its Applications,* 388 (10), 2061-2071. https://doi.org/10.1016/j.physa.2008.12.016

Huxley, J., Teissier, G. (*1936*). Terminology of Relative Growth. *Nature,* 137, 780-781. http://dx.doi.org/10.1038/137780b0

Hyndman, R.J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. Accessed on February 24, 2022.

Kaluza, P., Koelzsch, A., Gastner, M.T., Blasius, B. (2010). The complex network of global cargo ship movement. *Journal of The Royal Society Interface*, 7 (48), 1093-103. https://doi.org/10.1098/rsif.2009.0495

Kang, D.J., Woo S.H. (2017). Liner shipping networks, port characteristics and the impact on port performance. *Maritime Economics and Logistics*, 19 (2), 274-295 https://doi.org/10.1057/s41278-016-0056-2

Lam, W.H.K., Asce, M., Ng, P.L.P., Seabrooke, W., Hui, E.C.M. (2004). Forecasts and reliability Analysis of Port Cargo Throughput in Hong Kong. *Journal of urban planning and development*, 130 (3), 133-144.

Le, Y., Ieda, H. (2010). Evolution Dynamics of Container Port Systems with a Geo-Economic Concentration Index: A Comparison of Japan, China and Korea. *Asian Transport Studies,* 1(1), 46-61. https://doi.org/10.11175/eastsats.1.46

Lee, S.W., Song, D. W., Ducruet, C. (2008). A tale of Asia world ports: the spatial evolution in global hub port cities. *Geoforum* 39*,* 372-385. https://doi.org/10.1016/j.geoforum.2007.07.010

Li, Z., Xu, M., Shi, Y. (2015). Centrality in global shipping network basing on worldwide shipping areas. *GeoJournal* 80, 47–60. https://doi.org/10.1007/s10708-014-9524-3

Low, J.M.W., Lam, S.W., Tang, L.C., (2009). Assessment of hub status among Asian ports from a network perspective. *Transportation Research Part A*, 43, 593-606. https://doi.org/10.1016/j.tra.2009.04.004

McCalla, R., Slack, B. Comtois, C. (2005). The Caribbean basin: adjusting to global trends in containerization. *Maritime Policy and Management* 32 (3), 245-261. https://doi.org/10.1080/03088830500139729

Milenković, M., Milosavljevic, N., Bojović, N., Val, S. (2021). Container flow forecasting through neural networks based on metaheuristics. *Operational Research*, 21(2), 965-997. https://doi.org/10.1007/s12351-019-00477-1

Munim, Z. H., Schramm, H. J. (2017). Forecasting container shipping freight rates for the far East-Northern Europe Trade Line, *Maritime Economics & Logistics* 19 (1), 106-125. doi.org/10.1057/s41278-016-0051-7

Notteboom, T. E. (1997). Concentration and Load Centre Development in the European Container Port System. *Journal of Transport Geography*, 5, 99-115. https://doi.org/10.1016/S0966-6923(96)00072-5

Notteboom, T. E. (2006). Container Throughput Dynamics in the East Asian Container Port System. *Journal of International Logistics and Trade,* 4(1), pp. 31-52. https://doi.org/10.24006/jilt.2006.4.1.31

Notteboom, T. E. (2010). Concentration and the formation of multi-port gateway regions: An update. *Journal of Transport Geography*, 18, 567-583. https://doi.org/10.1016/j.jtrangeo.2010.03.003

Notteboom, T. E. (2022a). Ports and container shipping. In: Notteboom, T. E., Pallis, A. and Rodrigue, J.-P. (Eds.), *Port Economics, Management and Policy*, New York: Routledge. doi.org/10.4324/9780429318184

Notteboom, T. E. (2022b). Port planning and development. In: Notteboom, T. E., Pallis, A. and Rodrigue, J.-P. (Eds.), *Port Economics, Management and Policy*, New York: Routledge. doi.org/10.4324/9780429318184

Notteboom, T. E., Rodrigue, J. P. (2005). Port Regionalization: Towards a New Phase in Port Development. *Maritime Policy and Management,* 32(3), 297-313. https://doi.org/10.1080/03088830500139885.

Oliveira, G. F., Schaffar, A., Cariou, P., Monios, J. (2021). Convergence and growth traps in container ports. *Transport Policy*, 110, 17. https://doi.org/10.1016/j.tranpol.2021.05.027

Overman, H. G., Ioannides, Y. M. (2001). Cross-Sectional Evolution of the U.S. City Size Distribution. *Journal of Urban Economics*, 49, 543–566. https://doi.org/10.1006/juec.2000.2204

Pais-Montes, C. Freire Seoane M. J., Gonzalez-Laxe, F. (2012). General cargo and containership emergent routes: A complex networks description. *Transport Policy*, 24, 126-140. https://doi.org/10.1016/j.tranpol.2012.06.022

Pandey, P., Tripura, H., Pandey, V. (2018). Improving Prediction Accurancy of Rainfall Time Series by Hybrid SARIMA-GARCH Modeling, *Natural Resources Research,* 28 (1). https://doi.org/10.1007/s11053-018-9442-z

Pang G, Gebka B. (2017). Forecasting container throughput using aggregate or terminal-specific data? The case of Tanjung Priok Port, Indonesia. *International Journal of Production Research*. 2454-2469 http://dx.doi.org/10.1080/00207543.2016.1227102

Peng, W. Y., Chu, C. W. (2009). A comparison of univariate methods for forecasting container throughput volumes. *Mathematical and Computer Modelling*, 50 (7-8), 1045-1057. https://doi.org/10.1016/j.mcm.2009.05.027

Puertos del Estado. (2021, 2022). Monthly statistics. Retrieved May 15, 2021, and March 5, 2022, from https://www.puertos.es/en-us/estadisticas/Pages/estadistica_mensual.aspx

Quah, D. T. (1993). Empirical Cross-Section Dynamics in Economic Growth. *European Economic Review*, 37, 426-434. http://dx.doi.org/10.1016/0014-2921(93)90031-5.

Rashed, Y., Meersman, H., Van de Voorde, E., Vanelslander, T. (2017). Short-term forecast of container throughout: An ARIMA-intervention model for the port of Antwerp. *Maritime Economics and Logistics*, 19 (4). 749-764. https://doi.org/10.1057/mel.2016.8

Rimmer, P. (1967). The search for spatial regularities in the development of Australian seaports 1861 – 1961/2. *Geograkiska Annaler*, 49,42-54.

Rodrigue, J. P. (2020). What is Transport Geography? In Rodrigue, J. P. (Ed.), *The Geography of Transport Systems,* Fifth Edition, New York: Routledge. doi.org/10.4324/9780429346323

Rodrigue, J. P. and Notteboom, T. E. (2022). Maritime Shipping and International Trade. In: Notteboom, T. E., Pallis, A. and Rodrigue, J.-P. (Eds*.), Port Economics, Management and Policy*, New York: Routledge. doi.org/10.4324/9780429318184

Rosen, K. T., Resnick, M. (1979). The Size and Distribution of Cities: An Examination of Pareto Law and Primacy. *Journal of Urban Economics*, 8 (2), 165–186.

Saaty, R. W. (1987). The Analytic Hierarchy Process—What It Is and How It Is Used. *Mathematical Modelling*, 9, 161-176. http://dx.doi.org/10.1016/0270-0255(87)90473-8

Schaffar, A., Dimou, M. (2012). Rank-size City Dynamics in China and India, 1981–2004. *Regional Studies*, 46 (6), 707-721. https://doi.org/10.1080/00343404.2010.521146

Schulze P., Prinz, A., (2009). Forecasting container transshipment in Germany. *Applied Economics*, 41 (22), 2809-2815. https://doi.org/10.1080/00036840802260932

Shetty, D. K., Sumithra, Ismail. B., (2018). Hybrid SARIMA- GARCH Model for Forecasting Indian Gold Price. *RESEARCH REVIEW International Journal of Multidisciplinary*, 3 (08), 263-269.

Shingleton, A. (2010). Allometry: The Study of Biological Scaling. *Nature Education Knowledge*, 3 (10), 2.

Taaffe, E. J., Morril, R. L., and Gould, P. R. (1963). Transport expansion in underdeveloped countries: a comparative analysis. *Geographical Review,* 53*,* 503-529. https://doi.org/10.2307/212383

Tovar, B., Hernández, R., Rodríguez-Déniz, H., (2015). Container port competitiveness and connectivity: The Canary Islands main ports case, *Transport Policy* 38, 40-51.

Tran, N. K., Haasis, H. D. (2014). Empirical analysis of the container liner shipping network on the East-West corridor (1995–2011). *Netnomics* 15 (3),121-153. https://doi.org/10.1007/s11066-014-9088-x

Twrdy, E., & Batista, M. (2016). Modeling of container throughput in Northern Adriatic ports over the period 1990–2013. *Journal of Transport Geography*. 52(C), 131-142. https://doi.org/10.1016/j.jtrangeo.2016.03.005

UNCTAD. (2022). Review of maritime transport. United Nations: New York and Geneva.

United Kingdom Department of Transport. (2021). Freight traffic by port and year (filter by direction and cargo type). Retrieved May 15, 2021, from https://www.gov.uk/government/statistical-data-sets/port-and-domestic-waterborne-freight-statistics-port#port-level-statistics

Veenstra, A.W., Mulder, H.M., Sels, R.A. (2005). Analysing container flows in the Caribbean. *Journal of Transport Geography* 13(4), 295-305. https://doi.org/10.1016/j.jtrangeo.2004.07.006

Vining, D. R. Jr. (1976). Autocorrelated Growth Rates and the Pareto Law: A Further Analysis. *Journal of Political Economy,* University of Chicago Press, 84(2), 369-380, https://doi.org/10.1086/260440

Wang, Y., Cullinane, K. (2008). Measuring Container Port Accessibility: An Application of the Principal Eigenvector Method (PEM). *Maritime Economics & Logistics*, 10 (1-2), 75-89. https://doi.org/10.1057/palgrave.mel.9100192

Wang, *Y*., *Cullinane*, *K*. (*2014*). Traffic consolidation in East Asian container ports: A network flow analysis. *Transportation Research Part A: Policy and Practice*, 61, 152-163. https://doi.org/10.1016/j.tra.2014.01.007

Wang, Y., Cullinane, K. (2016). Determinants of ports centrality in maritime container transportation. Transportation Research Part E: Logistics and transportation Review, 95, 326-340. https://doi.org/10.1016/j.tre.2016.04.002

Wang, C., Ducruet, C. (2012). New port development and global city making: emergence of the Shanghai-Yangshan multi-layered gateway hub. *Journal of Transport Geography*, 25, 58-69. https://doi.org/10.1016/j.jtrangeo.2012.07.008

Wang, C., Ducruet, C. (2013). Regional Resilience and Spatial Cycles: Long-Term Evolution of the Chinese Port System (221 bc -2010 ad). *Tijdschrift voor Economische en Sociale Geografie*, 104 (5), 521-538. https://doi.org/10.1111/tesg.12033

Wang, C., Wang, J., (2011). Spatial pattern of the *global shipping network* and its hub-and-spoke system. *Reearch in Transport Economics*, 32, 54–63. https://doi.org/10.1016/j.retrec.2011.06.010

Wilsmeier, G., Monios, J. (2013). Counterbalancing peripherality and concentration: an analysis of the UK container port system. *Maritime Policy and Management,* 40 (2)*,* 116- 132. https://doi.org/10.1080/03088839.2012.756588

Wilsmeier, G., Monios, J., Pérez-Salas, G. (2014). Port system evolution: the case of Latin America and the Caribbean. *Journal of Transport Geography*, 39, 208-221. https://doi.org/10.1016/j.jtrangeo.2014.07.007

Wilsmeier, G., Notteboom, T.E. (2011). Determinants of liner shipping network configuration: A two-region comparison, *GeoJournal* 76(3), 213-228. https://doi.org/ 10.1007/s10708-009-9333-2

Xiang, Y., (2022). Using ARIMA-GARCH Model to Analyze Fluctuation Law of International Oil Price, *Hindawi Mathematical Problems in Engineering.* doi.org/10.1155/2022/3936414

Xie, G., Wang, S., Zhao, Y., Lai, K. K. (2013). Hybrid approaches based on LSSVR model for container throughput forecasting: A comparative study. *Applied Soft Computing*, 13 (5), 2232-2241. doi.org/10.1016/j.asoc.2013.02.002

Xu M, Li Z, Shi Y, Zhang X, Jiang S. (2015). Evolution of regional inequality in the global shipping network. *Journal of Transport Geography*, 44, 1-12. doi:10.1016/j.jtrangeo.2015.02.003

Xu, M., Pan, Q., Muscoloni, A., Xia, H., Cannistraci, CV. (2020). Modular gateway-ness connectivity and structural core organization in maritime network science. *Nat. Commun.* 11, 2849. https://doi.org/10.1038/s41467-020-16619-5

Zhang, W., Deng, W. and Li, W. (2018). Statistical properties of links of network: A survey on the shipping lines of worldwide marine transport network. *Physica A: Statistical Mechanics and its Applications,* 502, 218-227.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*, Cambridge, (Mass.).