

UNIVERSITY OF NAPLES FEDERICO II

Department of Civil, Architectural and Building Engineering

Ph.D program in Civil System Engineering XXXII cycle

Ph.D. Candidate

**Daniela Tocchi**

**Theoretical improvements in modelling worldwide container  
networks**

**Ph.D course Coordinator:**

Dr. Prof. Andrea Papola

**Supervisor:**

Dr. Prof. Vittorio Marzano

**Co-supervisor:**

Dr. Prof. Christa Sys

## ACKNOWLEDGMENTS

First and foremost, I would like to thank my supervisor, prof. Vittorio Marzano, for the time and the effort taken to go through this doctoral research, the kind understanding and the valuable advice.

I would also like to thank my co- supervisor, Prof. Christa Sys for her continuous, kind and strong support. I am deeply grateful for the time she dedicated to me and for her guidances, which have been precious throughout my entire Ph.D experience. Also, she inspired me during my visiting period at the University of Antwerp.

My gratitude also extends to Prof. Francesco Viti and Prof. Enrica Papa, who, through their comments, gave me the possibility to improve this work.

Special thanks to my family for their support in making this effort possible and my late mother for her teachings and values. In them I found the strength and the resilience to walk along this challenging path.

Most importantly, I would like to thank Antonio for his patience and endless support. I would like to dedicate this undertaking to him and our beloved daughter, Alice.

# CONTENTS

Introduction .....	8
1.1 Motivation and background.....	8
1.2 Research questions and thesis contribution.....	11
1.3 Co-authored papers published (or in course of publication) within the PhD research work.....	13
1.4 Outline of the thesis.....	15
2 Literature Review .....	16
2.1 National and international DSSs.....	16
2.2 Shortest paths in multimodal freight networks.....	21
2.3 Maritime (worldwide) container service network models.....	25
2.4 Network analysis and port centrality measures .....	26
2.4.1 Common graph metrics.....	26
2.4.2 Relevant metrics for the networks of maritime container services.....	28
2.4.3 Literature review on application of graph metrics for the maritime container market .....	30
2.5 Literature review findings .....	33
3 Rail freight supply modelling.....	34
3.1 Motivation and background.....	34
3.2 The proposed shortest unit cost path algorithm for rail freight.....	35
3.2.1 Calculation of train capacity.....	36
3.2.2 Calculation of train cost.....	37
3.2.3 Calculation of shortest unit cost path.....	39
3.3 An application: a new incentive scheme for rail freight undertakings.....	41
4 Modelling worldwide maritime container networks .....	48
4.1 The container shipping industry .....	48
4.1.1 Maritime container services: key features .....	51
4.2 A database of maritime container services from 2018 to 2020.....	58
4.2.1 Structure of the database.....	58
4.2.2 Implementation of the database .....	65
4.3 The supply model of container shipping services .....	69
4.3.1 Definition of the study area and its zoning.....	69
4.3.2 Topological model.....	70
4.4 Calculation of centrality metrics in hypergraphs for maritime container service networks.....	77
4.4.1 Degree centrality.....	77

4.4.2	Closeness centrality .....	77
4.4.3	Betweenness centrality .....	77
4.4.4	Results.....	79
4.5	Multimodal integration of port-to-port and landside supply models.....	86
5	Container shipping services: descriptive statistics and concentration analysis.....	91
5.1	Container shipping services supply .....	91
5.1.1	Services and fleet composition of top ten carriers.....	94
5.1.2	Services and liner fleet of ocean carriers' alliances.....	98
5.1.1	Container services supply: regional market and port-based analyses .....	107
5.2	Shipping market concentration analysis.....	112
5.2.1	Herfindahl-Hirschman index .....	112
6	A GLS-based procedure to estimate/update freight flows between ports .....	119
6.1	Introduction .....	119
6.2	Development of a mathematical tool for Ro-Ro/Ro-Pax supply.....	120
6.3	Estimation of the Ro-Ro/Ro-Pax freight flows between Italian ports given the port throughput: application to Italian ports.....	121
7	Summary and conclusions.....	125
8	References .....	130

## LIST OF FIGURES

Figure 1: Components of the Blue Economy .....	8
Figure 2: Container throughput of Italian ports .....	9
Figure 3: Shortest time path in multimodal networks: a counterexample (network structure: top; shortest time additive path: middle; shortest time overall path: bottom).....	23
Figure 4 : Characteristics of the Italian railway network in 2015: train length and loading gauge and weight class and slope.....	44
Figure 5 : Distribution of the infrastructural gap across o-d pairs – intermodal freight transport .....	45
Figure 6 : Distribution of the average regional infrastructural gap for Italian regions (only intermodal freight transport) .....	47
Figure 7: Development of international maritime trade by cargo type (Index: 1990 = 100)...	49
Figure 8: evolution of cellular ship size.....	54
Figure 9: Fleet capacity breakdown by TEU size range .....	55
Figure 10: Shipping alliances – source own composition.....	57
Figure 11: market share of shipping alliances on busiest container routes.....	57
Figure 12: structure of the database of the maritime container services.....	58
Figure 13: overview of regional macro-areas .....	66
Figure 14: application of splitting method .....	72
Figure 15:maritime graph.....	73
Figure 16: L-graph and P-graph to model maritime container service networks (links in the P-graph are two way .....	74
Figure 17: Illustration of a HL-graph related to in the example of Figure 1: representation by means of hyperlinks (top) and waiting links (detail for node A, bottom).....	75
Figure 18: Distribution of the closeness centrality in HP-graphs vs. P-graphs.....	81
Figure 19: Distribution of the betweenness centrality in HL-graphs vs. L-graphs (top) and in HP-graphs vs. P-graphs (bottom).....	85
Figure 20 : Illustrative example of the approach pursued by the NODUS model for a three-modes example (road, sea, maritime).....	88
Figure 21 : Flowchart for the implementation of the proposed approach .....	89
Figure 22: active ships by TEU size range – 2018.....	92
Figure 23: active ships by TEU size range – 2019.....	92
Figure 24: active ships by TEU size range – 2020.....	93
Figure 25: active ships by TEU size range between 2018 and 2020.....	93
Figure 26: Breakdown of operated capacity by TEU size range for Top 10 ocean carriers in 2018.....	97
Figure 27: Breakdown of operated capacity by TEU size range for Top 10 ocean carriers in 2019.....	97
Figure 28: Breakdown of operated capacity by TEU size range for Top 10 ocean carriers in 2020.....	98
Figure 29: number of shipping carriers’ liner services addressed to ocean alliances – 2018 ..	99
Figure 30: number of shipping carriers’ liner services addressed to ocean alliances - 2019.	100
Figure 31: number of shipping carriers’ liner services addressed to ocean alliances - 2020.	100
Figure 32: changes in the fleet composition (2018-2019) and (2019-2020) – 2M .....	104
Figure 33: changes in the fleet composition (2018-2019) and (2019-2020) – Ocean Alliance .....	104
Figure 34: changes in the fleet composition (2018-2019) and (2019-2020) – The Alliance.	105

Figure 35: fleet breakdowns of ocean alliances .....	105
Figure 36: weekly capacity provided by liner container services .....	108
Figure 37: weekly capacity provided by e shipping Alliances.....	109
Figure 38: HHI based on weekly capacity of shipping carriers in ports .....	117
Figure 39: HHI based on weekly capacity of ocean alliances in ports .....	118
Figure 40: Structure of the database of Ro-Ro/Ro-Pax services under analysis.....	120
Figure 41: cvRMSE between true and estimated port-to-port flows in the laboratory experiment to check performance of the proposed estimator (6.4).....	124
Figure 42: Comparison between model estimates (x-axis) and true values (y-axis) of attracted Ro-Ro freight flow by port.....	124

## LIST OF TABLES

Table 1 :Summary of most relevant European freight DSS.....	17
Table 2 : Yearly quantification of the infrastructural gap for intermodal transport.....	46
Table 3: global port throughput – source: Alphaliner Monthly Monitor, January 2020.....	49
Table 4: Cellular fleet breakdown.....	54
Table 5: ranking of top 30 shipping companies based on total operating capacity in 2020 ....	56
Table 6: structure of the table portable (primary key in bold) .....	59
Table 7: structure of the table country .....	59
Table 8: structure of the table macro area .....	59
Table 9: structure of the table trade lanes .....	60
Table 10: structure of the table carrier .....	60
Table 11: structure of the table alliance .....	60
Table 12: structure of the table service type (primary key in bold) .....	61
Table 13: structure of the table service .....	62
Table 14: structure of the table services trade lanes.....	63
Table 15: structure of the table services carrier .....	63
Table 16: structure of the table service string .....	63
Table 17: structure of the table vessel type .....	64
Table 18: structure of the table vessel.....	64
Table 19: structure of the table vessel services .....	65
Table 20: number of ports yearly called by worldwide services.....	66
Table 21: macro-areas details.....	67
Table 22: Number of active services per year.....	68
Table 23: top 30 ports Degree Centrality results in P-graph.....	79
Table 24: top 20 ports - closeness centrality ranking results in HP-graph vs. P-graph .....	80
Table 25: top 20 ports – closeness centrality ranking results in P-graph vs. HP-graph.....	80
Table 26: top 20 ports - betweenness centrality ranking results in HL-graph vs. L-graph.....	82
Table 27: top 20 ports - betweenness centrality ranking results in L-graph vs. HL-graph.....	83
Table 28: top 20 ports - betweenness centrality ranking results in HP-graph vs. P-graph .....	83
Table 29: top 20 ports - betweenness centrality ranking results in P-graph vs. HP-graph .....	84
Table 30: Number of active container services and their yearly variation.....	91
Table 31: Fleet capacity breakdown by TEU size range.....	94
Table 32: overview of container services operated by top ten shipping carriers .....	95
Table 33: active vessels and fleet capacity of top 10 ocean carriers.....	96
Table 34: overview of operated active services by ocean alliance .....	99
Table 35: Total active fleet capacity of ocean alliances.....	101
Table 36: fleet capacity share by carrier in the 2M alliance .....	101
Table 37: fleet capacity share by carrier in the Ocean Alliance.....	101
Table 38: fleet capacity share by carrier in The Alliance .....	102
Table 39: carrying capacity of each carrier participating in the 2M and percentage addressed to alliance services .....	102
Table 40:carrying capacity of each carrier participating in the Ocean Alliance and percentage addressed to alliance services.....	103
Table 41: carrying capacity of each carrier participating in The Alliance and percentage addressed to alliance services.....	103
Table 42: Share of each size range on the total fleet capacity by ocean alliance.....	106
Table 43: fleet capacity of 2M+ZIM agreement .....	107

Table 44: trade regions' compositions .....	107
Table 45: variations in weekly capacity supplied by liner container services by main trade region.....	108
Table 46: weekly capacity by port .....	110
Table 47: HHI based on weekly capacity of liner container services calling ports .....	114
Table 48: : HHI based on weekly capacity of shipping carriers .....	115
Table 49: : HHI based on weekly capacity of ocean alliances.....	116

# Introduction

The aim of this doctoral thesis is to model and to analyse worldwide container transport networks with a focus on the maritime liner shipping industry. This introductory chapter explains the motivations behind this project and outlines the research questions as well as the study contributions to the academic literature.

## 1.1 Motivation and background

The sea is one of the most important assets for the global economy. In Italy, the so-called "Blue economy" (Figure 1) contributes to 3% of the total GDP especially because of tourism and maritime freight transport (Unioncamere, 2019).

Figure 1: Components of the Blue Economy

Type of Activity	Ocean Service	Industry
Harvest of living resources	Seafood	Fisheries
		Aquaculture
	Marine biotechnology	Pharmaceuticals, chemicals
Extraction of non-living resources, generation of new resources	Minerals	Seabed mining
	Energy	Oil and gas
		Renewables
Fresh water	Desalination	
Commerce and trade in and around the oceans	Transport and trade	Shipping
		Port infrastructure and services
	Tourism and recreation	Tourism
		Coastal Development
Response to ocean health challenges	Ocean monitoring and surveillance	Technology and R&D
	Carbon Sequestration	Blue Carbon
	Coastal Protection	Habitat protection and restoration
	Waste Disposal	Assimilation of nutrients and wastes

Source: Blue Economy Development Framework, The World Bank, 2016.

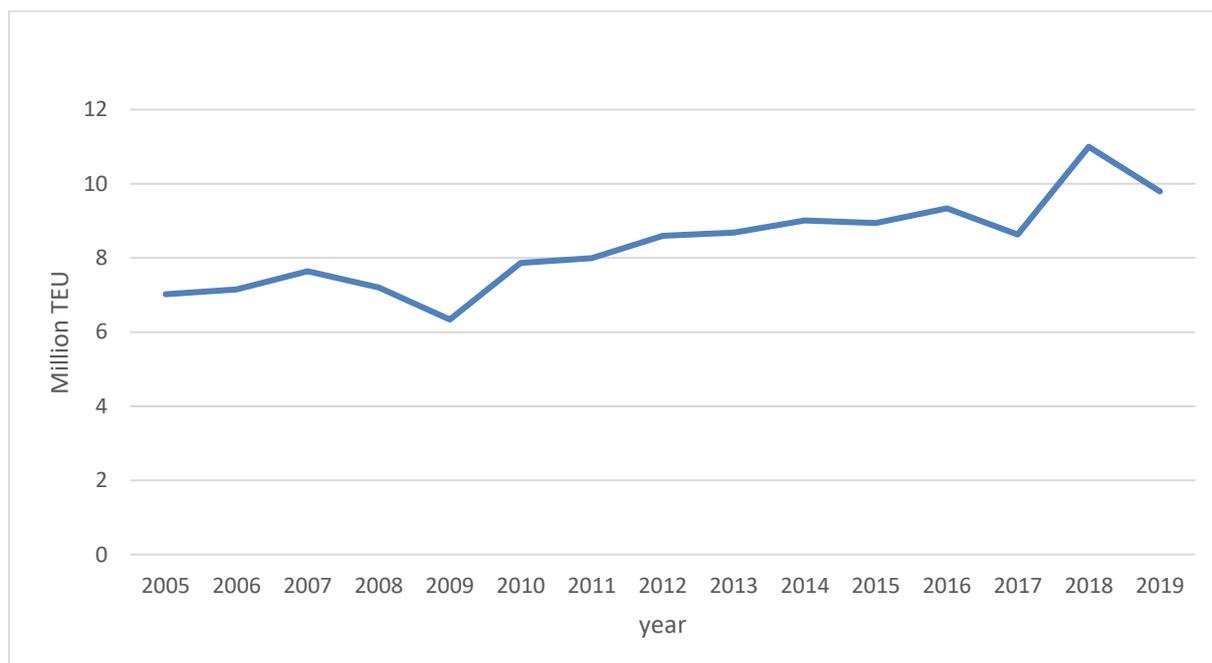
Transport infrastructures (i.e. road networks, railways, ports and logistics) as well as port governance play a key role in the growth of the maritime freight transport sector. Usually, planning and governance of ports and port systems is a very complex task.

In Italy, this difficulty has been magnified in the context of the "revolution" that the Government has been undertaking since 2016 to renew the national port system. Indeed, the Italian ports have been facing several drawbacks, mainly due to the lack of a proper planning at national level in the period 2001-2015 (Italian Ministry of Transport, 2016). Thus, the

Italian Ministry of Transport has deployed innovative measures to improve the national port system, as illustrated in the document “*Connettere l’Italia – Strategie per le infrastrutture di trasporto e logistica*” (Italian Ministry of Transport, 2016). The new planning approach provides a wide range of measures mainly aimed at improving the national port system, simplifying procedures and governance, and supporting sustainable freight transport.

Containerization plays a crucial role in freight transport. Before the recent COVID-19 pandemic, Italian ports handled around 10 million TEU/year (Figure 2).

**Figure 2: Container throughput of Italian ports**



*Source: own elaboration based on Eurostat data*

This growing market gained the attention of public and private operators, such as Port Authorities and carriers, who became more and more interested in increasing their container market share. To this purpose, they frequently ask the Italian Government for investments in port capacity and rail infrastructures. However, often the demand for additional investments is insufficiently based on quantitative analyses.

One key issue in assessing container-wise policy analyses relates with the issue of quantifying container-related freight flows in future scenarios including heterogeneous projects (new container terminals, capacity improvements on inland networks, immaterial facilitations to trade, ...).

Hence, it seems worthwhile adopting a proper tool to assess quantitatively the effects of planning choices related to maritime container transport.

To this purpose, this phd study aims to develop a worldwide container network model able to represents the current patterns of container flows and to identify strengths and weaknesses of container-related transport networks.

In literature, there are several models developed to represent freight transport systems (mono

and multimodal networks) both from public bodies and private operators. These so-called Decision Support System (DSSs) are widely used to support public and private stakeholders in planning and policymaking (Tavasszy, 1998). However, due to their simplified approach in the representation of multimodal networks, they usually fail to model peculiar phenomena related to container transport.

Besides, in the literature there are also studies dedicated to the maritime container transport. These tackled several issues such as the container network modelling, the implementation of network centrality and connectivity measures and the shipping market analysis.

In particular, the analysis of ocean container transport is a very wide line of research and it involves a large amount of closely related planning choices. As reviewed in Lee (2017), contributions available in the literature to date refer to several planning activities belonging to different timeframes (i.e. strategic planning, tactical choices, and operations management). Most of them are based on private operators' perspective (e.g. shipping companies and carriers' alliances). However, understanding features behind container network design and container cargo routing is of particular relevance especially to public bodies, as the outcome of planning choices determines the distribution of container flows in ports. Academic literature in this field presents a limited number of papers dealing with modelling of ocean container networks, as illustrated in the next section.

Also, many scholars focused on features affecting port choice and cargo routing adopted by shipping companies. Despite the great numbers of factors involved in port choice, some of the most relevant issues from a public perspective relate primarily to port infrastructure affecting accessibility over land and by sea, and port connectivity and centrality within container service lanes. Port connectivity and centrality are crucial indicators for policymakers, especially for Port Authorities (de Langen et al., 2007). These measures are important because of the attractiveness on potential users: a well-connected port provide access to a wide range of line trades. Moreover, as highlighted by several studies, not only did port connectivity affect trade cost at a regional level (Wilmsmeier and Hoffmann, 2008), but it also influenced costs at a global level (Arvis et al., 2013). The application of centrality measures to a network of maritime container services implies first deciding how to turn the network of maritime services into a graph. In this respect, the literature focuses primarily on synchronic graphs (Cascetta, 2009). However, these graphs exhibit an important limitation, as described by Bell et al. (2011) and Bell et al. (2013), who proposed the extension to container service networks of the approach based on the concept of hyperpath introduced by Spiess and Florian (1989). Notwithstanding this, there are no contributions in the related literature dealing with the implementation of network analysis on the hyperpath-based graphs.

## 1.2 Research questions and thesis contribution

Modelling worldwide container network is a key topic in assessing the effects of strategic and operational decisions on the port systems development. The most relevant current issues concern:

- the simplifications adopted in the available models which prevent adequate realism for many applications;
- the missing representation of non-additive key performance variables in the rail freight supply model;
- the fact that the hypergraph-based approach is not extended to the worldwide maritime container network;
- the estimation of port-to-port container flows on a global scale.

In this respect, this doctoral research aims to achieve distinct improvements over the state-of-the-art towards effective modelling of worldwide container transport networks in order to support port authorities, governments and public decision-makers in port planning activities. In detail, the research questions that this thesis aims to address are outlined in the following.

First, usually, in the design of a multimodal supply model, a challenging issue is given by the representation of peculiarities of each transport mode. In this case, the evaluation of system performances should account for non-additive path costs (e.g., dwell times of freight road vehicles, non-linear decreasing distance fees). This aspect is not currently taken into consideration by most of the available models. There are a few contributions in the literature trying to deal with the evaluation of the shortest paths in a multimodal network.

This issue has been investigated by the research group at the Department of Civil, Architectural and Building Engineering at the University of Naples Federico II during the last years. In detail, Vitillo (2011) focused on the design of a Decision Support System to model the multimodal freight transport network of the Euro Mediterranean basin, proposing sub-additive cost functions for road transport network. Later, Marzano et al. (2017) tackled this challenging issue proposing a practical approach to face the non-additive path costs using macrolink-based networks applied to the Euro-Mediterranean multimodal freight transport supply. This approach enables the evaluation of the shortest paths in a multimodal network including road and maritime transport legs yet neglecting road transport which is a key component of intermodal container transport chains.

This leads to the first research question:

1. *How to define a practical approach to evaluate shortest path with the non-additive costs in rail freight supply model?*

Second, the modelling of the worldwide container shipping services is a challenging task. In the literature there are only few contributions proposing a hyperpath -based approach (Bell 2011, Bell 2013). However, none of them has been applied to real maritime container service networks. Thus,

**2. *How to adapt the hypergraph-based approach to worldwide container service network modelling?***

Finally, whilst supply data (e.g., schedules of liner container shipping services and related vessel characteristics) are easily available, demand data on container flows on routes and port-to-port container flows matrices are commercially sensitive and difficult to collect. As a result, there is lack of demand data to inform freight demand models to assess both so-called matrix production-consumption (p-c) and primarily underlying origin-destination (o-d) flows. In this respect,

**3. *Is it possible to define a methodology to estimate port-to-port container flows on a worldwide scale?***

To answer these research questions this doctoral thesis proposed an innovative approach, which consists in the following steps:

1. implementation of a methodology to deal with non-additivity of key performance variables in the rail freight supply model;
2. definition of a hyperpath-based container maritime transport model;
3. analysis of liner container shipping services:
  - assessing effectiveness of the hyperpath-based centrality measures in container transport patterns;
  - investigating the shipping market structure through the evaluation of the HHI
4. definition of a practical approach to estimate port-to-port container flows.

### 1.3 Co-authored papers published (or in course of publication) within the PhD research work

Following the rationale of the thesis statement, the scientific contributions carried out in the framework of this PhD project are briefly outlined in the following papers published or under review:

- V. Marzano, A. Papola, F. Simonelli, R. Vitillo, D. Tocchi. *Shortest paths in freight multimodal networks with non-additive impedances: a practical approach*. Proceedings of the 96th Annual Meeting of the Transportation Research Board, Washington (D.C.), USA, Gennaio 2017 (No. 17-01856)

This study aims at tackling the issue of calculating shortest paths in multimodal freight networks, a key challenge for the presence of non-additive impedances (e.g. travel times, costs, fares) and also for the inherent heterogeneity of freight transport supply. Conceptually, since non-additive impedances can be handled only with explicit path enumeration, the proposed approach looks for a more tractable and less computationally demanding explicit path enumeration, based on two main steps. The former is the creation, starting from the monomodal networks for each freight mode, of so-called macrolinks representing direct connections between origins and destinations of that mode: this allows turning non-additive impedances in the initial monomodal network into additive impedances in the macrolink-based monomodal network, obviously at the price of increasing the number of links in the network. The latter is that all macrolink-based monomodal networks can be coupled together in a multimodal macrolink-based network, leveraging and slightly modifying earlier contributions in the literature, to achieve an effective model for the entire multimodal freight network. The proposed approach is applied to a very large real-size network representing the entire Euro-Mediterranean freight transport supply, demonstrating its capability to account for the effects of truck drivers' stop/rest time regulations (typically non-additive) on port choice and maritime service choice. Results show how the proposed approach actually improves the capability of calculating realistic shortest paths, with important policy implications, preserving a tractable

- Marzano, V., Tocchi, D., Papola, A., Aponte, D., Simonelli, F., & Cascetta, E. (2018). *Incentives to freight railway undertakings compensating for infrastructural gaps: Methodology and practical application to Italy*. **Transportation Research Part A: Policy and Practice**, 110, 177-188. (Q1)

The paper illustrates an incentive scheme to freight railway undertakings, proportional to the infrastructural gaps they experience on the rail network with respect to optimal performances in terms of loading gauge, length and weight. The proposed incentive is equitable and provided on an origin-destination pair basis, as opposed to the watering-can principle underlying current incentives schemes. In practice, it can be simply dispensed as a discount of the access charge the railway undertakings should pay to the rail network infrastructure manager. It also differs by type of train, to account for their different infrastructural needs, and can be adjusted on a yearly basis to account

for ongoing network improvements. The main methodological challenge lies in the quantification of the infrastructural gap, defined as the difference of the unit transport costs in the current and in the optimal scenario, as a consequence of the non-additivity of concerned costs. For this aim, a specific procedure is illustrated and applied to the railway intermodal transport in Italy, to show the feasibility of the approach and highlight the differences with respect to the current incentive schemes.

- Marzano, V., Tocchi, D., Fiori, C., Tinessa, F., Simonelli, F., & Cascetta, E. (2020). *Ro-Ro/Ro-Pax maritime transport in Italy: A policy-oriented market analysis. Case Studies on Transport Policy. (Q1)*

This paper proposes an in-depth analysis of Ro-Ro/Ro-Pax services in the Western Mediterranean, with a focus on Italy. Following an already consolidated research track, a database of Ro-Ro/Ro-Pax services has been built and analysed, enabling policy insights on port connectivity, market positioning of shipping companies across ports and routes, and congestion of Ro-Ro/Ro-Pax port terminals. Furthermore, an adaptation of the well-known GLS-based procedure that updates/estimates o-d flows from traffic measurements has been proposed, to estimate Ro-Ro/Ro-Pax freight flows between ports in a study area, given the total inbound/outbound port throughput and the total weekly capacity of port-to-port services. Application to both a laboratory experiment and to a real case study yielded very effective results. Overall, the presented analyses update earlier contributions in the literature and set the basis for an observatory on Ro-Ro/Ro-Pax services that might be regularly brought up to date and applied also to other countries.

- Tocchi D., Sys C., Papola A., Tinessa F., Simonelli F., Marzano V. (2021). *Hypergraph-based centrality metrics for maritime container service networks: a worldwide application. Submitted to Journal of Transport Geography. (Q1)*

Centrality metrics are commonly applied to analyse maritime container service networks, usually modelled as L-graphs (with links representing legs of each service) or P-graphs (with links representing direct port-to-port connections enabled by each service). In fact, a hypergraph-based approach would be more appropriate given the inherent nature of container services, as highlighted in concerned literature, however at the price of greater complexity. Notably, extension of common centrality metrics to hypergraphs is not straightforward and deserves attention: this paper aims to contribute to this topic, with a theoretical analysis and with an application to a worldwide network of container services relating to year 2019.

## 1.4 Outline of the thesis

Consistent with the research statement, the present monograph is structured as described below:

- the present Chapter 1 illustrates the scope of this doctoral research, outlining the project background, the research questions and the approach that will be adopted;
- Chapter 2 provides a deep review of the academic and scientific publications dealing with the research questions, highlighting the gaps in the literature so far;
- Chapter 3 outlines the methodological advances achieved in modelling landside rail freight networks, overcoming the drawbacks highlighted in the literature review;
- Chapter 4 shows the methodology adopted to build the supply model of worldwide maritime container services, based on the hyperpath approach. Also, it illustrated the application of well-known graph metrics to the hypergraph of the container services.
- Chapter 5 illustrates the results of the worldwide container liner services analysis based on:
  - descriptive statistics: which investigate on key features of the actual liner container network and their changes over a three-year period;
  - concentration measures, to assess the shipping market concentration at different levels (i.e. macroareas and ports).
- Chapter 6 proposes an o-d matrix estimation/upgrade technique to improve the accuracy of container demand data in ports. By a way of example, the procedure is illustrated referring to the estimation of Ro-Ro/Ro-Pax freight flows between ports, given the total inbound/ outbound port throughput and the total weekly capacity of port-to-port services;
- Chapter 7 summarises the main research outcomes of this doctoral project and proposes further research lines related to the transport network analyses.

## 2 Literature Review

The first step in the research is an in-depth literature review related to the container transport modelling. Given the interdisciplinary nature of this topic, ranging from economics through to transport engineering and geography, this section illustrates and groups all the relevant studies according to different lines of research, namely:

1. National and international DSSs;
2. Shortest paths in multimodal freight networks;
3. Maritime (worldwide) container service network models;
4. Network analysis and port centrality measures.

In detail, Section 2.1 reports a literature review of national and international freight transport models used by public authorities for transport planning purposes. Usually, these Decision Support Systems (DSSs) are comprehensive models representing a wide range of transport modes such as road, rail, maritime, inland waterways, and pipeline. However, due to the difficulty to represent in a comprehensive multimodal model all the peculiarities of each transport mode, usually these DSSs resort of several simplifications preventing adequate realistic results. Indeed, the review of DSSs is primarily finalized to show that their underlying supply models, intended to evaluate performances and externalities of freight transport system, are largely simplified. In particular, referring to a multimodal network, the costs associated to each path are not additive (e.g., dwell times of freight road vehicles, non-linear decreasing distance fees).

The issue of ameliorating the capability of freight supply models to deal with the inherent characteristics of multimodal freight transport has been faced by some scholars, who tried to propose some contributions, illustrated in Section 2.2. With an increasing level of detail, Section 2.3 focuses on ocean container transport modelling, highlighting the issues related to its implementation (e.g. network modelling, port choice, availability of data). Finally, the last section introduces relevant concepts of graph theory and concerned applications to measure network accessibility and connectivity.

### 2.1 National and international DSSs

In the literature, there are different Decision Support Systems designed for freight transport modelling. They mainly differ for the geographical coverage considered in the model (e.g. international, national, regional and urban models), the underlying assumptions on how key aspects of the phenomenon are modelled (e.g. capacity on maritime, port and inland networks; economies of scale; non-additivity in costs and times), the different stakeholders' perspectives considered (public authorities for transport planning purposes, private operators for logistics planning and operations) (de Jong et al., 2012).

A first review of the most relevant freight models developed for public authorities reveals that they are mainly based on the same approach usually adopted to model passenger transport systems. Indeed, many authors considered worthwhile adapting the four-step approach (Cascetta 2006, Blauwens et al. 2008) to freight transport as well, even if each of the four models can be very different from that in passenger transport. In the context of freight transport modelling, the four steps can be specialized as follows:

1. **production and attraction:** the quantities of goods to be transported from the various origin zones and the quantities to be transported to the various destination zones are determined (the marginals of the origin–destination (OD) matrix).
2. **distribution:** the flows in goods transport between origins and destinations (cells of the OD matrix) are determined.
3. **modal split:** the allocation of the commodity flows to modes (e.g. road, train, combined transport, inland waterways) is determined.
4. **network assignment:** after converting the flows in tons to vehicle-units, they can be assigned to networks (in some models this is about assigning truck flows together with passenger cars to road networks).

However, phenomena underlying freight flows and concerned transport/logistics choices are difficult to model. As pointed out by de Jong et al (2004), compared to passenger transport, in freight transport system there are remarkable differences in the decision-makers to model, in the heterogeneity of freight moved and in the complexity of the transport and supply chains. Indeed, often additional modules are needed to adapt the four-step approach to freight modelling. (e.g. conversion of freight flows in money units into tonnes, implementation of logistic choices).

Several authors highlighted the importance of including supply chain features in freight transport models (Southworth and Wigan, 2006; Turnquist, 2006) and emphasized the strong influence of logistics<sup>1</sup> on transport demand (Tavasszy et al. 2012). For the interests of the thesis, attention has been focused in a first stage on DSS developed for European and/or Euro-Mediterranean study area: Table 1 shows the main features of the aforementioned models.

**Table 1 :Summary of most relevant European freight DSS**

MODEL	GEOGRAPHICAL COVERAGE	CHOICES MODELLED	TRANSPORT MODES
SIMPT	Italy	Generation, Distribution, Modal split, Assignment	Road, Rail, Combined (road-rail)
SMILE	Netherlands	Generation, Distribution, Modal split, Logistics, Assignment	Road, Rail, Sea, Inland waterways, Air, Pipeline

<sup>1</sup> All the activities related to planning and implementing the movement of raw materials, inventory and finished goods from origin to final destination. Logistics choices include inventory control, material handling, ordering processes, plant and warehouse selection as well as transport mode choice (Mitra et al., 2015).

MODEL	GEOGRAPHICAL COVERAGE	CHOICES MODELLED	TRANSPORT MODES
BVWP	Germany	Generation, Distribution, Modal split, Assignment	Road, Rail, Inland waterways – separate models for Sea and Air transport
TRANS-TOOLS	Europe	Generation, Distribution, Modal split, Logistics, Assignment	Road, Rail, Sea, Inland waterways, Air, Pipeline
MODEV	France	Generation, Distribution, Modal split and assignment	Road, Rail, Combined (road-rail), Inland waterways
NODUS	Belgium - Europe	Modal split and Assignment	Road, Rail, Inland waterways
BASGOED	Netherlands	Generation, Distribution, Modal split, Assignment	Road, Rail, Inland waterways
LOGIS	Europe (focus on France)	Generation, Distribution, Modal split, Assignment	Road, Rail, Combined (road-rail), Inland waterways
WORLDNET	Europe	Generation, Distribution, Modal split, Assignment	Road, Rail, Sea, Inland waterways, Air
Norway	Norway	Generation, Distribution, Logistics, Assignment	Road, Rail, Sea, Air
Sweden (SAMGODS)	Sweden	Generation, Distribution, Logistics, Assignment	Road, Rail, Sea, Air
Flanders	Flanders and Brussels	Generation, Distribution, Logistics, Assignment	Road, Rail, Sea, Inland waterways, Air

Source: own composition

The Italian national transport model **SIMPT** - *Sistema Informativo per il Monitoraggio e la Pianificazione dei Trasporti* - (Cascetta et al. 1995, Marzano and Papola, 2004) has been developed to control both passenger and freight transport system and evaluate alternative policies. Starting from the representation of transport supply (road and rail networks), macro-economic and sociodemographic scenarios it enables the forecasting of national and international travel demand for different time periods, traffic flows on infrastructures and services, operating and investment costs, traffic returns and impacts on users, territorial accessibility, pollutant emissions, safety and energy consumption. In detail, the SIMPT uses a Multi Regional Input-Output model with elastic coefficients in generation and distribution steps. This model does not provide an explicit representation of logistic choices, but some main features are included in the modal split model (e.g. frequency of shipments, shipment weight, ...).

The **SMILE** model - *Strategic Model for Integrated Logistics and Evaluations* – (Tavasszy et al. 1998, Bovenkerk 2005), developed for the Netherlands, is the first national transport model including endogenous logistics. The second release, the SMILE+ model, included a multiregional input-output model for production and attraction, a gravity model for the distribution model and multimodal stochastic network assignment for modal split and route choice. However, in 2009, the Dutch MoT invested in the development of a new and simpler

model, namely the BASGOED. This is a basic freight transport model based on the four-step approach, which use a limited number of zones and commodity types and the existing unimodal transport models for the assignment. To have a more complete version of the model a roadmap to include logistics in the BASGOED model was proposed, on the basis of the previous SMILE+ model (Tavasszy 2011 and Tavasszy et al. 2012).

The national French transport model **MODEV** does not account for explicit logistic components. In this DSS, production and attraction are estimated based on regression models, distribution is represented by a gravity model and the modal split (between road, rail, combined road-rail and waterways) is given by a logistic model based on aggregated data. Similarly to the BASGOED model, in MODEV the assignment is unimodal.

The **BVWP** is a combined freight and passenger transport model, developed for German federal infrastructure planning. In this DSS, the generation and the distribution models are represented in the same way of MODEV. Whereas the modal split model, including road, rail and inland waterways, is based on disaggregated data. Separate modules represent sea and air freight transports. Similarly to most of the aforementioned models, in the BVWP the assignment is carried out unimodally. Also, this model does not account for logistics choices.

**TRANS-TOOLS** is a widely used tool in policy analysis developed by the European Commission (i.e. cost-benefit analysis of infrastructure policies, pricing policies, vehicle dimensions regulation and transport and accessibility analysis). The first version of the EU TRANS-TOOLS (TT1) combines different modelling techniques in transport generation and assignment, economic activity, trade, logistics, regional development and environmental impacts. It mainly consists of a system of models, representing both freight and passenger transports, following the well-known four-step approach. Logistics choices are modelled in a separate module similar to the Dutch SMILE+. Due to the inclusion in a separate module, the use of logistic model is not mandatory. It includes also a spatial computable general equilibrium (SCGE) model (CGEurope) which provides inputs to the freight model (GDP changes) and can also receive zone-zone generalised transport cost information from the freight model.

In a further version, TRANS-TOOLS 2 (TT2) includes various adjustments, namely an unconstrained gravity model for international trades, an aggregate logit to model the modal split and a jointly assignment for cars and trucks (at NUTS3 level). Furthermore, the European Commission Joint Research Centre's Institute for Prospective Technological Studies (IPTS) and DG TREN are currently working on a new release of TRANS-TOOLS 3 (TT3) which aim at increasing the level of detail with regard to the rail, maritime and air transport modules, to better analyse issues of cost, capacity and externalities of transport.

Also, the **WORLDNET** model has been developed as part of the European Commission projects. It has been used mostly to assess EU's Motorways of the Sea initiative. This model reproduces long distance freight transport within the European region, as well as intercontinental sea and air cargo routes. Differently from other models, it is able to build multimodal logistic chains, and reproducing both unimodal or multimodal routes (road, rail, sea air and inland waterway transport). The choice between the alternatives is given by a multinomial logit model. While the OD flow matrices are derived by using gravity models.

Moreover, Norwegian and Swedish transport authorities developed their logistics models for

national DSSs as part of a joint project to renew their previous multimodal DSSs. These models share the same theoretical overall approach and ADA (aggregate-disaggregate-aggregate) structure, but they are based on different settings (e.g. zoning, modes and vehicle types included, commodity classification used, costs functions modelled). In the Swedish model the logistic module the input PC matrices are estimated on the basis of National Commodity Flow Survey (CFS) through the application of gravity models. Whereas the Norwegian model uses a SCGE model to produce forecasts of PC matrices. Finally, in both cases the logistic model reproduces the o-d matrices (in terms of vehicles) that are assigned to the road, rail, sea and air networks (by unimodal assignment).

The same ADA structure followed by Norwegian and Swedish logistic models has been adopted for the Mobility Masterplan of Flanders. Based on the input PC matrices derived from an existing trade model, the logistics model reproduces the shipment size choice and the transport chain. Beside this, Flanders also developed a conventional multimodal freight transport (K + P Transport Consultants and Tritel 2006), while the Walloon region adopted a multimodal network model based on the NODUS software, which includes the whole Europe at a NUTS2 level.

**LOGIS** is another freight transport model developed for the European Commission, which has been widely applied in several EU works (e.g. Trans-European Networks projects, development of dedicated Rail Freight Networks, transport corridor-evaluations). The four-step approach is given by: regression models for generation, gravity models for distribution, aggregate logit models for modal split and unimodal assignment. Even it is widely used, it has no specific logistic module.

The **Great Britain Freight Model (GBFM)** is the freight component of the UK Department for Transport's National Transport Model. It was developed by MDS Transmodal, and includes 2,650 domestic and 350 foreign zones and 10 commodity groups (NSTR1). It replicates multi-modal flows along highway, railway and maritime networks and through the ports including connections with the Continent and Ireland but does not include any logistics module. Later, a metropolitan variant of this national model was developed for London, namely the Freight in London Model (FiLM) including some logistics elements.

Summarizing, the review of this literature reveals that although the DSSs developed so far are comprehensive models involving a wide range of transport modes, they usually do not properly represent the peculiarities of each single mode in a multimodal network approach.

To properly evaluate the performances of a freight transport system it is necessary to take into account that the costs associated to each path are not additive (e.g. dwell times of freight road vehicles, non-linear decreasing distance fees).

The next paragraph illustrates academic contributions trying to overcome this issue.

## 2.2 Shortest paths in multimodal freight networks

Calculating shortest paths in multimodal freight networks is important for a wide range of applications, including policy-making and governance (see for instance Cascetta, 2009a and de Dios Ortúzar and Willumsen, 2011) and also for supply chain and transport operations' optimization (e.g. Steadieseifi et al. 2014). In general, both synchronic and diachronic approaches<sup>2</sup> can be applied, as argued by Cascetta (2009b) amongst others. However, focusing attention on decision support systems at national and international level for transport policy and governance, multimodal freight supply models largely resort on synchronic networks (de Jong et al. 2013), and also embed remarkable simplifications to preserve full integration with equilibrium-based demand-supply interactions and to reduce calculation times. In fact, albeit acceptable for some types of policies, such simplifications might lead to considerable modelling errors in key policy applications, for instance in the identification of the catchment areas of ports and in the analysis of competition amongst freight modes.

In this respect, a first necessary characteristic of a multimodal freight supply model is its capability to model impedances associated to terminal nodes (i.e. nodes allowing transfer between modes), which can change depending on the connected modes and on the type of multimodal leg (e.g. first access, transshipment): for instance, a container transshipment operation in a port is normally associated with times and costs different from a container import/export operation. A second necessary characteristic is the capability to represent the remarkable heterogeneity of freight transport options (own account/third party, type of commodity, type of vehicle, type of loading unit...), which can lead to substantially different performances. A noteworthy approach capable to handle both issues is the NODUS model by Jourquin et al. (1996) and Beuthe et al. (2001). NODUS exhibits two key features: first, a specific topological representation of terminals –based on “exploding” all within-terminal connections between modes – enables differentiation of impedances for all possible transfers between modes; second, “virtual links” are created to model connections between terminals, being each virtual link representative of a specific freight service, e.g. characterized by different types of vehicles and/or different costs/prices. NODUS has been applied in some contexts, mainly for the analysis of freight elasticities and for the location of new freight terminals.

Mainly, the characteristics and the performances of a multimodal freight supply model depend much upon the specific impedance chosen for the calculation of the shortest path. From a fairly general viewpoint, four types of impedances can be taken into account: travel times, costs (i.e. in the light of the operators of the freight transport service), fares (i.e. in the light of the users of the freight transport service), and generalized costs. Generalized costs are normally given by a linear combination of costs and other factors (e.g. travel times, reliability) opportunely harmonized through monetary coefficients, e.g. value of travel time saving, value of reliability, (see for instance Wardman et al., 2012). Calculating the shortest

---

<sup>2</sup> In *synchronic networks*, nodes are not identified by a specific time coordinate, and the same node represents events occurring at different moments (instants) of time. In *diachronic networks*, nodes may have an explicit time coordinate and therefore represent an event occurring at a given instant (Cascetta, 2009). See further details in section 4.3.2.

path by generalized cost is generally not possible without explicit path enumeration, i.e. without exploring the full set of feasible paths. In fact, to circumvent this problem, one normally calculates sub-optimal shortest generalized cost paths by considering the generalized cost of the shortest time and of the shortest cost paths. When calculating shortest paths with respect to a standalone impedance, i.e. travel times, costs, or fares, a key issue in modelling freight supply is dealing with the presence of non-additive impedances.

By definition, a non-additive impedance cannot be associated with any specific links in the networks, thus preventing calculation of path impedances as sum of impedances of all links belonging to that path. Two main examples of non-additive freight impedances are non-linear freight fares, whose unit value (e.g. €/km or US\$/mile) usually decreases by distance, and the regulation on driving times and rest periods for truck drivers, leading to non-additive total travel times, see for instance Regulation (EC) No 561/2006 of the European Parliament and of the Council of 15 March 2006 on the harmonisation of certain social legislation relating to road transport and amending Council Regulations (EEC) No 3821/85 and (EC) No 2135/98 and repealing Council Regulation (EEC) No 3820/85. Travel costs can be non-linear as well with respect to freight flows, as a result of the superposition of two contrasting effects, the economies of scale on one hand and the congestion on the other hand.

In general, a shortest path with respect to a purely non-additive impedance can be calculated only by enumerating explicitly all feasible paths. However, a noteworthy exception is represented by a particular sub-class of non-additive impedances, called sub-additive impedances. By definition, an impedance  $i$  is sub-additive (Cascetta et al., 2013) if it can be decomposed into an additive component  $i_{add}$  and a non-additive component  $i_{nadd}$ , such that  $i_{nadd}$  is a non-decreasing function of  $i_{add}$ . This means that, given any two paths indexed by 1 and 2, the following holds:

$$i^1_{add} < i^2_{add} \Rightarrow i^1_{nadd}(i^1_{add}) < i^2_{nadd}(i^2_{add}) \Rightarrow i^1 = i^1_{add} + i^1_{nadd} < i^2 = i^2_{add} + i^2_{nadd} \quad (2.1)$$

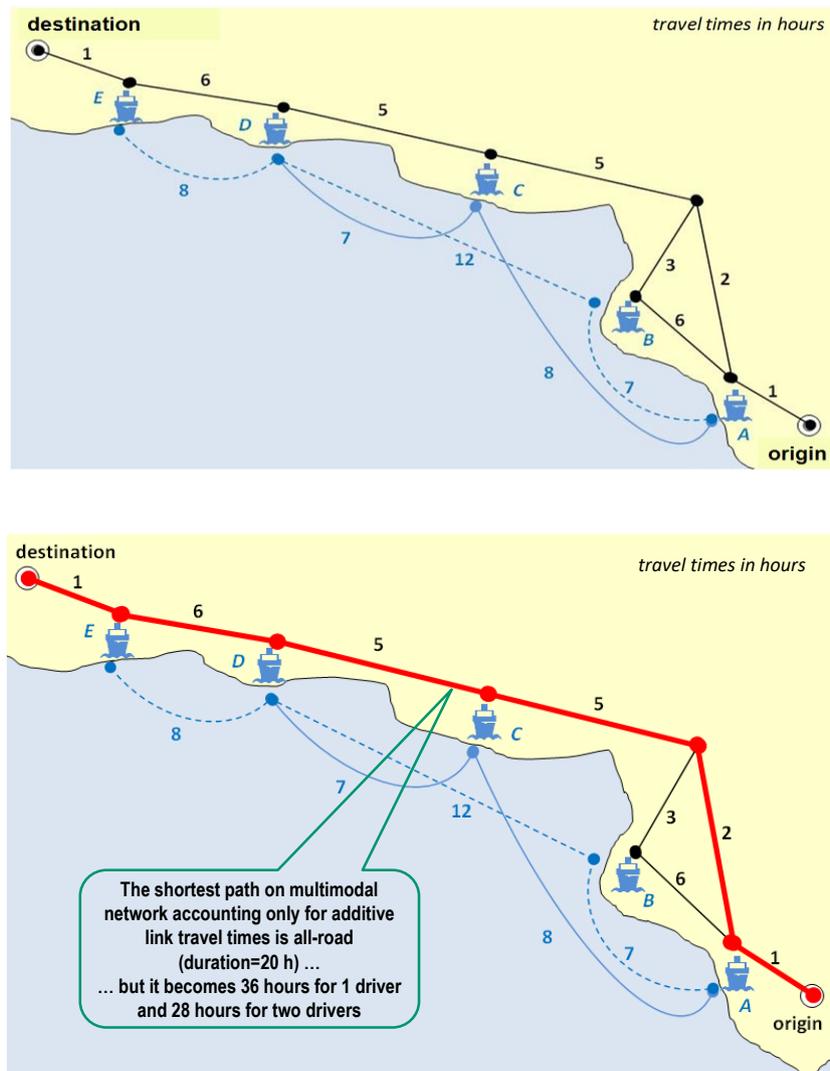
The key advantage of the sub-additivity property is that, thanks to (1), the shortest path with respect to the impedance  $i$  is still the shortest path with respect to the sole additive component  $i_{add}$ . In other words, it can be determined by calculating the shortest additive path with respect to the additive component  $i_{add}$  via a traditional shortest path algorithm (e.g. Dijkstra, 1959), and then adding in post process the corresponding non-additive component  $i_{nadd}$  in order to calculate its overall impedance.

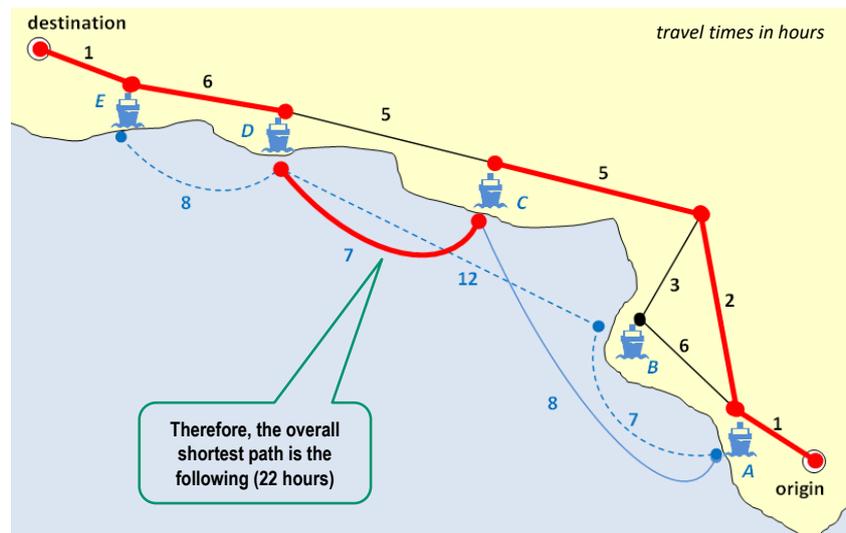
A very relevant example of sub-additive impedances in modelling freight supply is represented by travel times: indeed, all regulations on truck drivers' rest and stop times are such that, given the net road driving time  $tr_{road}$ , the mandatory additional rest/stop time  $ts_{road}$  is a non-decreasing function of the net road driving time  $tr_{road}$ , i.e. equation (2.1) holds. In this respect, the work by Shah (2008) and the subsequent application by Min (2011) leverage exactly this property to derive a modification of the Dijkstra's algorithm capable to account for rest/stop times in a monomodal road freight network. Similarly, also the structure of many freight service fares is such that the unit fare  $p_u(d)$  (e.g. US\$/mile or €/km) is a decreasing function of the distance  $d$ , in a way such that the overall cost is subadditive. That is, given two paths 1 and 2 with length  $d_1$  and  $d_2$ , their respective fares  $p_1 = d_1 \cdot p_u(d_1)$  and  $p_2 = d_2 \cdot p_u(d_2)$  are

such that  $d_1 < d_2 \Rightarrow p_1 = d_1 \cdot p_u(d_1) < p_2 = d_2 \cdot p_u(d_2)$ , i.e. the property (1) still holds. Thus, being the distance an additive impedance, the shortest fare path in a monomodal network can be still calculated simply by calculating the shortest distance path and then calculating the corresponding fare, which will be the minimum possible thanks to the subadditivity property.

The key issue is that the sub-additivity does not hold anymore in multimodal networks, i.e. when a path may include a sequence of different modes, each with specific non-additive impedances. In other words, one may take advantage of the sub-additivity property only if all links and paths in the network are characterized by the same impedance  $i$ . By way of example, Figure 3 provides a simple counterexample of a road-maritime freight multimodal network with sub-additive impedances (driving stop times) only on the road network: it is easy to recognize that the shortest additive time path (bold red in the middle of Figure 3) is actually not the shortest overall path (bold red in the bottom of Figure 3).

**Figure 3: Shortest time path in multimodal networks: a counterexample (network structure: top; shortest time additive path: middle; shortest time overall path: bottom)**





This example also clarifies that, if the objective of the modelling is for instance to understand the catchment areas of ports or the flows on maritime services, not accounting for the non-additivity of freight impedances might lead to significant biases in policy making (in the example, in the throughput of ports C and D and in the maritime flows between C and D).

In such cases, explicit path enumeration is required to calculate the shortest path with respect to a non-additive impedance. Since the number of paths is normally very large, even for medium-sized networks, proper heuristics should be applied. For instance, one might adopt a random generation algorithm, whose steps at the generic iteration are: (a) sampling additive link impedances from a normal distribution; (b) calculating the shortest additive path with respect to that impedance, and (c) calculate the total path impedance by adding the non-additive components. A sufficiently large variance of the sampling distribution will increase the chance of detecting good candidate paths, however at the price of increasing substantially the computational burden.

Alternatively, Jourquin (2007) reproduced approximately the impact of road rest times in multimodal freight networks using NODUS by clustering o-d pairs by distance band and applying multi-class assignment with separate cost functions by distance band. However, a proper clustering should be based on travel time bands rather than on distance bands, in accordance with the above; also, this approach does not allow capturing the complex effects on non-additivity along the entire transport chain.

### 2.3 Maritime (worldwide) container service network models

In the literature, there are also freight models developed specifically for container transport system. In this respect, a relevant contribution is given by Tavasszy's research on worldwide container traffic (Tavasszy et al., 2011). Tavasszy et al. (2011), propose a worldwide container model to analyse how changes in supply transport chains affect shipping companies' behaviours. Using a super-network approach (Sheffi, Y., 1985), the maritime network based on routing data available by shipping companies' websites is modelled. To represent the competition between ports belonging to neighbouring countries, a simplified inland transport network is provided. Also, to model routing choice behaviour of shipping companies, a path size logit is proposed. The generalized cost of each route is given by a linear combination of time and monetary costs. Port costs (e.g. port dues and terminal costs) are not explicitly included in the model specification. In this regard, an interesting supply chain cost model has been provided by van Hassel et al. (2015), as illustrated in the following. However, despite its innovative approach, Tavasszy's work (2011) introduces several approximations that allow modelling global container transport patterns, at the price of not properly capturing the dynamics behind changes in container port throughputs that are crucial inputs for planning.

Also, Bell et al. (2011), transfer the frequency-based transit assignment method of Spiess and Florian (1989) to global maritime container assignment model, to investigate the effects of sailing time, service frequency and port capacity on port choice. To model shipping maritime services, the virtual network approach is used; unlike Tavasszy et al. (2011), the inland transport network is neglected. The model is formulated by means of a linear program to allow large-scale network applications, i.e. to global maritime network. In a further contribution, Bell et al. (2013) introduce the concept that containers are more likely to be assigned minimising the expected cost than the travel time. Thus, the same authors formulate a cost-based container assignment model. In this model, ship operating costs are fixed and the objective is to minimise container handling costs, container rental and inventory costs. The model includes port capacities constraints as well and it is formulated as a linear program problem.

Moreover, other authors studied key features affecting container transport chain costs. Indeed, due to the evolution of cooperative behaviours (vertical and horizontal integration strategies) in container market, the competition between shipping companies and ports moved from a single transport leg to the entire supply chains. Notably, van Hassel et al. (2016) investigate on two key aspects of this change: how mega-ships affect the total generalised costs and which leg of total chain they affect the most. To answer these questions, these authors provide a very detailed set of cost models, referred to the different legs forming the supply chain. The ship model allows calculations of maritime generalised costs per TEU from 3 main costs: operational cost, voyage cost and capital cost. Furthermore, in the port model, all costs related to operating the vessel in the port, dues and third-party costs are provided. Also, diseconomies of scale (due the increasing time for loading and unloading a mega-ship) are modelled. As a result, it is possible to calculate the total time and costs per ship for each terminal in the port and thus the generalised cost during the port phase. The last model captures the inland

network transport costs (road, railway, and inland waterways).

In a recent publication, van Hassel et al (2020) adjusted the aforementioned model to analyse the impact of the expanded Panama Canal on the potential shift of cargo flows from US to Europe. To this purpose the authors extended the geographical coverage of the model, including some of the major ports of the US West Coast, accounting for more detailed port features (e.g. technical data related to maritime access to the port and terminal infrastructure characteristics).

## 2.4 Network analysis and port centrality measures

Network analysis is a classical and interdisciplinary research topic, spanning over a variety of consolidated quantitative methods and approaches, see e.g. Newman (2010). The paradigm of complex networks and its applications is increasingly applied as well, see e.g. Estrada (2011). In the transport sector, network metrics and indicators are widely used to inform inherent characteristics of transport networks (e.g. structure) and of their components (e.g. centrality/importance of nodes), see e.g. Rodrigue et al. (2006). For the purposes of this research, Section 2.4.1 and Section 2.4.2 recall the mathematical definition of relevant graph metrics and approaches, whilst Section 2.4.3 reports a literature review of their applications for the analysis of maritime container networks.

### 2.4.1 Common graph metrics

Let  $G \equiv \{N, L\}$  be a directed<sup>3</sup> graph characterised by a set of nodes  $N$  and a set of links  $L$ ; for any  $n \in N$ , let  $FWS(n)$  the forward star of node  $n$ , that is the set of links whose tail is  $n$  and similarly let  $BWS(n)$  the backward star of  $n$ , that is the set of links entering node  $n$ .

Freeman (1979) classified three main types of common graph metrics for nodes, here particularised for the case of a directed graph, that is:

- *degree centrality*. It is expressed by means of two simple indicators  $DC_{in}(n)$  e  $DC_{out}(n)$ , named *indegree* and *outdegree* centrality, respectively given by the cardinality of the backward star and of the forward star of a node  $n$ . The overall degree centrality  $DC(n)$  equals:

$$DC(n) = DC_{in}(n) + DC_{out}(n) = |BWS(n)| + |FWS(n)| \quad (2.2)$$

Indicator (1) can be also normalised by considering that at most a node can be connected to all other nodes in the network but itself, so it can be divided by  $|N|-1$ .

---

<sup>3</sup> Reference is made to a directed graph which is of interest for the paper. Also, in general, as clarified later, the graph does not necessarily correspond to the physical structure of the transport networks and/or of its services, because it might represent an abstraction.

- *closeness centrality*. Given a pair of nodes  $n, m \in N$ , let  $d_{nm}$  the impedance of the shortest path between  $n$  and  $m$ . The closeness centrality, normalised to account for the dimension of the network, is defined by Bavelas (1950) as:

$$CC(n) = \frac{|N|-1}{\sum_m d_{nm}} \quad (2.3)$$

- *betweenness centrality*. Given a node  $n$ , it is defined as the percentage of shortest paths between any other pairs of nodes  $i, j \in N - \{n\}$  passing through  $n$ . Letting  $n_{ij}$  be the number of (possible multiple) shortest paths between  $i$  and  $j$  and  $n_{ij}(n)$  the number of shortest paths between  $i$  and  $j$  passing through  $n$ , it occurs:

$$BC(n) = \frac{1}{(|N|-1)(|N|-2)} \sum_{i, j \in N - \{n\}} \frac{n_{ij}(n)}{n_{ij}} \quad (2.4)$$

These graph metrics have been extensively applied to a variety of research fields and have inspired also relevant generalisations and extensions, including the following:

- *harmonic centrality*. As discussed by Boldi and Vigna (2014), the presence of many non-connected pairs of nodes might yield a misleading interpretation of the closeness centrality (2.3), in such cases the normalised *harmonic centrality* is a more effective metric, given by:

$$HC(n) = (|N| - 1) \sum_n \frac{1}{d_{nm}} \quad (2.5)$$

wherein is assumed conventionally  $1/0 = \infty$ .

- *node strength*. The degree centrality (2.2) was generalised by Barrat et al. (2004) and Newman (2004) by replacing the count of ingoing and outgoing links with the sum of corresponding link weights, that is:

$$NS(n) = NS_{in}(n) + NS_{out}(n) = \sum_{l \in BWS(n)} w_l + \sum_{l \in FWS(n)} w_l \quad (2.6)$$

wherein  $w_l$  is the weight (e.g. a throughput or the reciprocal of an impedance) of link  $l$ , see also Barthelemy (2011). Opsahl et al. (2010) generalised further, by creating a metric depending upon a parameter  $\alpha \in [0, 1]$  such that  $\alpha=0$  yields (2.2) and  $\alpha=1$  yields (2.6).

- *eigenvector centrality*. As recalled by Mishra et al. (2012), the degree centrality of a node  $n$  can be modified by accounting for the importance of its adjacent nodes, measured through their degree centrality. This induces a circular dependence amongst the degree centralities of all nodes in the graph, that can be mathematically formulated as the search of eigenvalues of the transformation induced by the adjacency matrix of the graph. A power method can be applied to calculate the eigenvector centrality

$EC(n)$  of a node  $n$ , usually adopting the largest eigenvalue  $\lambda^*$  as input. For directed graph, separated eigenvector centralities can be calculated with reference only to the tail nodes of links of  $BWS(n)$ , that  $EC_{in}(n)$ , or to the head nodes of links belonging to  $FWS(n)$ , that is  $EC_{out}(n)$ . Usually,  $EC_{in}(n)$  is called prestige of node  $n$  and  $EC_{out}(n)$  importance of node  $n$ . The eigenvector centralities might suffer from issues when applied to directed graphs wherein nodes exist such that  $BWS(n) \equiv \emptyset$ , this issue is circumvented by the Kats metric. Further variants and extensions exist as well, e.g. the page rank algorithm.

- *hubs and authorities (HITS algorithm)*. It resembles the concept of eigenvector centrality, see Kleinberg (1999), by defining the concept of hub and authority nodes: an authority is a node such that the tail nodes of links of its backward star are hubs, and an hub node is such that the head nodes of links of its forward star are authorities. Calculation can be performed via a recursive algorithm and, similarly with (2.2) and (2.6), both a simple count of links and the adoption of a weight can be introduced.

Other graph metrics are common and not reported here only for the sake of brevity, e.g. the rich-club rank discussed by Ducruet (2013). Recalled metrics can be also extended to the graph, for instance by defining the *degree centrality of a graph*. Application of (2.2) to all nodes of the graph yields  $DC_{max} = \max_{n \in N} \{DC(n)\}$  and the possibility of setting a graph metric given by  $\sum_n |DC_{max} - DC(n)|$ . This quantity is maximal for a star-shaped graph wherein each node is connected with all other nodes in the network, and it can be proved that in this case it attains the value  $|N|^2 - 3|N| + 2$ . Thus, the centrality of a graph can be defined as:

$$DC(G) = \frac{\sum_n |DC_{max} - DC(n)|}{|N|^2 - 3|N| + 2} \quad (2.8)$$

At the graph level many other analyses, based on the paradigm of the complex networks, can be applied: for instance, the  $k$ -core decomposition finds the largest subgraph of a network, in which each node has at least  $k$  neighbours in the subgraph. Clustering algorithms based on the above metrics are also very common as well.

#### 2.4.2 Relevant metrics for the networks of maritime container services

Taylor et al. (2006), Low et al. (2009), and Tang et al. (2011) after revising metrics illustrated in Section 2.4.1, introduced a further port-specific graph metric, based on pairwise comparisons between pairs of ports. Specifically, given a port  $n$ , for any port  $m$  let  $n_n$  the number of ports reached only from  $n$ ,  $n_m$  the number of ports reached only from  $m$  and  $n_{mn}$  the number of ports in common, that is connected with both  $m$  and  $n$ . Overall, the pair of ports  $m$  and  $n$  allows connecting together  $2(n_n + n_{mn})(n_m + n_{mn})$  pairs of ports, which allows setting the following normalised *port connectivity* metric:

$$PC(n) = \frac{\sum_{m \neq n} 2(n_n + n_{mn})(n_m + n_{mn})}{\sum_k \sum_{m \neq k} 2(n_k + n_{mk})(n_m + n_{mk})} \quad (2.9)$$

Jiang et al. (2015) developed connectivity port measures based on optimization formulation and accounting for time and throughput considerations, from the perspective of a single global carrier.

Probably the most common metric in container shipping is the Liner Shipping Connectivity Index (LSCI) released yearly by UNCTAD since 2004 and described by Hoffman (2005). The LSCI is country-based and considers four main aspects: number of container vessel calls; container vessel carrying capacity; number of shipping companies, liner services and vessels; average and maximal vessel size. Further details are discussed by Fugazza and Hoffman (2017). The relationship between the LSCI and other logistics indicators and trade data was investigated amongst others by Ojala and Hoffman (2010) and by Arvis et al. (2013).

Some variants have been also proposed in the literature, such as the Liner Shipping Bilateral Connectivity Index (LSBCI) by UNCTAD (2016), that accounts for pairwise container liner shipping service analysis between countries, with an upper bound threshold on the maximum number of intermediate transshipments. Bertholdi et al. (2016) proposed a new indicator, termed Container Port Connectivity Index (CPCI), leveraging the LSCI and the HITS algorithm.

It is also worth mentioning other indicators not directly related to container but rather to the Ro-Ro market, see amongst others the connectivity, costs, and congestion indicators by PORTOPIA (2014) and by de Langen et al. (2016).

Although not straightforwardly related to the graph metrics and connectivity indices, another powerful measure to analyse container shipping services market is given by *Herfindahl-Hirschman index*. This is a commonly accepted measure of market concentration, evaluated by squaring the market share of each firm competing in the market and then summing the resulting numbers:

$$HHI = \sum_{f=1}^{n_f} ms_f^2 \quad (2.7)$$

wherein  $n_f$  is the number of firms in the market and  $ms_f$  the market share of the  $f$ -th firm. Intuitively, a node in a graph can be considered a firm and  $ms_f$  a related weight (e.g. the port throughput in case of maritime transport). Interpretation of the HHI is twofold: if  $ms_f$  is expressed in decimals between 0 and 1,  $HHI=1$  implies a monopolistic market whilst  $HHI \rightarrow 0$  corresponds to perfect competition; in particular, usually  $HHI < 0.20$  indicates strong competition,  $0.20 < 0.60$  indicates oligopoly,  $HHI > 0.60$  indicates monopoly. Similarly, if  $ms_f$  is expressed as a whole number (i.e. not in decimals), the US Department of Justice suggests  $HHI < 1500$  to indicate a competitive market,  $1500 < HHI < 2500$  a moderately concentrated market,  $HHI > 2500$  a highly concentrated market.

### 2.4.3 Literature review on application of graph metrics for the maritime container market

The maritime container market is only one of the application fields of the indicators reported in Section 2.4.1: by way of example, analysis of passenger and/or freight airline networks is very common as well, see e.g. Guimerà et al. (2005), Alderighi et al. (2007), Paleari et al. (2010), Scholz (2011), Arvis and Sheperd (2011), Roucolle et al. (2018). With reference to the container market, the concerned literature is vast, and contributions can be classified depending upon various inherent characteristics, including:

- *data source(s) to build the network.* Usually, data are available in the form of container services, either reconstructed based on company or port information or provided by specialised companies, mainly Alphaliner, Containerisation international, and Lloyd's marine intelligence unit. Since its deployment in 2001, Automatic Identification System (AIS) is also very popular. The former provides richer information but are cumbersome to collect, the latter are very easy to manage but, being mostly passive information, do not reveal a full picture of the phenomenon under analysis, see Ben-Akiva et al. (2016).
- *subject and focus of the analysis.* Container networks are studied from many standpoints, including topology of the network itself, port connectivity analysis, competition amongst ocean carriers. Consistently, subject of the analysis might be ports, ocean carriers, and so on.
- *network under analysis.* Literature focuses on synchronic graphs (Cascetta, 2009), and two options are usually available. The former reproduces precisely container service strings, that is each service calling at  $n$  ports is represented via  $n-1$  links representing subsequent port-to-port sailing. The latter adopts a so-called multigraph – see Ducruet (2013) – wherein each service calling at  $n$  ports is represented by  $n(n-1)$  links, modelling connections made available by the service for any pairs of called ports in its string. The two approaches are also termed in various ways, e.g. Tovar et al. (2015) define Graph of Direct Links (GDL) the former and Graph of All Links (GAL) the latter; Hu and Zhu (2009) used the terminology “L space” for the network of container services and “P space” for the multigraph network. The second approach should be clearly preferred for a connectivity analysis, because a direct connection between a pair of ports enabled by a common service string should be of course considered. However, there are other applications wherein the former aspect is preferable, e.g. a vulnerability analysis wherein interest is in identifying single critical links of strings. Interestingly, to the authors' knowledge, none has applied for the calculation of network indicators the hyperpath approach by Spiess and Florian (1989), see also Cascetta (2009), Ortuzar and Willumsen (2011), that was revealed to be important in modelling container network by Bell et al. (2011) and Bell et al. (2013).
- *impedance (weight) of network links and nodes.* As recalled in Section 2.4.1, basic graph metrics deal with only the cardinality of backward and forward stars, i.e. links are all equal in their contribution to the centrality of a node. Of course, links can be also weighted, e.g. leading to the node strength metric (5): in this respect, many weights are applied in the literature, including deployed capacity in TEU, service

frequency, sailing time and so on.

- *geographical and time coverage.* Due to the inherent difficulty in collecting comprehensive worldwide data, studies often focus only on some areas (e.g. a specific world region) or on key ports. Notwithstanding, as recalled later, there are some studies dealing with the entire worldwide container network. For the same reason, many studies refer to analyses related to a specific year, and only few studies investigate worldwide evolution over time of container networks.

Earlier literature reviews on this field are available and consolidated, see amongst others Wang and Cullinane (2008), Cullinane and Wang (2009), Ducruet et al. (2010), Tran and Haasis (2014) and Tovar et al. (2015), who also provide an interesting review of papers applying graph theory to analyse container worldwide networks along the classifications described above.

Hu and Zhu (2009) built a worldwide container network based on a Containerisation International database and adopted both network approaches recalled above. Links in the network were considered either unweighted or weighted based on the number of direct services between ports, i.e. not accounting for capacity or frequency. Indicators presented in Section 2.4.1 were calculated, and they found that the worldwide container network is a small world network, and that degree centrality metrics follow truncated power-law distributions in the L space and an exponential decay distribution in the space P. Also, the container network showed a hierarchy structure and rich-club phenomenon, i.e. presence of a small subset of crucial nodes. Centrality measures were found to have strong correlations with each other. No policy or market indications are provided, being the study more focussed only on the topological aspect of the network. Similar results on the network structure of container worldwide flows are reported by Kaluza et al. (2010) and are a clear consequence of the prominent presence of transshipment, mainly hub and spoke.

Relevant papers dealing with worldwide container network analysis are by Angeloudis et al. (2007), Wang (2008), Pais et al. (2012), Gonzalez et al. (2012), and Ducruet and Notteboom (2012). Ducruet and Zaidi (2012) dealt with the role of both container hubs and regional ports, with an application of a  $k$ -clusters approach to identify relevant sub-networks. Ducruet et al. (2014) applied degree and betweenness indicators to analyse changes in the maritime container networks between 1996, 2006 and 2011. Kosowska-Stamirowska et al. (2016) analysed changes in the maritime network structure between 1890 and 2000, by means of metrics calculated on an unweighted maritime network based on historical database of worldwide merchant vessel movements. Kutin et al. (2018a)

Other studies offered a worldwide analysis limited only to most relevant ports and trade lanes. Lam and Yap (2011) investigated the relationship between the number of container vessel calls and their capacity in selected ports. Kang et al. (2014) calculated four centrality measures for top 5 container ports from 2006 to 2011. Mengqiao et al. (2015) applied degree and betweenness to the network of main trade routes between world regions. Wang and Cullinane (2016) applied port centrality and betweenness measures to 39 worldwide ports on a network weighted by the weekly transportation capacity deployed by top 20 liner shipping companies, presenting a correlation analysis of centrality measures with port throughput.

Chen et al. (2016) built a network of direct links amongst the first 100 container ports in the world, each link representing the traffic in TEU between these ports. Kutin et al. (2018b) collected a dataset of 153 ports from 50 countries, including statistics for more than six thousand maritime routes in 2014.

Many papers deal with case studies related to limited geographical areas. Amongst others, Low et al. (2009), Ducruet et al. (2010b), Tang et al. (2011), and Song et al. (2019) applied degree and betweenness metrics to some Asian ports. Mou et al. (2018) proposed an application to the so-called silk road along the Europe-Far East trade lane. McCalla et al. (2005) focussed on Caribbean ports. Ducruet et al. (2010) analyse a network of container services in the Atlantic for years 1996 and 2006 based on AIS data, adopting the multigraph approach. Tran and Haasis (2014) propose an empirical analysis of the container liner shipping network on the East-West corridor for the period 1995-2011. Calatayud et al. (2017) introduced trade strength relationship measures in addition to the usual metrics calculated on network of container services, with an application to port connectivity in America. Earlier studies on the Mediterranean are reported amongst others by Notteboom (2010), Kitsos (2014) and Elsayeh (2015). Varan and Cerit (2014) dealt with Turkish ports. Arvis et al. (2019) developed an analysis of container services in the Mediterranean, leveraging the LSCI and centrality metrics. Elbayoumi and Dawood (2016) dealt with Middle Eastern ports.

Other relevant contributions are reported in the special issue of Transportation Research Part E: freight and logistics on *Maritime logistics and port connectivity in the globalised economy* (2016). Variation on the theme include Leicht and Newman (2008) and Kaluza et al. (2010), who applied clustering algorithms at the port level, based on the metrics reported in Section 2.4.1, aim to identify areas of port competition and cooperation. Sys (2009) analysed the degree of concentration by estimating the following coefficients: the HHI, the Lorenz curve and the Gini coefficient as well as the Hymer–Pashigan index of market share instability. The multiple linkage analysis by Cullinane and Wang (2012) provides insights on the hierarchical configuration of the container port market. Ducruet (2013) and Ducruet (2017) investigated maritime networks considering multiple commodities altogether, finding the distribution of maritime traffics among ports to be influenced strongly by the concerned commodity diversity. Lange and Bier (2019) propose a Principal Component Analysis to cluster graph nodes based on a set of graph theory metrics. The HHI index is widely applied in the sector: amongst only the most recent papers, it is worth citing Goulielmos (2017), Hanafy et al. (2017), and Haralambides (2019). Hirata (2017) questioned the relationship between HHI and freight rates, finding absence of correlation between the two variables.

## 2.5 Literature review findings

The analysis of the available contributions dealing with the aim of the thesis focused on four main line of research:

1. National and international DSSs;
2. Shortest paths in multimodal freight networks;
3. Maritime (worldwide) container service network models;
4. Network analysis and port centrality measures.

A first review of the most relevant freight models developed for public authorities reveals that even if they are comprehensive models (representing a wide range of transport modes) they usually resort of many simplifications preventing adequate realistic results. Indeed, their underlying supply models, intended to evaluate performances and externalities of freight transport system, are largely simplified. These simplifications might lead to considerable modelling errors in key policy applications, for instance in the identification of the catchment areas of ports and in the analysis of competition amongst freight modes.

To properly evaluate the performances of a freight transport system it is necessary to take into account that the costs associated to each path are not additive (e.g. dwell times of freight road vehicles, non-linear decreasing distance fees). This issue has been faced by some scholars, who tried to overcome this issue, limited to the road and maritime transport. Indeed, none of the existing research deal with the non-additivity of key performance variables in the rail freight supply mode.

Also, in the literature, there are freight models developed specifically for container transport system. An innovative approach, proposed by Bell et al. (2011), transfer the frequency-based transit assignment method of Spiess and Florian (1989) to global maritime container assignment model. However, the application of the hypergraph-based approach to model global container service networks has not been sufficiently investigated in the literature.

Usually, maritime container service networks are modelled by implementing classic approaches (i.e. *L-graphs* or *P-graphs*) already consolidated in the literature. Also, they usually leverage simplified and/or partial representations of the worldwide maritime container service networks.

The aforementioned approaches are usually adopted to assess port centrality in global networks by applying well-known graph metrics. However, given the inherent nature of container services, it could be worth assessing the potential benefits of a hypergraph-based approach in port centrality evaluation.

### 3 Rail freight supply modelling

The modelling of multimodal freight transport system is a very complex task. Usually, the available models resort of simplifications preventing adequate results.

A recognized drawback, also hindering effective implementation of decision support systems for freight planning and policymaking, is the absence of approaches and tools to deal with the non-additivity of key performance variables of rail freight transport, namely the unit cost of transport and the total capacity of a freight train. In fact, both depend upon best/worst rail freight link variables across a path, a circumstance that prevents applying standard shortest path algorithms.

In the present chapter the methodological advances achieved in modelling landside rail freight networks are illustrated and the novel approach developed in the context of the doctoral study<sup>4</sup> is described.

For the sake of completeness, an application to a new scheme of incentives to railway undertakings – chosen for better highlighting the features and the effectiveness of the proposed approach – is also presented.

#### 3.1 Motivation and background

Rail freight transport is a key component of many intermodal container-related transport chains worldwide. Top European ports leverage arrival/departure of dozens of trains/day to/from their terminals to enlarge their catchment areas. Improving railway connections of container ports is also a key policy strategy for the European Union and also for Italy, as reported in the official Italian planning documents of the *Piano strategico nazionale della portualità e della logistica* (2015) and of *Connettere l'Italia* (2016). By way of example, the port of Trieste in Italy managed to double its railway traffic from 2015 to 2019, reaching the threshold of 10.000 trains/year, and its container port throughput also increased remarkably.

Also, the evolution of ocean carriers from simple maritime carrier companies (*merchant haulage*) towards integrated door-to-door freight transport providers (*carrier haulage*) has been almost completed for all top worldwide companies. Interestingly, two main phases can be identified in this process of vertical integration (Bologna, 2011): in a first stage, ocean carriers were buyers of rail freight services from landside carriers and/or multimodal transport operators (MTO), whilst in a more recent phase at least the very top ocean carriers have started making railway services by their own. After the pioneering experience of *ERS Railways* by the Maersk group in early '2000, the most relevant example of this approach is by MSC, that recently established a multimodal transport operator (*MedLog*) and a railway company (*MedWay*) operating services in Italy.

---

<sup>4</sup>Marzano, V., Tocchi, D., Papola, A., Aponte, D., Simonelli, F., & Cascetta, E. (2018). *Incentives to freight railway undertakings compensating for infrastructural gaps: Methodology and practical application to Italy*. *Transportation Research Part A: Policy and Practice*, 110, 177-188.

Apart from relevant policy implications, this significant market evolution poses interesting research questions. In the light of this thesis, two aspects are worth mentioning:

- from an ocean carrier perspective, the optimization of the overall door-to-door intermodal cost of container is a key factor to improve transport solutions;
- from a planning perspective, decision support systems for planning and policymaking should be capable to model jointly maritime and landside container networks.

In both cases, whilst supply modelling of road freight transport offers already consolidated and effective approaches, the same does not occur for rail freight transport, which is the subject investigated in the next sections. In particular, Sections 3.2 and 3.3 deal with a novel rail supply freight model capable to interact with the port-to-port supply model presented in Chapter 4 with adequate landside railway supply models.

### 3.2 The proposed shortest unit cost path algorithm for rail freight

In general, four main infrastructural characteristics affect operations of freight railway undertakings: permissible train weight per axle and/or per train/meter, permissible maximum train length, loading gauge, slope (Woodburn, 2011; Islam and Mortimer, 2017). Some of these characteristics lead to hard constraints (e.g. a maximum train length limit) whilst others to soft constraints, possibly surmountable with *ad hoc* technical solutions usually at the price of higher operational costs. For instance, high-cube containers can be transported with standard freight railcars (height over tracks of 1175 mm) on railway lines with at least a so-called PC45 loading gauge (UIC, 2006), with more expensive low-loader railcars (height over tracks of 945 or 825 mm) with a PC22 loading gauge, and cannot be transported by rail if the loading gauge is lower than PC22. Similarly, a slope can be climbed by appropriate dimensioning of the overall power traction, i.e. by choosing the appropriate type and number of locomotives. Clearly, infrastructural characteristics impact differently on the type of freight train  $t$ : for instance, freight trains carrying industrial bulk goods are usually shorter and heavier, thus affected by weight and slope constraints, whilst intermodal trains are longer and lighter, thus depending more upon length and loading gauge limits.

Overall, in the light of railway undertakings, the current infrastructural characteristics along a path  $k$  connecting a pair of stations  $o$  and  $d$  influence two key factors for any freight trains of type  $t$ : the total cost  $c^{tot}_{tk}$  to operate the train  $t$  on the path  $k$  and the train payload capacity  $cap_{tk}$ , the latter expressed in tons for all freight train types and also in terms of number of intermodal units (or TEUs) for intermodal trains, which represents the main focus for this thesis. Consistently, the following unit weight cost can be defined:

$$cW^{unit}_{tk} = c^{tot}_{tk} / cap_{tk} \quad (3.1)$$

In practice, given the cost structure faced by railway undertakings (Janic, 2007 and 2008; Woodburn, 2011), the total cost  $c^{tot}_{tk}$  to operate a freight train increases sub-linearly with respect to its capacity  $cap_{tk}$ , that is augmenting train capacity enables potential economies of

scale. Thus, freight railway undertakings try to maximize train capacity – consistent with the demand level on  $k$  to achieve a satisfactory load factor – given the constraints imposed by the infrastructural characteristics along  $k$ .

### 3.2.1 Calculation of train capacity

The calculation of the train capacity requires firstly modelling the freight railway network by means of a graph  $G \equiv \{A, N\}$  encompassing two ordered sets of arcs  $A$  and nodes  $N$  respectively. An arc  $a \in A$  represents a portion of rail infrastructure with homogeneous characteristics: loading gauge  $g_a$ , maximum permissible train weight  $w_a$  (usually expressed in tons per train/meter), maximum slope  $s_a$ , maximum permissible train length  $l_a$ . Let  $k$  be a rail path, that is an ordered sequence of arcs connecting a pair of stations  $o$  and  $d$  with  $o, d \in N$ . The infrastructural performance of a path  $k$  is determined by the corresponding worst performance of its arcs, i.e.  $w_k = \min_{a \in k} \{w_a\}$ ,  $g_k = \min_{a \in k} \{g_a\}$ ,  $s_k = \max_{a \in k} \{s_a\}$ ,  $l_k = \min_{a \in k} \{l_a\}$ .

The payload capacity  $cap_{tk}$  of a train of given type  $t$  on the path  $k$  is a function of the infrastructural characteristics of the path  $k$ , plus some train characteristics, namely: number of locomotives  $n_{loc}$ , average locomotive weight  $w_{tloc}$  and length  $l_{tloc}$ , average total weight  $w_{tcar}$ , unladen weight  $w_{tcar}$  and length  $l_{tcar}$  of a freight railcar. A necessary condition for the train  $t$  to operate on  $k$  – that is to have a payload capacity – is the path loading gauge  $g_k$  to be greater than the minimum loading gauge  $g_t^{min}$  needed by train  $t$ , i.e.  $g_k \geq g_t^{min}$ . Then, the maximum train length  $l_{tk}$  and weight  $w_{tk}$  – and the corresponding payload capacity  $cap_{tk}$  – should be determined by recognizing which characteristic between length and weight is the prevailing bottleneck. For this aim, if the train length were the maximum allowed by  $k$ , i.e.  $l_{tk} = l_k$ , the corresponding maximum number of freight railcar would be:

$$n_{cars} = \text{int} \left[ \frac{l_k - n_{loc} \cdot l_{tloc}}{l_{tcar}} \right] \quad (3.2)$$

yielding the following candidate train weight:

$$w_{tk}^* = n_{loc} \cdot w_{tloc} + n_{cars} \cdot w_{tcar} = n_{loc} \cdot w_{tloc} + \text{int} \left[ \frac{l_k - n_{loc} \cdot l_{tloc}}{l_{tcar}} \right] \cdot w_{tcar} \quad (3.3)$$

In fact, the actual maximum permissible weight of the train  $t$  might be lower than the candidate weight (7), because of possible attainment of an upper bound imposed by the infrastructural characteristics along path  $k$ , consistent with the following two constraints:

- the former is that  $w_{tk}$  should be consistent with the maximum permissible weight per train/meter on  $k$ , that is:

$$w_{tk} \leq w_k \cdot l_{tk} \quad (3.4)$$

- the latter depends upon the maximum slope  $s_k$  and on the corresponding applied tractive force, function of the number of locomotives  $n_{loc}$ , of their adhesive weight

and of other mechanical characteristics (e.g. the adhesion coefficient). From a practical perspective, the network rail infrastructure manager usually provides a relationship  $w_{tk}(s_k)=\Psi_t(s_k)$  expressing the maximum weight for a train of type  $t$  to operate on  $k$  with a single locomotive given the slope  $s_k$ . In the case of more locomotives  $n_{loc}$ <sup>5</sup>, a proper adjustment factor  $\phi(n_{loc})$  is also provided, yielding:

$$w_{tk} \leq w_{tk}(n_{loc}, s_k) = \Psi_t(s_k) \cdot \phi(n_{loc}) \quad (3.5)$$

As a result, if the candidate weight (3.3) is lower than the minimum between (3.4) and (3.5), that is:

$$w_{tk}^* \leq \min\{w_k \cdot l_{tk}, \Psi_t(s_k) \cdot \phi(n_{loc})\} \quad (3.6)$$

the maximum train length  $l_{tk}=l_k$  is the bottleneck for the maximum train weight, thus given by (3.3), i.e.  $w_{tk}=w_{tk}^*$ . Vice versa, if (3.6) does not hold, the bottleneck is represented by weight limits. In this case, the maximum train weight is given by:

$$w_{tk} = \min\{w_k \cdot l_{tk}, \Psi_t(s_k) \cdot \phi(n_{loc})\} \quad (3.7)$$

and the corresponding number of railcars and maximum train length are respectively:

$$n_{cars} = \text{int} \left[ \frac{w_{tk} - n_{loc} \cdot w_{tloc}}{w_{tcar}} \right] \quad (3.8)$$

$$l_{tk} = n_{cars} \cdot l_{tcar} + n_{loc} \cdot l_{tloc} = \text{int} \left[ \frac{w_{tk} - n_{loc} \cdot w_{tloc}}{w_{tcar}} \right] \cdot l_{tcar} + n_{loc} \cdot l_{tloc} \quad (3.9)$$

Finally, once determined  $w_{tk}$  and  $l_{tk}$  in accordance with the above, the resulting train payload capacity  $cap_{tk}$  is given by:

$$cap_{tk} = \begin{cases} w_{tk} - n_{cars} \cdot w_{tcar} - n_{loc} \cdot w_{tloc} & \text{if } g_k \geq g_t^{\min} \\ 0 & \text{otherwise} \end{cases} \quad (3.10)$$

Overall, equations (3.2) - (3.10) allow calculation of the payload capacity  $cap_{tk}$  as a function of the infrastructural path characteristics ( $g_k, l_k, w_k, s_k$ ) and train characteristics ( $n_{loc}, w_{tloc}, w_{tcar}, l_{tcar}$ ). Obviously, the same procedure described above can be repeated under the assumption of optimal (i.e. EU-standard) infrastructure performance, yielding the optimal payload capacity  $cap_{tk}^{opt}$  on path  $k$ .

### 3.2.2 Calculation of train cost

The cost to operate a freight train of type  $t$  on path  $k$ , in the light of railway undertakings, can

---

<sup>5</sup> It is worth noticing that the upper bound for the number of locomotives is very low in Europe – usually 2 in many circumstances and no more than 3 in particularly steep railway stretches.

be calculated, similarly for the capacity, as a function of the path/infrastructure characteristics ( $g_k, l_k, w_k, s_k$ ) and the train performances ( $n_{loc}, w_{iloc}, w_{tcar}, l_{tcar}$ ) defined in the previous section. In general, the procedure illustrated is general and allows embedding a variety of assumptions on the calculation of train costs. By way of example, the following cost components can be considered:

- cost of train driver(s)  $c_{driv}$ . This cost is assumed independent of the number of locomotives  $n_{loc}$ , and is determined based on the number of drivers  $n_{driv}$  imposed by the current regulations on railway transport (e.g.  $n_{driv}=2$  in Italy). Thus, the total cost is  $n_{driv} \cdot tt_k \cdot c^h_{driv}$ , wherein  $c^h_{driv}$  is an hourly cost rate, based on the yearly salary of the driver, on the average number of working hours per year, and  $tt_k$  the total travel time in hours on the path  $k$ ;
- cost of locomotives  $c_{loc}$ , given by  $n_{loc} \cdot tt_k \cdot c^h_{loc}$ , being  $c^h_{loc}$  the hourly cost of the locomotive, determined based on the initial cost of the locomotive (purchase or lease), the maintenance costs, the number of years of operation and the average number of working hours per year;
- cost of rolling stock  $c_{car}$ . With the same rationale of the cost of locomotives, it is given by  $n_{cars}(l_k) \cdot tt_k \cdot c^h_{car}(g_k)$ , being  $n_{cars}(l_k)$  the number of railcars allowed by the train length  $l_k$ , calculated as described in Section 3.2.1, and  $c^h_{car}(g_k)$  the hourly cost of the rolling stock, determined as a function of the loading gauge  $g_k$ . Indeed, provided that  $g_k \geq g_t^{min}$  (see Section 3.2.1), the loading gauge  $g_k$  determines also the height  $h_k$  of the freight railcars (normal or low-loader) needed to operate the train  $t$  on  $k$ ;
- energy consumption and toll payment to the railway network operator  $c_{netw}$ . A proper calculation, in practice cumbersome and country-dependent, can be reasonably approximated by a toll per km  $\tau^{km}(w_{tk})$ , usually function of  $w_{tk}$ , yielding  $c_{netw} = \tau^{km}(w_{tk}) \cdot td_k$ , being  $td_k$  the travel distance between  $o$  and  $d$  along  $k$ . Notably, the toll usually incorporates also the energy consumption, which thus becomes, in the light of the railway undertaking, a flat cost per km irrespective of the actual traction energy consumed corresponding to the kinematic profile of the train;
- other fixed costs  $c_{fix}$ , occurring irrespectively of the train characteristics and of the path  $k$ : they include, for instance, shunting costs at origin and/or destination, or other fixed administrative costs.

Overall, the total cost of running the train  $t$  of type  $r$  on path  $k$  between  $o$  and  $d$  is given by the sum of the above cost components, yielding:

$$\begin{aligned} c^{tot}_{tk} &= c_{driv} + c_{loc} + c_{car} + c_{netw} + c_{fix} = \\ &= n_{driv} \cdot tt_k \cdot c^h_{driv} + n_{loc} \cdot tt_k \cdot c^h_{loc} + n_{cars}(l_k) \cdot tt_k \cdot c^h_{car}(g_k) + \tau^{km}(w_{tk}) \cdot td_k + c_{fix} \end{aligned} \quad (3.11)$$

The previous equation obviously holds only if  $\rho_{tk}=1$ , i.e. loading gauge allowing that type of train. Equation (3.11) can be also conveniently rewritten in the following form, considering that a commercial speed  $v_t$  can be defined for the train  $t$ , such that  $td_k = v_t \cdot tt_k$ , yielding:

$$c^{tot}_{tk} = [n_{driv} \cdot c^h_{driv} + n_{loc} \cdot c^h_{loc} + n_{cars}(l_k) \cdot c^h_{car}(g_k) + \tau^{km}(w_{tk}) \cdot v_t] \cdot tt_k + c_{fix} \quad (3.12)$$

### 3.2.3 Calculation of shortest unit cost path

Calculation of the shortest unit cost path requires first noticing that a higher number of locomotives does not necessarily reduce the unit cost (3.1). This might occur, for instance, on a path with high slope and limited maximum train length: in this case  $n_{loc}=1$  would likely allow operating a train with – let us assume for the sake of argument – a length close to the maximum train length because of the weight limits imposed by the slope, whilst  $n_{loc}=2$  would overcome the weight limit, but the train capacity would not increase appreciably because of the attainment of the length limit. As a result, the increased cost for  $n_{loc}=2$  would not be compensated by a sufficient increased capacity. Overall, the procedure illustrated in the following should be repeated for each number of locomotives  $n_{loc}$ , to find the global shortest path minimizing (3.1).

The methodological complexity in the calculation of the proposed incentive is that both the total cost  $c^{tot}_{tk}$  and the capacity  $cap_{tk}$  depend upon the performance of the worst link for each of the relevant infrastructural characteristics (loading gauge, length, weight, and slope). As a result, path costs (3.12) are non-additive, i.e. they cannot be expressed as the sum of corresponding link costs: this prevents using standard shortest path algorithms to calculate the incentive. To overcome this issue, a brute force approach could be adopted, by enumerating all paths for each o-d pair and then calculating the corresponding train capacity and costs via the methods described in Sections 3.2.1 and 3.2.2 respectively; this is practically infeasible also for small-size networks.

Fortunately, the specific nature of the problem allows setting a more efficient solution algorithm for effective calculation of the incentive. Indeed, given a train type  $t$  and a number of locomotives  $n_{loc}$ , if the infrastructural characteristics of each link (slope, loading gauge, weight, and length) were homogeneous (i.e. the same) across all links of the network, also the train performances (length, weight, capacity) would be homogeneous, i.e. irrespective of the specific path  $k$ . In this case, the terms in square brackets in equation (3.12) would not depend anymore on the specific path  $k$ , and hence they could be used to define the cost of all the links of the network. In other words, a feasible conceptual vehicle to quantify the infrastructural gap is to calculate the lowest unit cost paths for each combination of fixed infrastructural characteristics, and then to find the lowest of the lowest unit cost paths across all those combinations of parameters.

Obviously, working with fixed infrastructural characteristics means dealing with homogeneous subnetworks characterized exactly by such infrastructural performances. Thus, this approach can be operationalized by discretizing the relevant infrastructural characteristics based on a limited number of thresholds. In formal terms, let  $L$  be a set of  $n_l$  train length thresholds,  $G$  a set of  $n_g$  loading gauge thresholds,  $S$  a set of  $n_s$  slope thresholds, and  $W$  a set of  $n_w$  permissible weight thresholds. By way of example, the train length might be discretized based on  $n_l=5$  thresholds  $L=\{750 \text{ m}, 600 \text{ m}, 500 \text{ m}, 400 \text{ m}, 300 \text{ m}\}$ . Consistently, a set  $\Omega$  with cardinality  $n_\omega=n_g \cdot n_l \cdot n_s \cdot n_w$  encompassing all possible combinations of such thresholds can be defined. The generic combination  $\omega_i \in \Omega$  defines an appropriate bound for infrastructural characteristics (an upper bound for the slope and a lower bound for all the others), let  $w_{\omega_i} \in W$  the bound for the weight,  $l_{\omega_i} \in L$  for the length,  $s_{\omega_i} \in S$  for the slope and  $g_{\omega_i} \in G$  for the loading

gauge. In turn, each  $\omega_i \in \Omega$  identifies a subset  $A_{\omega_i} \subseteq A$  of arcs of the network such that  $a \in A$  belongs to  $A_{\omega_i}$  if  $w_a \geq w_{\omega_i}$ ,  $g_a \geq g_{\omega_i}$ ,  $l_a \geq l_{\omega_i}$  and  $s_a \leq s_{\omega_i}$ . Clearly, the resulting sets  $A_{\omega_i}$  will overlap, because each link  $a \in A$  will belong to all  $A_{\omega_i}$  sets satisfying the conditions  $w_a \geq w_{\omega_i}$ ,  $g_a \geq g_{\omega_i}$ ,  $l_a \geq l_{\omega_i}$  and  $s_a \leq s_{\omega_i}$ .

In practice, the combination of infrastructural characteristics  $\omega_i$  identifying each subnetwork  $A_{\omega_i}$  leads to specific performances for each train type  $t$  – i.e. a payload capacity  $cap_t$ , a total cost  $c^{tot}_t$  and hence a unit cost  $c^{unit}_t$  – calculated by means of the procedure illustrated in the previous sections. Importantly, letting  $A^*_{\omega_i}$  be the subset including all links  $a$  such that  $w_a = w_{\omega_i}$ ,  $g_a = g_{\omega_i}$ ,  $l_a = l_{\omega_i}$  and  $s_a = s_{\omega_i}$  and  $A_{\omega_i} - A^*_{\omega_i}$  be the subset including all links for which at least a characteristic is better than the one defining  $\omega_i$ , the shortest path  $ks^{od}_{\omega_i}$  in the subnetwork  $A_{\omega_i}$  for a given  $od$  pair could possibly include only links belonging to  $A_{\omega_i} - A^*_{\omega_i}$ , i.e. none of the links  $A^*_{\omega_i}$  representing the bottleneck in accordance with the infrastructural characteristics underlying  $\omega_i$ . In this case, the cost associated to that path would be wrong, because carried out on the basis of infrastructural bottlenecks not actually present in that path. However, this is not an issue, because that specific path  $k$  will be also present in another combination  $\omega_j \neq \omega_i \in \Omega$  representing exactly the combination of infrastructure performance characterizing  $k$ . As a result, each path  $k$  in the network will be appropriately processed with all correct associated infrastructural characteristics by a combination in  $\Omega$ , and thus a comparison of the shortest paths across all combinations will ensure finding in any case the correct global shortest path.

That said, the algorithm that operationalizes the calculation of the infrastructural gap consists of the following steps, to repeat for each train type  $t$  (for the sake of simplicity, the loop across train types  $t$  is omitted) and for each number of locomotives  $n_{loc}$ . Please also notice that, in accordance with the above rationale, path  $k$  subscript is replaced with a reference to combination  $\omega_i$ , having assumed homogeneous network characteristics for  $\omega_i$  irrespective of the specific path  $k$ :

```

 $\forall \omega_i \in \Omega$ 
  If  $g_k = g_{\omega_i} \geq g_t^{min}$ 
     $\forall n_{loc}$ 
      set  $l_{tk} = l_{\omega_i}$ 
      calculate  $n_{cars}$  (eq. # 3.2) and candidate weight  $w^*_{t\omega_i}$  (eq. #3.3)
      if  $w_{t\omega_i} \leq \min\{w_{\omega_i} \cdot l_{tk}, \Psi_t(s_{\omega_i}) \cdot \phi(n_{loc})\}$  (eq. #3.6)
         $w_{t\omega_i} = w^*_{t\omega_i}$ 
      else
         $w_{t\omega_i} = \min\{w_{\omega_i} \cdot l_{tk}, \Psi_t(s_{\omega_i}) \cdot \phi(n_{loc})\}$  (eq. #3.7)
         $l_{t\omega_i} = \text{int} \left[ \frac{w_{t\omega_i} - n_{loc} \cdot w_{tloc}}{w_{tcar}} \right] \cdot l_{tcar} + n_{loc} \cdot l_{tloc}$  (eq. #3.9)
      end if
      calculate payload capacity  $cap_{t\omega_i}$  (eq. #3.10)
     $\forall \text{ arc } a \in A$ 

```

```

if  $a \in A_{\omega_i}$ 
     $c_a = [n_{driv} \cdot c^h_{driv} + n_{loc} \cdot c^h_{loc} + n_{cars}(l_{t\omega_i}) \cdot c^h_{car}(g_{\omega_i}) +$ 
         $+ \tau^{km}(w_{t\omega_i}) \cdot v_t] \cdot tt_a$ 
else
     $c_a = \infty$ 
end if
 $\forall$  o-d pair  $od$ 
    calculate the shortest paths  $ks^{od}_{\omega_i}$  based on arc costs  $c_a$ 
    add  $c_{fix}$  to get the shortest unit path cost
end if
 $\forall$  o-d pair  $od$ 
    calculate the shortest unit path as the minimum across  $n_{loc}$ 
 $\forall$  o-d pair  $od$ 
    Calculate the shortest path  $k^*$  of the shortest paths across all combinations6,
    yielding the optimal combination  $\omega^*$  and  $n_{loc}$  for that o-d pair, and calculate the
    current minimum unit cost on  $k^*$ 

```

In terms of calculation times, the only demanding step is, in principle, the first. However, calculation of shortest additive paths is a standard in most of transport applications and does not represent a noticeable computational bottleneck, leading to a sufficiently fast calculation procedure, as it will be shown in a real application to Italy in the next section.

### 3.3 An application: a new incentive scheme for rail freight undertakings

To highlight the viability of this procedure in support to a variety of planning and policymaking applications, this subsection illustrates an example related to a new incentive scheme for rail freight undertakings.

The rationale behind this new incentive is that, if a path  $k$  had optimal infrastructural conditions – i.e. EU targets for freight train performance on the core Trans-European Transport (TEN-T) network, given by 750 m length, 2000 tons weight and PC80 gauge – railway undertakings would achieve an optimal train capacity  $cap^{opt}_{tk}$  and a corresponding

---

<sup>6</sup> Importantly, a hierarchy can be defined across combinations: given two combinations  $\omega_i$  and  $\omega_j$ , it might occur that for each parameter  $\omega_i$  dominates  $\omega_j$  (e.g. same gauge, weight and length, lesser slope for  $\omega_i$  rather than for  $\omega_j$ ), or vice versa. It might be also obviously the case that none of the two dominates the other (e.g.  $\omega_i$  has better gauge but worse slope than  $\omega_j$ ), being therefore both Pareto-efficient combinations. In this respect, one might think of restricting the comparison only to all Pareto-efficient combinations for that o-d pair, that is discarding all existing (i.e. non-cost-infinite) dominated combinations, to speed up calculation times. However, it is easy to recognize through a counterexample that this would lead to errors: in fact,  $\omega_i$  dominating  $\omega_j$  implies  $|A_{\omega_j}| > |A_{\omega_i}|$ , i.e. the network associated with  $\omega_j$  contains more links than the network associated to  $\omega_i$ , being less restrictive in the parameters. As a consequence, the travel time associated with the shortest path of  $\omega_j$  might be much lower than the shortest path associated with  $\omega_i$ , thus compensating the fact that the parameters associated with  $\omega_i$  are better than  $\omega_j$ . By way of example,  $\omega_i$  might dominate  $\omega_j$  in terms of train length (e.g. 600 metres maximum against 500 metres), but the shortest path corresponding to  $\omega_i$  might imply a larger detour, resulting in a significantly larger travel time, avoided by a 500 metres train length shortcut related to the shortest path of  $\omega_j$ , such that to compensate the better performance offered by the parameter combination  $\omega_j$ .

total cost  $c^{tot,opt}_{tk}$ , such that:

$$cW^{unit,opt}_{tk} = c^{tot,opt}_{tk} / cap^{opt}_{tk} \leq cW^{unit}_{tk} \quad (3.13)$$

that is lower or at most equal to the actual cost currently faced on that path  $k$ . The above yields a straightforward definition of the *infrastructural gap* experienced by railway undertakings on path  $k$  as the difference  $cW^{unit,gap}_{tod}$  of unit weight cost between current and optimal conditions, that is:

$$cW^{unit,gap}_{tod} = \min_{k \in K_{od}} \{cW^{unit}_{tk}\} - \min_{k \in K_{od}} \{cW^{unit,opt}_{tk}\} \quad (3.14)$$

wherein  $K_{od}$  is the set of paths  $k$  available for the o-d pair  $od$ <sup>7</sup>. Equation (3.14) suggests the definition of an o-d based incentive to railway undertakings operating trains of type  $t$ , proportional to such infrastructural gap, as opposed to an incentive based on a watering-can principle, i.e. equal across o-d pairs irrespective of the infrastructural gap.

Importantly, incentives to railway undertakings are usually operationalized in the form of discounts to the network toll paid by railway undertakings to the railway network infrastructure manager, i.e. they are expressed in currency/(train·km). Thus, also the infrastructural gap should be expressed more conveniently in the same way. To do so, the unit weight cost (3.14) can be converted into a unit distance cost  $cd^{unit,gap}_{tod}$  through the following formula:

$$cd^{unit,gap}_{tod} = cW^{unit,gap}_{tod} \cdot cap^*_{tk} / td_{od} \quad (3.15)$$

being  $td_{od}$  the travel distance between  $o$  and  $d$  and  $cap^*_{tk}$  the capacity corresponding to the path  $k^*$  minimizing the  $cW^{unit}_{tk}$  cost. As a result, the proposed incentive can be expressed as a percentage  $\beta$  of the infrastructural gap in terms of unit distance cost (3.15), to account for possible budget constraints. In addition, EU regulations upper-bound the incentive (3.15) by a maximum percentage  $\alpha$  of the total toll·km  $t^{km}$ , hypothesized independent of the type of train  $t$ , yielding the following formula for the proposed unit incentive per o-d pair:

$$inc^{km}_{tod} = \min \{ \beta \cdot cd^{unit,gap}_{tod}, \alpha \cdot t^{km} \} \quad (3.16)$$

If the current infrastructural characteristics are so poor to prevent operating a given type of train – for instance, as mentioned before, a container freight train cannot be operated on a railway lines with less than PC22 loading gauge – the infrastructural gap (3.14) and thus the incentive (3.16) become infinite. This is correct, because no incentives can be provided in a situation wherein, whatever incentive, the railway undertaking would not be able to operate the train without infrastructural upgrading.

From a practical perspective, exactly as for the operationalization of the current incentive schemes, the provider of the incentive – usually the network rail infrastructure manager – will

---

<sup>7</sup> Notice that equation (3) reflects the fact that the path corresponding to the minimum cost might change between current and optimal conditions.

declare at the beginning of each year the amount of incentive provided to railway undertakings on each o-d pair. Based on this information, railway undertakings will be able to plan their business in fair and equitable market conditions, i.e. irrespective of the infrastructural gaps they will face. At the closure of the year, the network rail infrastructure manager will then provide the total amount of incentive to each railway undertaking, based on its actual amount of operated trains·km. A further nice feature of the proposed incentive is that, for the subsequent year, the calculation of the infrastructural gap (and thus of the incentive) can be updated, to account for the effects of possible infrastructural upgrading works. In this sense, the proposed incentive can be dynamically adjusted and eventually run out if the completion of the infrastructural upgrading will enable running EU-standard freight trains.

That said, the methodology described in Section 3.2 has been applied to a real case study in Italy, considering the entire national railway network and, without loss of generality, only intermodal freight trains as type  $t$ . The case study refers to year<sup>8</sup> 2015 in terms of network characteristics and freight demand. As illustrated in Figure 4, the characteristics of the rail network are very heterogeneous in terms of weight class, slope, loading gauge and train length, thus representing an ideal case study for the proposed methodology. The railway network has been modelled with a graph including 11270 links and 11408 nodes, of which 356 freight terminals/stations with intermodal traffic; overall, 2358 o-d pairs with intermodal trains have been considered. Infrastructural characteristics have been disaggregated, in accordance with the procedure illustrated in Section 3.2.3, into the following interval thresholds:

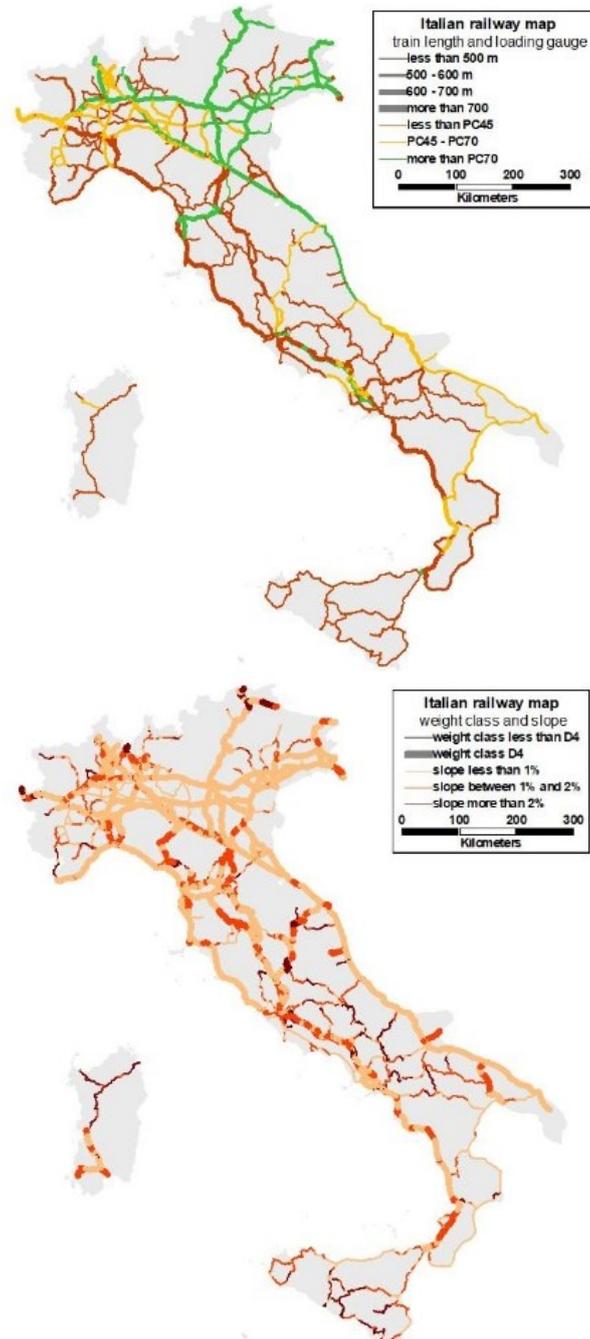
- length:  $n_l=5$ , set  $L \equiv \{700 \text{ m}, 600 \text{ m}, 500 \text{ m}, 400 \text{ m}, 300 \text{ m}\}$ ;
- loading gauge:  $n_g=4$ , set  $G \equiv \{\text{PC80}, \text{PC45}, \text{PC22}, \text{PC00}\}$ ;
- slope:  $n_s=4$ , set  $S \equiv \{10\%, 15\%, 21\%, 50\%\}$ ;
- weight:  $n_w=2$ , set  $W \equiv \{8.0 \text{ tons/m}, 2.0 \text{ tons/m}\}$ , with the threshold 8.0 tons/m corresponding to the EU standard for the core network, classified as D4/D4L according to the Italian railway network weight classification.

Overall,  $n=160$  combinations of infrastructural characteristics lead to the set  $\Omega$ .

---

<sup>8</sup> Year 2015 was the one with most recent data at the time of carrying out this part of the doctoral research.

Figure 4 : Characteristics of the Italian railway network in 2015: train length and loading gauge and weight class and slope.



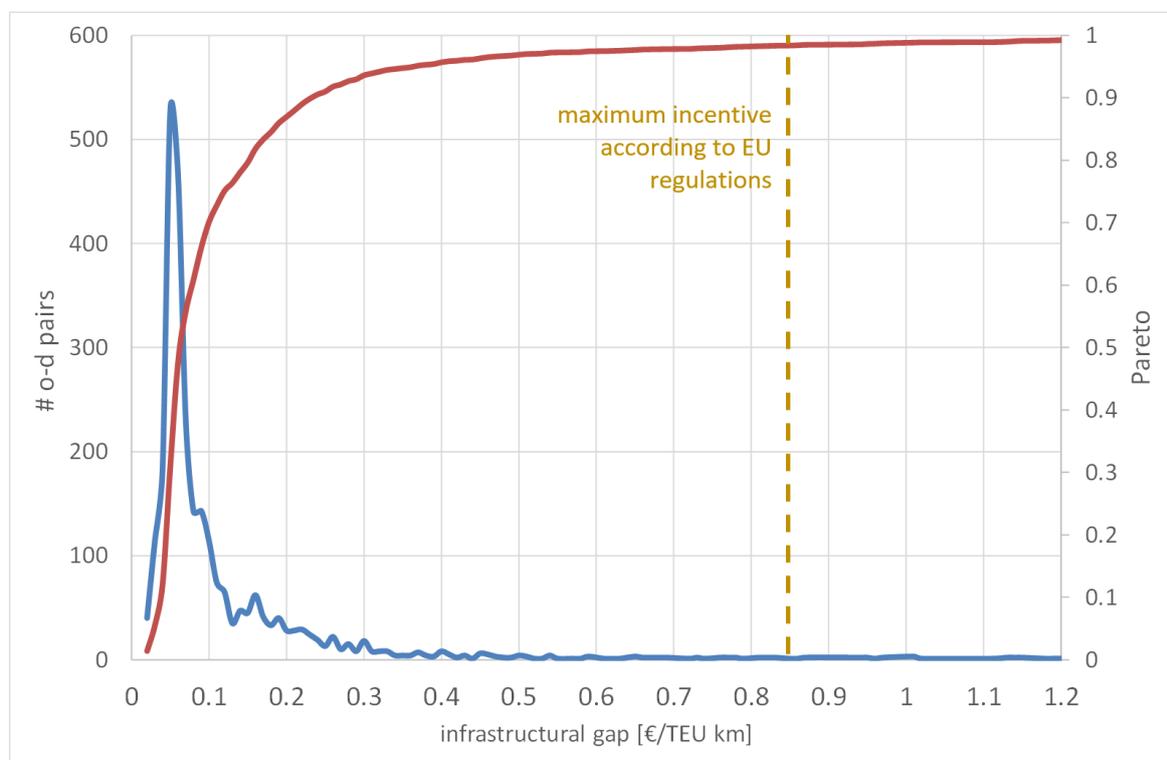
In terms of costs, in accordance with Section 3.2.3, the cost of drivers ( $n_{driv}=2$  according to the Italian regulations) is calculated by assuming  $c^h_{driv}=25$  €/h, corresponding to a gross yearly salary of 55.000 € and 2200 working hours/year. The cost of locomotives is calculated assuming  $c^h_{loc}=22.83$  €/h, corresponding to an initial cost of 4 M€ and 20 years' lifetime. Similarly, for the rolling stock it is assumed  $c^h_{car}=0.36$  €/h for a standard railcar (initial cost of 63.000 €) and  $c^h_{car}=0.43$  €/h for a low-loader railcar (initial cost of 75.600 €), both with 20 years' lifetime. The energy and toll costs are assumed  $\tau^{km}=3.26$  €/train·km (of which toll

accounts for 2.82 €/train·km), based on the average costs by the Italian national railway operator. For the sake of simplicity,  $\tau^{km}$  is assumed independent of the train weight. Finally, fixed costs include 2200 €/trip for shunting in origin and destination, plus 200 €/trip to account for administrative costs. In terms of typical intermodal train characteristics, the mass of the locomotive is assumed  $w_{loc}=106$  tons and the length  $l_{loc}=18$  m, the total weight of each freight railcar  $w_{car}=50$  tons (of which 17.5 tons of unladen weight) and a length of railcar of  $l_{car}=20$  m.

The procedure described in Section 3.2 has been coded in Visual Basic for Application, and the overall calculation times are less than 10 minutes on a standard laptop.

reports the distribution of the infrastructural gap, expressed in €/TEU·km, across o-d pairs, and its corresponding cumulative distribution<sup>9</sup>. For about 90% of o-d pairs the infrastructural gap is lower than 0.20 €/TEU·km and in almost all cases lower than the maximum incentive threshold allowed by EU regulations (30% of toll). Importantly, such infrastructural gap is only apparently low: a 0.20 €/TEU·km gap for an o-d pair with a 1000 km distance – this is the case, for instance, of an intermodal train connecting a Southern Italian port with an inland terminal in Northern Italy – would correspond to a total gap of around 200 €/TEU. Considering that the transport fare on that o-d pair is around 800 €/TEU (average market value in 2016), this means that the infrastructural gap might correspond to 25% of the service fare.

**Figure 5 : Distribution of the infrastructural gap across o-d pairs – intermodal freight transport**



<sup>9</sup> Results are presented in €/TEU·km rather than in €/train·km, because TEU-based calculations are common and more easily interpretable when dealing with combined transport. The conversion between the two measurement units is straightforward, based on the train capacity.

An estimation of the total amount of the proposed incentive concept is provided in Table 2, with a load factor of 85% to convert the number of trains into number of TEUs and using a conversion factor of 3 TEU/railcar and 12 tons/TEU. O-d pairs have been clustered accordingly with their number of trains/year; the corresponding percentage distribution of the total number of intermodal trains by cluster is reported in the first column of Table 1. The second and third columns report respectively the mean and the standard deviation of the infrastructural gap in €/TEU·km across o-d pairs in each cluster. As expected, the gap is lower for the cluster with the highest traffic share, i.e. one would expect that better infrastructural characteristics would lead to higher traffic.

**Table 2 : Yearly quantification of the infrastructural gap for intermodal transport**

# trains/year per o-d pair	# trains/year % distribution	infrastructural gap [€/TEU km]		smart incentive
		mean	st. dev.	M€/year
0-49	12%	0.14	0.23	23.28
50-99	5%	0.10	0.08	10.18
100-199	12%	0.14	0.27	23.29
200-499	12%	0.13	0.20	18.16
500 and more	58%	0.10	0.05	48.58
<b>total</b>	<b>100%</b>	<b>0.14</b>	<b>0.22</b>	<b>123.49</b>

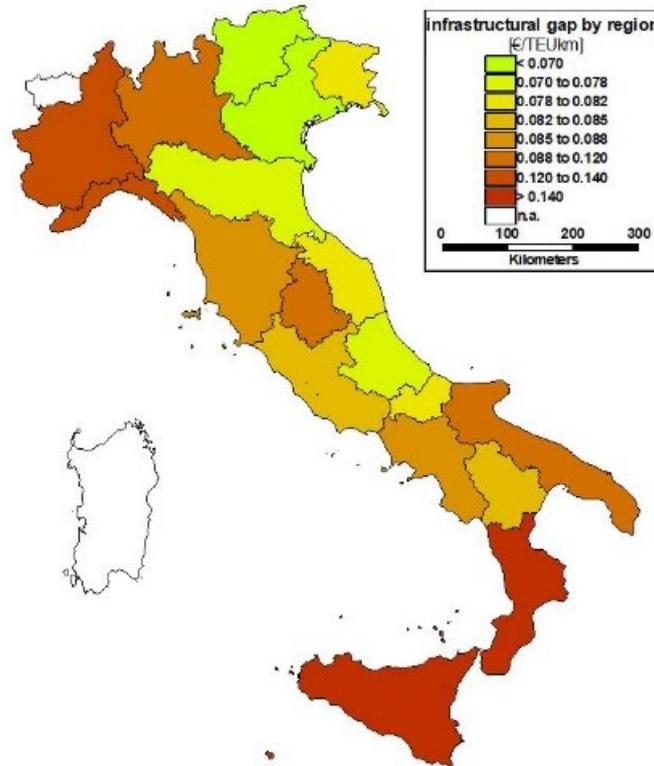
The total amount of such incentive is 123.49 M€/year, which can be interpreted as the cost of the “intermodal divide” of railway transport in Italy with respect to the optimal standards of the EU. In this respect, the overall contribution provided by the Italian Government in 2015 to sustain railway transport is 100 M€/year. Considering that the share of intermodal railway traffic, which calculations in Table 1 refer to, is slightly higher than 50%, a full coverage of the infrastructural gap would likely require more than doubling the currently allocated budget. This is obviously just a rough estimate, due on the one hand, on the cost assumptions made at the beginning of this section for the calculation of the incentive and, on the other, by assuming that the removal of the infrastructural gap would correspond to an increase of the demand such that the additional offered capacity would be saturated. However, the methodological approach and the results are still valid and interpretable: on one hand, the proposed approach can be easily and flexibly adapted to any changes/updates in the cost calculation; on the other, it can be interestingly embedded in demand models and decision support systems, such as those mentioned in Chapter 1 of the thesis.

Finally, Figure 6 visualizes the infrastructural gap: for this aim, letting  $R$  be the set of Italian regions<sup>10</sup> and  $t$  a subscript identifying intermodal trains, each pair of regions  $r, s \in R$  has been associated with an average infrastructural gap  $cd^{unit, gap}_{trs}$ , obtained as the average of the infrastructural gaps of type (4.15) between any pairs of stations  $o \in R$  and  $d \in S$ , calculated with the procedure illustrated in Section 3.2. In turn, each region  $r \in R$  has been associated with an average infrastructural gap  $cd^{unit, gap}_r$  calculated as the weighted average of the average

<sup>10</sup> The definition of infrastructural gap introduced in the thesis is on an o-d pair basis, thus its geographical distribution can be represented naturally by means of network-based indicators and maps. Results are presented here with less granularity because of data sensitivity constraints.

infrastructural gaps  $ca^{unit,gap}_{trs}$  with  $s \in \mathbb{R}$ , based on the total freight demand between regions available from Italian national official statistics. Figure 6 illustrates a map of the resulting geographical distribution of  $ca^{unit,gap}_{trs}$ , which highlights again the heterogeneity of the network performance and thus the need for a smarter o-d based incentive.

**Figure 6 : Distribution of the average regional infrastructural gap for Italian regions (only intermodal freight transport)**



The procedure illustrated in the previous sections allows calculating the shortest unit path cost on a railway network. This has direct effect on the research of this thesis, because such freight railway network modelling can be straightforwardly coupled with the maritime container supply model described in Chapter 4.

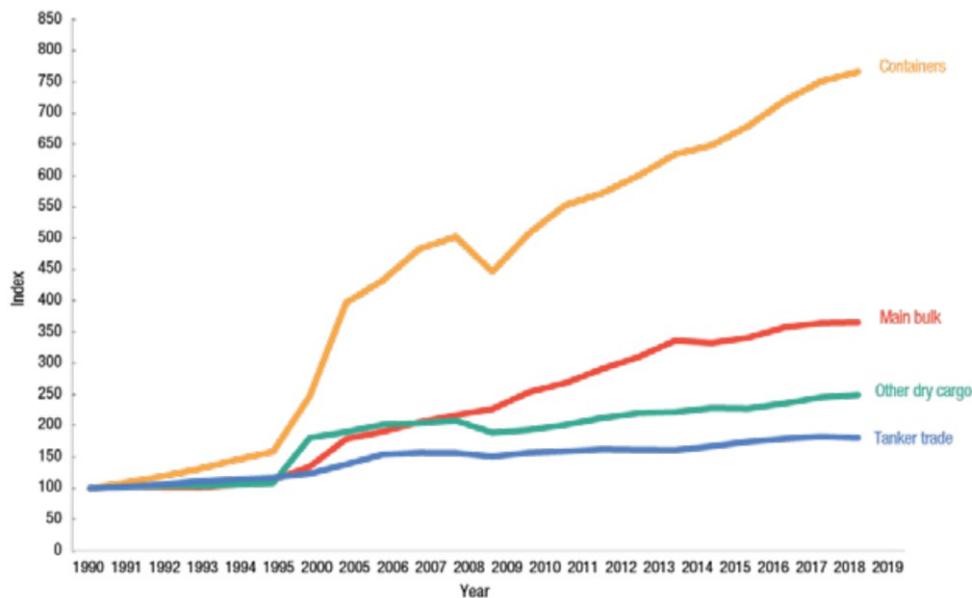
## **4 Modelling worldwide maritime container networks**

The methodology applied to build a supply model of worldwide maritime container network services and its practical implementation are illustrated in this Chapter. In detail, the study area encompasses the whole World, including 1196 ports belonging to 188 different Countries and the supply network is built considering the active container services operating from 2018 to 2020.

The chapter is structured as follows. First, Section 4.1 reports on a brief overview on the maritime container service market, with the objective to frame the reference context. The design of the database of liner shipping services, crucial to implement the model and run the applications, is outlined in Section 4.2 as well as the steps followed in gathering data and populating the database. Following Cascetta (2009), the transport supply model has been developed. Section 4.3 goes through all steps of this approach and describes all concerned methodological and operational issues for their application to the worldwide network of container services. The issue of effective linkage between port-to-port container model and the landside railway model (Chapter 3) is dealt in Section with the appropriate generalization of existing approaches in the literature. Finally, the last section illustrates the application of centrality metrics to the proposed hyperpath-based approach.

### **4.1 The container shipping industry**

Since its appearance on the international market, intermodal transport based on the so-called “container revolution” has undoubtedly boosted the expansion of the global economy, representing a key enabling factor for the whole globalization of world trade (Levinson, 2006 Khanna, 2015). Since the 1990, compared to the other market segments, the container sector played a crucial role in the growth of the maritime trade ( Figure 7 ).

**Figure 7: Development of international maritime trade by cargo type (Index: 1990 = 100)**

\*source: UNCTAD, Review of Maritime Transport, 2020.

Nowadays, more than 70% of worldwide general cargo and approximately the 60% of world seaborne trade is freighted in containers, accounting for 837 Million TEUs of global port container throughput in 2019 (Table 3).

**Table 3: global port throughput – source: Alphaliner Monthly Monitor, January 2020**

Global port throughput includes empty container and transshipment cargo handling

Million TEU by Region	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018E	2019E
China & HK	40.7	44.5	56.4	69.1	83.8	98.2	117.2	138.4	152.8	143.4	169.8	188.1	200.6	212.6	224.7	231.6	239.9	259.2	270.7	279.4
Other NE Asia	32.4	34.2	38.0	41.4	45.8	47.4	49.7	52.8	52.5	46.7	53.5	57.0	58.6	60.2	62.9	62.5	63.6	66.6	69.3	69.9
SE Asia	36.4	37.9	42.5	46.9	53.0	56.5	61.1	68.8	73.5	69.5	78.0	84.5	89.1	91.8	97.4	97.3	100.7	106.4	114.2	118.2
Middle East	9.3	10.4	11.9	14.5	17.8	20.1	21.6	24.9	28.4	28.5	30.6	32.8	35.4	35.2	37.0	38.3	39.4	40.0	39.5	38.3
S Asia	5.6	6.1	6.8	7.9	8.9	10.0	11.9	13.7	14.7	14.6	16.8	17.9	17.9	18.8	21.1	22.4	24.7	27.4	30.2	31.3
N Europe	33.6	34.6	37.3	40.4	45.2	49.0	53.3	60.3	61.5	51.3	58.1	62.2	62.8	64.0	67.3	65.9	67.5	71.1	73.9	76.2
S Europe	21.0	21.9	24.5	27.4	30.8	32.9	35.6	40.2	41.1	36.0	39.3	42.7	44.9	46.6	48.2	47.8	50.6	54.0	57.5	59.7
N America	33.0	33.3	35.9	39.8	43.3	46.8	49.1	50.0	47.7	41.8	47.1	47.7	49.0	50.0	52.0	54.2	54.7	59.1	62.4	64.3
C+S America	15.6	17.0	18.0	20.1	23.6	26.7	30.1	33.7	36.4	32.6	37.7	42.6	45.5	46.0	46.3	47.2	46.3	50.0	53.7	55.3
Africa	7.8	8.6	9.6	11.2	12.8	14.4	16.2	18.5	21.4	22.0	24.7	26.6	26.7	28.7	30.6	30.0	29.3	31.0	32.8	32.1
Oceania	5.5	5.4	6.1	6.7	7.4	7.6	8.1	8.9	9.7	9.1	9.9	10.4	10.7	10.9	11.3	11.6	11.9	12.6	13.2	12.9
Total	240.9	253.9	286.9	325.5	372.4	409.7	453.9	510.4	539.6	495.5	565.5	612.4	641.1	664.6	698.8	708.9	728.5	777.2	817.4	837.6
Growth (YoY)	12.2%	5.4%	13.0%	13.4%	14.4%	10.0%	10.8%	12.5%	5.7%	-8.2%	14.1%	8.3%	4.7%	3.7%	5.1%	1.4%	2.8%	6.7%	5.2%	2.5%

The unprecedented success of containerization leverages several factors. The most important likely relates to the standardised dimensions of maritime containers, which in turn implied standardisation of cellular ships and all equipment for container loading/unloading (e.g. spreaders) and landside transport (e.g. railcars and truck beds). This enabled a more efficient cargo handling in ports: containers can be loaded/unloaded, stacked and transported over long distances, and shifted between transport modes (e.g. sea, inland, rail and road) without being opened and with very limited cost and time effort. Overall, maritime transport of general cargo became cheap, efficient, safe and seamless: unpublished preliminary estimates by the research group at the University of Naples suggests that the freight transport cost by sea of the

goods in a standard 40' high cube container (2 TEU) fell by a more than 100 factor. Another key factor, crucial for the research presented in the thesis, is that the container triggered a structural change in the concerned maritime shipping industry: shipping companies became capable to pursue economies of scale and scope, and the structure of maritime container service networks is nowadays such that more than one thousand ports are connected amongst themselves with very reliable and stable services.

Since the early 1960s, international container shipping networks have changed significantly, mainly due to a gradual geographical expansion of the major shipping lines. Firstly, shipping carriers focused on the East-West routes connecting the main global economies, namely Asia-North America (Transpacific), Asia-Europe (FE-EU), and North America-Europe (Transatlantic) (Stopford, 2009). The transpacific is the biggest deep sea liner route, with services operating between the North American ports on the East Coast, the Gulf and the West Coast and the industrial centres of Asian countries, with some services extending to the Middle East. The Europe – Far East trade lane, includes services connecting the Northern Europe to Japan, East Asia to the Mediterranean, as well as services operating between India and European ports. Finally, the last of the three major container trade routes is the Transatlantic lane, which includes connections between North America - Northern Europe and the Mediterranean basin.

Beside this so-called Main Street, other three trade lanes gained attention from major shipping companies, especially due to progressive liberalization of maritime transport in the 1980s: the north-south offshoots connecting Europe to Africa, Asia to Australasia and North America to Central and South America. Indeed, these north-south connections were strongly linked to the major east-west routes and enabled expansion into regional markets. Thus, main shipping companies added feeder routes to serve areas outside the core Europe-North America-Far East trade lane. Moreover, the growing volume of goods traded on the Main Street led the increasing in ocean-going vessels size deployed on East-West routes and produced a cascade effect: to avoid the scrapping of smaller vessels, shipping companies shifted the replaced vessels to the North-South markets. To provide a global coverage of both the main East-West trade lanes with the secondary North-South routes, major shipping companies branched out into offering a mixture of direct and transshipment services.

A more thorough look at the evolution of the market of maritime container shipping companies – the equivalent term “ocean carriers” will be used interchangeably throughout the thesis – also reveals interesting development patterns. In general, the maritime container shipping business involves a remarkable financial effort – by way of example, a 18.000 TEU containership costs more than 120 MUS\$ (SYS 2008a.,Clarkson Research, 2020), and the top ocean carriers manage a fleet of hundreds of vessels – and, as a consequence, a strong tendency towards aggregation into an oligopolistic market (Sys 2009) has been observed since 2007 (Sys 2008, Meersman 2018, Merk, 2018, Brooks 2019). As a result, the top-10 ocean carriers control 83% of the global container market, according to Alphaliner’s ranking of all the shipping lines. (Alphaliner monthly monitor, January 2020). The tipping point is represented by the consolidation of the top container shipping companies into three main

alliances, as illustrated in section 4.1.1.4.

Establishing a strategic partnership in the form of an alliance allows ocean carriers to share and optimise integrated functions, such as transactions with insurers and legal services, to manage and use vessels and containers, to contract with inland transport and ocean feeder services and investment in port terminals, however keeping their individual corporate identities defined by administrative and marketing functions. (Hoffmann, 1998). In the light of the thesis, the concentration of the worldwide maritime container market into a very limited number of key players impacts significantly also on planning and governance of this system from a public perspective: often, huge investments are required to improve container terminals and other container-related infrastructures, and the power of key ocean carriers influences usually key planning decisions, as highlighted by Merk (2013).

#### 4.1.1 Maritime container services: key features

Before diving into the exercise of implementing a worldwide maritime container supply model, it is worth identifying the main features of the container shipping market that the model has to face/incorporate. Thus, the following section describes the topology of the liner services for container transport, the main operational variables and the composition of the global container fleet and its evolution over the last decade. Finally, the market size of the main shipping carriers as well as their co-operational agreements (i.e. ocean alliance, service sharing) are illustrated.

##### 4.1.1.1 Topology of container services

As a result of an evolutionary path spanning over four decades, since the beginning of containerisation, the structure of container services operated today by shipping companies is characterised by some relevant features,

First, in terms of *network structure*, all ocean carriers resort to a network of services pivoted on key hub ports, with a structure based on the concept of transshipment. Basically, ocean carriers cluster the world into geographical subareas (corresponding to the concept of macro-areas, see section 4.2.2.2) and identify a limited number of hub ports within each subarea, thus acting as “home ports” of that ocean carrier in that subarea. Then, a network of *deep-sea* container services is deployed between these hubs, with the rationale described below in this sub-section. In turn, other ports in each subarea are connected to concerned reference hubs via *feeder* services. The resulting *hub-and-spoke* network structure allows maximising the economies of scale on deep-sea services, which are usually operated with very large containerships, up to 24.000 TEU on the Europe-Far East trade lane, and allows achieving a complete connectivity amongst ports with a relatively low number of container services. On the other side, the need for transshipment operations in hub ports yields a non-negligible improvement of port-to-port transit times with respect to a hypothetical network with direct port-to-port links.

Although largely prevailing, hub-and-spoke is not the sole type of transshipment operated by ocean carriers, that sometimes also resort to *interlining* and *relay* transshipment, that occur

when two or more deep-sea services are connected with each other with the aim of rationalising port calls within each subarea (interlining) or of creating connections between trade lanes (relay)<sup>11</sup>. Notably, also due to the so-called *cascade effect* (see below), especially in late 2000 and early 2010 a network of direct services with medium-size ships, in between 6.000 TEU and 10.000 TEU, was also deployed by ocean carriers; however, this tendency has been systematically decreasing for years.

Second, in terms of *structure of deep-sea container services*, based on the sequence of ports of call – usually defined *port string* – deep-sea container services can be classified either as *end-to-end* services, when they connect two continents, or as *pendulum* services, when they connect three continents. *Round-the-world* services, characterised by one-way sailing through Suez and Panama canals, operated up to early 2000 and now are disappeared. As reported from various sources (e.g. Alphaliner, Drewry) some 85% of deep-sea services are end-to-end and the remaining are pendulum services.

It is worth noting that the growing trend in maximizing economy of scale deploying mega vessels (operating *hub-and-spoke* networks) made the system highly exposed to economic crisis. During the recent Covid-19 pandemic, the demand for container freight has fallen by approximately 30%, and the current oversupply of vessels has gotten worsen due to the planned container ship supply increase (+5% in 2020 and expected +3% in 2021). Consequently, to face the crisis shipping carriers started to reduce container supply by cancelling services. Indeed, the so-called *blank sailings* have increased significantly during the last year, especially on the Asia-North America West Coast trade lane on the Far East-North Europe trade lane.

#### 4.1.1.2 Operational variables of container services

The liner shipping services are usually operated on fixed schedules calling defined port string. To be competitive in the shipping industry the ocean carriers rely on the improvement of some key features affecting operational efficiency and service effectiveness (Song, 2015).

To this purpose, the companies tackle different level of choices. From a tactical perspective, the main operational variables include:

- service frequency;
- fleet deployment;
- sailing speed optimization;
- vessel schedule design.

Finding the optimal trade off among these variables is a challenging task.

The total travel time of a service route can be evaluated as

---

<sup>11</sup> Typically hubs are points of convergence of regional shipping, essentially linking separate hierarchies and interfacing global and regional freight distribution systems. Whereas relay and interlining locations connect the same hierarchy levels and improve connectivity within the network. (Rodrigue, 2010)

$$t_{rs} = \frac{d}{s} + \sum_p tt_p \quad (4.1)$$

where

$d$ , is the total travelled distance (in nautical miles)

$s$ , is the vessel speed (in knots)

$tt_p$ , is the turnaround time spent in each port (in days).

Given the total travel time, carriers define the service frequency and evaluate the number of vessels needed to operate the service as follows:

$$n_v = \text{int}(\varphi_s \cdot t_{rs}) + 1 \quad (4.2)$$

where

$n_v$  is the number of vessels needed to operate the service;

$\varphi_s$  is the service frequency.

The frequency is one of the most important KPIs which impacts on the effectiveness of a service: rather than fast deliveries, customers need reliable supply chains (to get their cargo on time).

The vessel speed is a crucial operational variable to manage the fuel consumption and the carbon emissions. The shipping carriers often adopt the so-called *Slow Steaming* which is a process based on the vessel speed reduction (from 20-24 knot to 12-19 knot) to cut down the engine power and consequently limit the fuel consumption (which usually amounts to 25% of operational costs). Thus, the slow steaming help the ship owners to face the rising of fuel prices.

However, to ensure a certain frequency of service while reducing the vessel speed, the shipping companies need to provide more vessels to compensate for a longer travel time. Since the number of vessels in a company's fleet is usually limited, they usually charter vessels from other carriers, resulting in additional costs. These costs of course depend on the type of vessels needed to accommodate the container demand.

### 4.1.1.3 Ships

Maritime shipping services provide container carrying capacity by operating a wide range of vessels in the liner market. The bulk of container are transported by large cellular vessels. As shown in Table 4, in 2019 the global cellular fleet comprised over 5.300 vessels, with a total amount of slot capacity of 23.2 million TEUs.

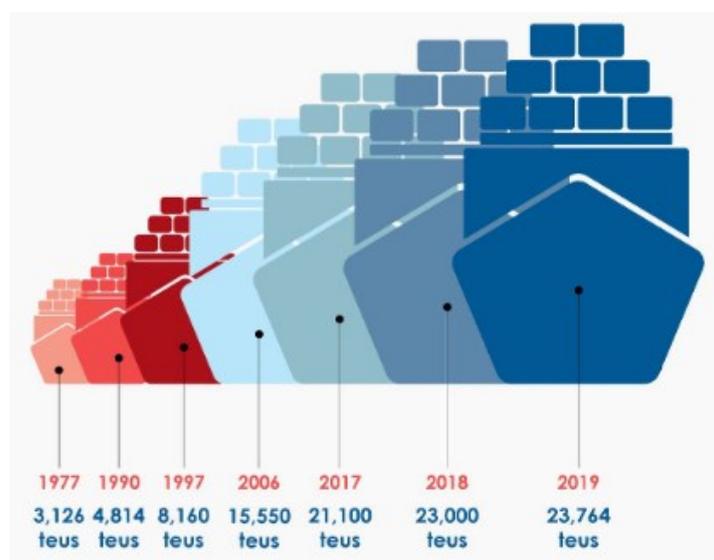
**Table 4: Cellular fleet breakdown**

Current TEU	All cellular ships Units	TEU
18,000-24,000	115	2,311,433
15,200-17,999	42	703,681
12,500-15,199 NPX	252	3,469,931
10,000-12,499	164	1,760,937
7,500-9,999	480	4,235,592
5,100-7,499	448	2,791,383
4,000-5,099	631	2,859,792
3,000-3,999	252	878,393
2,000-2,999	675	1,723,100
1,500-1,999	601	1,034,715
1,000-1,499	709	817,562
500-999	786	582,565
100-499	182	59,413
<b>TOTAL</b>	<b>5,337</b>	<b>23,228,497</b>

\*Source: Alphaliner Monthly Monitor, January 2020

In the last thirty years, according to the growing of container trade, also the size of cellular fleet increased at an exponential rate, as larger container vessels potentially allow for economies of scale through lower cost per TEU (Figure 8 ).

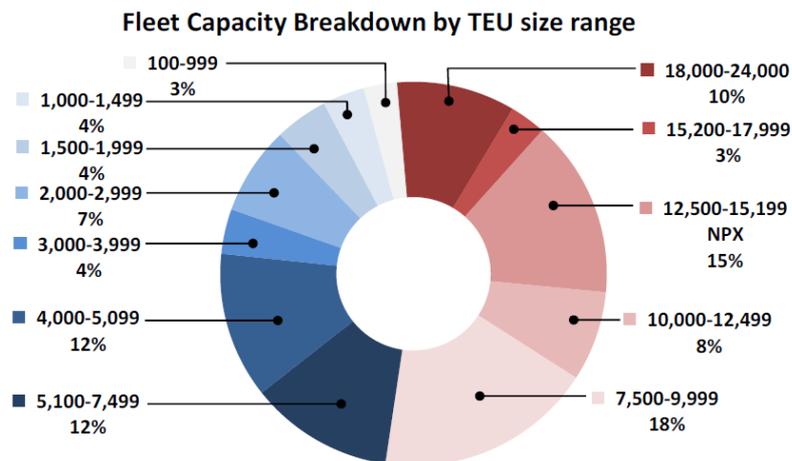
**Figure 8: evolution of cellular ship size.**



\*source: Studi e Ricerche per il Mezzogiorno (SRM), 2020

Nowadays, the global cellular fleet is very heterogeneous, as shown in Figure 9. The share accounted by mega vessels (sized between 18.000-24.000 TEUs) reached 10% during the last year. Whereas vessels ranging above 10.000 TEUs have become the workhorses on the main trades.

**Figure 9: Fleet capacity breakdown by TEU size range**



*\*source: Alphaliner Monthly Monitor, January 2020*

The fast growth in mega ships led to a higher risk in filling the overcapacity supplied (Meersman, 2015). To overcome this issue, the major shipping companies adopted different strategies. Firstly, services operating main trade lanes were restructured, in order to keep the vessels' capacity highly used, leading to the cascade effect described in the section above. Secondly, the vessel gigantism resulted in the formation of strategic alliances between shipping carriers, in order to spread the risk associated with new investment among the participants involved in the alliance, as illustrated in the following section.

#### 4.1.1.4 Carriers and alliances

Despite a multitude of carriers operating worldwide services, the container shipping market is dominated by few bigger players. Table 4 shows the rank of top 30 shipping carriers based on their fleet capacity operated in 2019 (*Alphaliner, 2020*). These numbers highlight the high concentration of the market: the share of biggest 30 companies account for 92% of the global fleet capacity. Furthermore, focusing on the top ten carriers, the market results even the more concentrated with the top ten carriers reaching 83% of the total market share.

**Table 5: ranking of top 30 shipping companies based on total operating capacity in 2020**

Rank	Operator	TEU	ships	market share
1	APM-Maersk	4192742	707	0.178
2	MSC	3766049	565	0.159
3	COSCO Group	2938030	481	0.124
4	CMA CGM Group	2695863	507	0.114
5	Hapag-Lloyd	1717889	240	0.073
6	ONE	1581368	221	0.067
7	Evergreen	1276568	200	0.054
8	Yang Ming	646630	100	0.027
9	PIL	392410	119	0.017
10	HMM	388526	63	0.016
11	Zim	292303	61	0.012
12	Wan Hai	274036	97	0.012
13	Zhonggu Logistics	161068	113	0.007
14	IRISL Group	152419	48	0.006
15	KMTC	146734	63	0.006
16	Antong (QASC)	146152	117	0.006
17	SITC	117812	82	0.005
18	X-Press Feeders	115137	79	0.005
19	UniFeeder	85067	57	0.004
20	TS Lines	79673	35	0.003
21	SM Line	73882	17	0.003
22	Arkas / EMES	64517	38	0.003
23	Sinokor	63937	48	0.003
24	Sinotrans	59339	37	0.003
25	RCL	54058	29	0.002
26	Salam Pacific	52121	52	0.002
27	Matson	45634	25	0.002
28	Emirates Shg	42672	8	0.002
29	Grimaldi (Napoli)	41756	38	0.002
30	Swire Shg	40815	25	0.002

*\*source: Alphaliner Monthly Monitor, January 2020.*

During the last decades, this oligopolistic trend led shipping companies to cooperate and make agreements in order to act like monopolist in the market. Mainly, there are two different forms of consolidation, namely ocean alliances and mergers and acquisitions (Sys, 2010).

There are three main ocean alliances operating in the market (Figure 10). The biggest one is the 2M, formed by Maersk Line and MSC on the basis of a vessel sharing agreement started in 2014. The Ocean Alliance is the second major cooperation, formed in the 2017 for an initial period of 5 year, later extended until the 2027. It currently includes the CMA CGM,

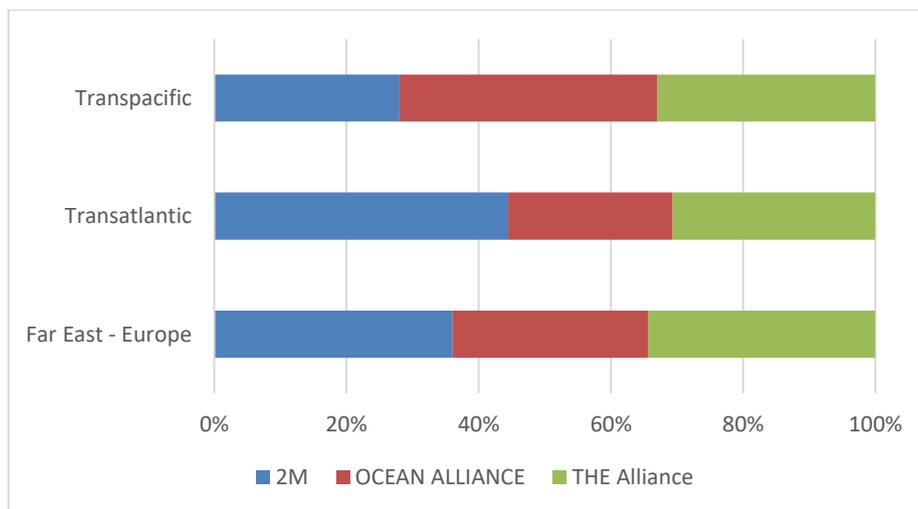
COSCO Shipping, OOCL and the Evergreen Line. Whereas, The Alliance, launched in 2017, firstly included Hapag-Lloyd, ONE (K Line, MOL, NYK), and Yang Ming. Lately, the HMM joined the group considerably boosting the global alliance market share, as will be illustrated in Chapter 5 ).

**Figure 10: Shipping alliances – source own composition**



Establishing shipping alliances helps to reach greater market shares and control more effectively the trade routes, serving a wider geographical area and offering more frequent services. Nowadays, the bigger players control the busiest container trade lanes thanks to the established strategic alliances. Figure 11 shows the market share of the aforementioned shipping alliances on the three Main Street, highlighting the dominant position held by the Ocean Alliance and the 2M on the Transpacific and the Transatlantic routes respectively.

**Figure 11: market share of shipping alliances on busiest container routes**



\*source: own composition - data 2020 Q3 – week 26.

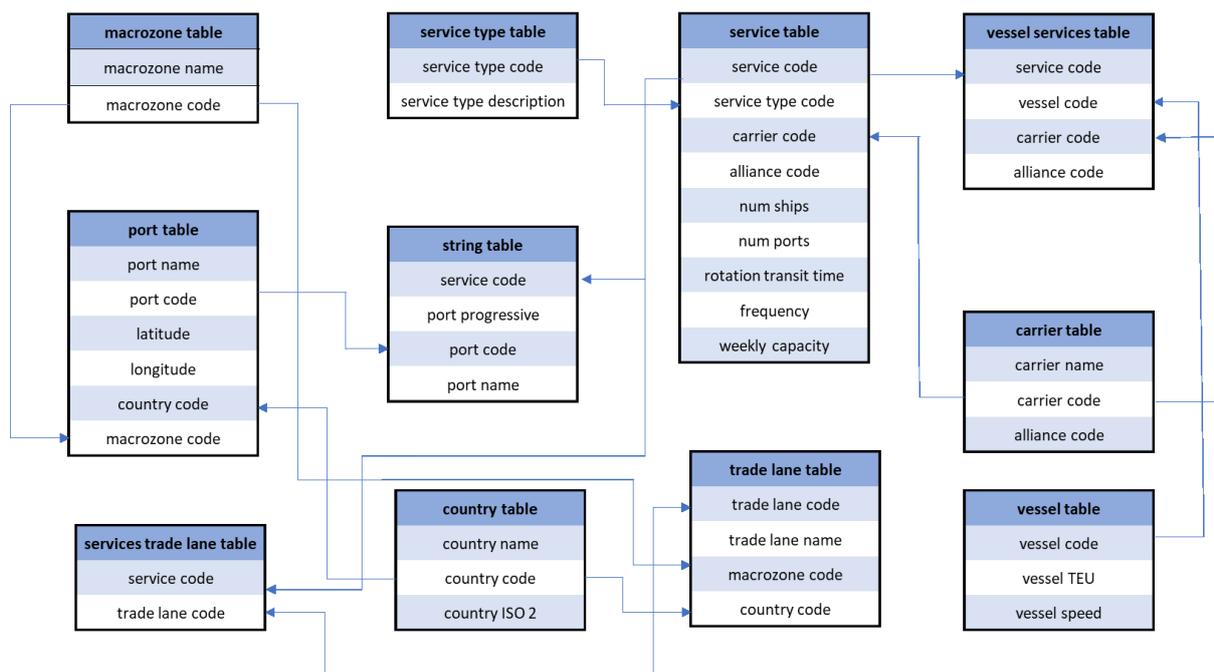
## 4.2 A database of maritime container services from 2018 to 2020

This section illustrates the structure and the practical implementation of a database of worldwide container shipping services for the years 2018 to 2020. The database represents the backbone of many applications of the thesis, starting from the implementation of the maritime container supply model described in Section 4.3. Specifically, the structure of the database is described in Section 4.2, and its implementation in Section 4.2.2. .

### 4.2.1 Structure of the database

The database leverages the traditional approach of relational databases, with data arranged in tables, each characterized by a primary key given by a field (or a combination of fields) allowing unique identification of each record of the table. Importantly, the database is also perfectly consistent with the maritime supply model described in Section 4.3. In a nutshell, the structure of the database is illustrated in Figure 12 .

Figure 12: structure of the database of the maritime container services



Overall, the database encompasses 14 tables, whose detailed structure and description of its concerned fields is reported in the next sub-sections.

#### 4.2.1.1 Port Table

Table 5 contains all information related to ports in the network of maritime container services. Each port is identified by a unique *portCode*, with a digit structure of type *xxxxyy* wherein *xxx* is the *countryCode* of the country of the port and *yy* is a sequential number identifying the

port, given the country, assigned with alphabetical order. The structure of the *portTable* is reported in the following.

**Table 6: structure of the table portable (primary key in bold)**

field name	data type	description
<b>portName</b>	character	name of the port
<b>portCode</b>	<b>integer</b>	<b>unique port code identifier</b>
<b>countryCode</b>	integer	country code of the port
<b>macroAreaCode</b>	integer	code of the geographical region including the port

#### 4.2.1.2 Country Table

All the countries included in the model are listed in this table and provided with a unique numerical code identifier (*countryCode*) as well as the ISO 3166-1 alpha-2 code (*countryCodeISO2*). The next table shows the structure of the table.

**Table 7: structure of the table country**

field name	data type	description
<b>countryName</b>	character	name of the country
<b>countryCode</b>	<b>integer</b>	<b>unique country code identifier</b>
<b>countryCodeISO2</b>	character	country ISO 3166-1 alpha-2 code

#### 4.2.1.3 Macro Area Table

This table illustrate the geographical belonging of spatial characteristic of the shipping services, with a different level of aggregation. Thus, each country belongs to a specific geographical region (*macroArea*) included in a certain trade region. Details of the different fields are listed in as follows:.

**Table 8: structure of the table macro area**

field name	data type	description
<b>macroAreaName</b>	character	name of the geographical region
<b>macroAreaCode</b>	<b>integer</b>	<b>unique geographical region code identifier</b>
<b>countryCode</b>	integer	code of the country belonging to the macroArea
<b>tradeRegionCode</b>	integer	code of the trade region including the macroArea

#### 4.2.1.4 Trade Lanes Table

Commercial routes operated by liner services are listed in the trade lane table, and for each of them, a unique progressive numerical code is defined as well as the trade regions belonging to the trade lane (*tradeRegionCode*, *tradeRegionName*) These commercial lanes are defined accordingly to the geographical coverage reported in the commercial name of the services.

**Table 9: structure of the table trade lanes**

field name	data type	description
<b>tradeLaneName</b>	character	name of the shipping route
<b>tradeLaneCode</b>	integer	<b>unique code identifying the trade lane</b>
<b>tradeRegionCode</b>	integer	<b>code of the trade region</b>
<b>tradeRegionName</b>	character	name of the trade region

#### 4.2.1.5 Carrier Table

Each liner service is operated by a shipping carrier univocally identified in the *carrier table*. This table provides a complete list of active shipping carriers, identified by the *carrier Name* and a numerical code (*carrierCode*). Also, to represent merges and acquisitions occurred in the shipping market over the last three years, each carrier is associated with a *carrierHoldingCode*, referred to the holding group to whom they belong. Finally, to represent whether or not a carrier belongs to a shipping alliance, each operator is associated with an *alliance code*.

**Table 10: structure of the table carrier**

field name	data type	description
<b>carrierName</b>	character	name of the shipping operator
<b>carrierCode</b>	integer	<b>unique code identifying the single carrier</b>
<b>carrierHoldingCode</b>	integer	code of the main carrier holding group
<b>allianceCode</b>	integer	code of the related ocean alliance

#### 4.2.1.6 Alliance Table

Information related to the participants involved in the ocean alliances are listed in this table. Importantly, to highlight changes in cooperative agreements, companies participating in the alliances are classified on a yearly basis (from 2018 to 2020). The table structured is shown in the following:

*Table 11: structure of the table alliance*

field name	data type	description
<b>allianceName</b>	character	name of the ocean alliance
<b>allianceCode</b>	integer	<b>unique code identifying the alliance</b>

field name	data type	description
<b>participants 2018</b>	character	carriers involved in the shipping alliance in the 2018
<b>participants 2019</b>	character	carriers involved in the shipping alliance in the 2019
<b>participants 2020</b>	character	carriers involved in the shipping alliance in the 2020

#### 4.2.1.7 Service Type Table

Table 12 includes information about the types of services collected in the database. Each service type is uniquely identified through the *serviceTypeCode* a progressive integer numerical code and a shortened name reported in the *serviceTypeRaw* field. A full description of each type is provided in the *serviceTypeDescr*, helping the identification of liner container services. Finally, each service type is associated with a Boolean variable (in the *ifContainerService* field) to make easier the selection of container services in the application coding.

**Table 12: structure of the table service type (primary key in bold)**

field name	data type	description
<b>serviceTypeCode</b>	integer	<b>unique code identifying the service type</b>
<b>serviceTypeRaw</b>	character	shortened name
<b>serviceTypeDescr</b>	character	full description of the service type
<b>ifContainerService</b>	boolean	1= container service – 0 = others

#### 4.2.1.8 Service Table

This table lists detailed information about global liner shipping services. Each liner service, uniquely identified by a numerical code (*serviceCode*), is associated firstly with general information such as the service (commercial) name, the type (e.g. breakbulk, ro-ro, container), the trade regions covered and the participants (operators and/or alliances) involved in the service. Secondly, operational features are listed, such as the sailing frequency/ headway of the service (*sailingFrequencyOrHeadwayRaw*, *sailingHeadwayDays*, *sailingWeeklyFreq*) and the weekly capacity, both gathered by web sources. In a further step, these characteristics are revised (see section 4.2.2) and the adjusted values reported in service table as well (*reconstructionRotationDays*, *finalFrequency*, *finalNumOfShips*, *weeklyCapacity*, *weeklyCapacityByPort*), as can be observed in Table 13 .

Importantly, to set a time dimension and select active services for each investigated year, each service is associated with the information about its lifetime, from the entry into service (start year) to the final dismissal or substantial operational changes (end year).

Table 13: structure of the table service

field name	data type	description
<b>serviceCode</b>	integer	<b>unique code identifying the service</b>
<b>serviceCodeRaw</b>	character	alphanumerical code identifying the vessel
<b>serviceName</b>	character	name of the service (given by the operator)
<b>serviceType</b>	character	shorted name of service type
<b>allianceName</b>	character	name of the alliance operating the service (if any)
<b>participants</b>	character	carriers involved in the service
<b>coverage</b>	integer	trade regions covered
<b>sailinghFrequencyOrHeadwayRaw</b>	integer/ character	frequency of the service or headway reported by the operator
<b>sailingHeadwayDays</b>	real	adjusted service headway in days
<b>sailingWeeklyFreq</b>	real	adjusted weekly frequency of the service
<b>rotationDurationDaysRaw</b>	integer	voyage duration in days
<b>weeklyCapacityRaw</b>	character	size and number of vessels per week (source: shipping companies)
<b>numShipsFromDbVessel</b>	integer	number of ships perating the service (extracted from the <i>vesselServices</i> table)
<b>numPorts</b>	integer	number of ports called by the service
<b>distNm</b>	real	total distance of the service loop
<b>teuMin</b>	integer	carrying capacity of the smallest ship operating the service
<b>teuMax</b>	integer	carrying capacity of the biggest ship operating the service
<b>teuAv</b>	integer	average carrying capacity
<b>reconstructionRotationDays</b>	real	estimated voyage duration in days
<b>finalFrequency</b>	real	frequency associated to the service
<b>finalNumOfShips</b>	integer	number of vessels associated to the service
<b>weeklyCapacity</b>	real	weekly capacity supplied by the service
<b>weeklyCapacityByPort</b>	real	Weekly capacity deployed by the service on each port call
<b>startYear</b>	integer	first year of service
<b>endYear</b>	integer	last year of service (changing included)

#### 4.2.1.9 Services Trade Lanes Table

Table 14 then gives an overview of each shipping service is associated with the operated trade lane (via numerical code defined in table 12). In order to represent changes occurred in the supplied services over the investigated period, three different tables are defined.

**Table 14: structure of the table services trade lanes**

field name	data type	description
<b>serviceCode</b>	integer	<b>unique code identifying the service type</b>
<b>serviceCodeRaw</b>	character	temporary service code
<b>tradeLane</b>	character	name of the trade lane
<b>tradeLaneCode</b>	integer	<b>unique code identifying the trade lane</b>

#### 4.2.1.10 Services Carrier Table

This table defines connections between active services and carriers operating them. Each service, uniquely identified by its *serviceCode*, is associated with the related *carrierCode*, the *carrierHoldingCode* and the *allianceCode* (if any). If a service is jointly operated by two or more carriers, the same *serviceCode* is repeated and each row is associated with the single carrier Code. According to the yearly gathered data, three different service carrier tables are created.

**Table 15: structure of the table services carrier**

field name	data type	description
<b>serviceCode</b>	integer	<b>code of the shipping service</b>
<b>carrierCode</b>	integer	code of the single carrier
<b>carrierHoldingCode</b>	integer	code of the main carrier holding group
<b>allianceCode</b>	integer	code of the ocean alliance

#### 4.2.1.11 Service String Table

The service string table is a key table in the shipping services database. It listed detailed information on the service loop operated by each service. To enable the use of coded application, the table is structured as illustrated below. For each service uniquely identified by the *serviCode*, the full loop is reported (*portProg*, *portName*, *portCode*). Also, the estimated travel distance (in nautical miles) between adjacent ports (*p2pDist*) in the loop and the progressive travel distance (*totDist*) of the service loop is provided. The procedure to evaluate port-port distances is illustrated in Section 4.3.2.

**Table 16: structure of the table service string**

field name	data type	description
<b>serviceCode</b>	Integer	<b>unique code identifying the vessel type</b>
<b>portProg</b>	Integer	position of the port call in the loop
<b>portName</b>	Character	name of the port
<b>portCode</b>	Integer	<b>code identifying the port</b>
<b>p2pDist</b>	Real	estimated distance between adjacent ports
<b>totDist</b>	Real	estimated progressive distance between successive ports in the string

#### 4.2.1.12 Vessel Type Table

Data gathered on worldwide liner services refer to several types of vessels. In this table each type is described and identified with a numerical code, as reported below. Importantly, the vessel type table reports information about all the available types of vessels operated liner container service (e.g. the Con Bulkers vessels and the cellular ships).

**Table 17: structure of the table vessel type**

field name	data type	description
<b>vesselTypeCode</b>	<b>integre</b>	<b>unique code identifying the vessel type</b>
vesselType	character	vessel type name
vesselTypeDescr	character	vessel type full description

#### 4.2.1.13 Vessel Table

In this table, all the available information about features of yearly active vessels is reported. For each vessel the reported features are:

- the *vessel Name* and its numerical progressive code (the *vessel Code*), representing the primary key of the table;
- the vessel type, identified by a shorted name in the *vesselTypeRaw* field and a numerical code reported in the *vesselTypeCode* (both linked to the *vessel type table*);
- the flag (*vesselFlagRaw*) and the flag code (*vesseFlagCode*) associated to the flag;
- the vessel's carrying capacity in TEU (*vesselTEU*) and the deadweight tonnage in metric tons (*vesselDWT*);
- the commercial speed of the vessel (*vesselSpeed*);
- the vessel Gear, which indicates whether or not a ship is equipped with cranes for loading and off-loading a port, through a Boolean variable.

The database includes three different *vessel tables* based on yearly data.

**Table 18: structure of the table vessel**

field name	data type	description
vesselNameRaw	character	name of the vessel
<b>vesselCode</b>	<b>integer</b>	<b>unique code identifying the vessel</b>
vesselTypeRaw	character	shorted name of vessel type
vesselTypeCode	integer	code of the vessel type
vesselFlagRaw	character	shorted name of the Country associated with the flag
vesselFlagCode	integer	progressive flag code
vesselTEU	integer	carrying capacity of the vessel in TEU

field name	data type	description
vesselDWT	integer	vessel deadweight tonnage
vesselSpeed	real	commercial speed (knots)
vesselGear	boolean	1= cranes – 0 = no cranes

#### 4.2.1.14 Vessel Services Table

The *Vessel Service Table* associates each active vessel with information about the carrier (i.e. *vesselCarrierName*, *vesselCarrierCode*, *vesselCarrierHoldingCode*) and the service/services operated (*serviceCode*). Also, information about vessels deployed on co-operated services (e.g. services operated by shipping alliances) is specified in the *allianceName* and *allianceCode* fields. Also, as mentioned before, three different tables are defined in order to represent changes in vessels' deployment over the investigated period.

**Table 19: structure of the table vessel services**

field name	data type	description
vesselNameRaw	character	name of the vessel
<b>vesselCode</b>	<b>integer</b>	<b>unique code identifying the vessel</b>
vesselCarrierName	character	name of the vessel operator
vesselCarrierCode	integer	code of the vessel operator
vesselCarrierHoldingCode	integer	code of the main carrier holding group
allianceName	integer	name of the shipping alliance operating the service
allianceCode	integer	code of the shipping alliance
serviceCode	integer	service operated by the vessel

## 4.2.2 Implementation of the database

The practical implementation of the database described in Section 4.2.1 has been cumbersome and time consuming. In general, many data sources provide information on all concerned aspects of maritime container services, including e.g. ships, string of port calls, level-of-service attributes, and so on. From a general perspective, four types of data sources can be identified:

- official information on services by ocean carriers e.g. Maersk, MSC, COSCO Shipping and so on;
- vessel data from automated identification services (AIS) given by free vessel tracking web sites (i.e. [www.vesselfinder.com](http://www.vesselfinder.com), [www.marinetraffic.com](http://www.marinetraffic.com)) displaying real time ship positions and marine traffic detected by global AIS network;
- details about worldwide ports are gathered from the World Port Index which contains information about location, characteristics, known facilities, and available services;
- port call data are collected from shipping companies' website and real time ships

mapping tools (i.e. [www.marinetraffic.com](http://www.marinetraffic.com));

- data from relevant consultancies (i.e. Alphaliner, Drewry, Clarkson, Lloyd’s list).

A comprehensive scouting of all freely accessible data sources has been performed<sup>12</sup> for each year between 2018 and 2020, leading to a first raw version of the database. Data coverage was very satisfactory for deep-sea services, practically entirely covered, and satisfactory for feeder services, whose volatility clearly is more challenging. The following sub-sections report basic information on each table, preparatory for the analyses reported in subsequent sections and chapters of the thesis, and also describes concerned implementation issues and solutions.

#### 4.2.2.1 Ports

Overall, the database contains information on 1994 ports, including the Panama and the Suez canals which are represented to model actual routes of container services’ loops. The number of ports served during each year of the investigated period, is outlined in the following table:

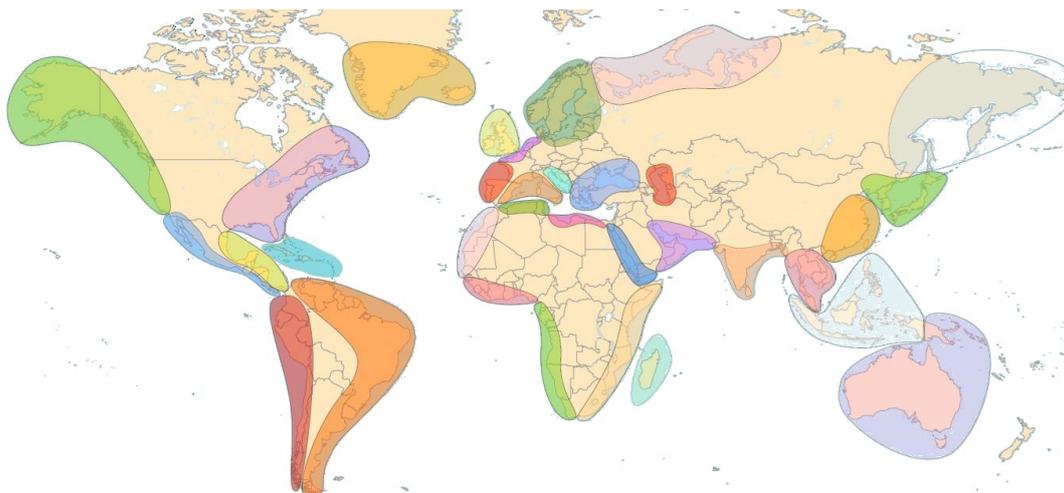
**Table 20: number of ports yearly called by worldwide services**

	2018	2019	2020
# called ports	1115	1050	1045

#### 4.2.2.2 Macro- areas

Consistent with the general remarks reported in Section 4.1, 33 macro-areas have been identified to cluster worldwide ports belonging to 188 countries, as depicted in Figure 13.

**Figure 13: overview of regional macro-areas**



<sup>12</sup> This part of the research benefited from the collaboration of several MSc students in Management Engineering and in Hydraulics and Transport Engineering at the University of Naples Federico II, namely Filomena Di Lorenzo and Ivana Manco for the year 2018, Maria Luongo for the year 2019 and Chiara Esposito for the year 2020.

Notably, consistent with the purposes of the research, a more disaggregated clustering has been adopted to represent Euro-Mediterranean ports, for a more thorough analysis of their role in global container networks, as well as for some regions of the East-Asia exhibiting significant domestic container liner services (e.g. the southern region of the Pearl River Delta served by the Port of Shenzhen, a container hub including several relevant port terminals). In particular, Italian ports belong to two different macro-areas, namely the West-Med-Europe and the East-Med-Adriatic. Table 21 shows the extent of each macro-area in terms of numbers and ports included.

**Table 21: macro-areas details**

macro-area	# countries	# ports
Greenland & Iceland	2	22
North America West	1	27
North America East	3	44
Central America East	6	24
Central America West	2	15
South America West	3	28
South America East	14	68
West Med Africa	2	19
East Med Africa	1	11
North West Africa	14	34
South West Africa	9	24
South East Africa	3	17
Red Sea	7	14
West Med Eu	3	34
Scandinavia & Baltic Sea	8	136
South West Europe	2	26
East Med Adriatic	5	15
East Med Egeo Black Sea	6	39
Northern Range	3	31
Caraibi	24	47
East Med	4	8
India Sri Lanka Bangladesh	4	26
Indocina	10	111
Far East	3	67
Korea & Japan	3	84
Russia East	1	14
Russia West	19	9
Oceania	8	73
Persian Gulf	3	41
UK	6	41
Indian Ocean	5	15
South East Asia	188	24

### 4.2.2.3 Container Services

The *serviceTable* represents one of the key tables of the database, with many important information on the global network of maritime container services.

It is important to underline that the implementation of the database has been performed trying to include all container-related maritime services, that is not only those operated with cellular containerhips but also with multipurpose ships enabling transport of containers (e.g. the Con-Ro vessels, Con-Bulkers vessels,...). At a glance, the breakdown of container services per year is reported in the following.

**Table 22: Number of active services per year**

# active services	year
1900	2018
1867	2019
1756	2020

Importantly, some information required for the implementation of the *serviceTable* had to be reconstructed via the supply model (see section 4.3). In particular, in many cases, relevant operational features of liner services (i.e. sailing time, service frequency, weekly capacity) derived from web sources were found to be inconsistent or missing. Thus, to overcome this issue and set up the fullest and most reliable database, missing data are evaluated by leveraging the supply model implementation. By way of example, to calculate the total distance covered by each service, port-to-port shortest paths are evaluated based on the model of sea routes described in section 4.3.2.1. Similarly, an adjusted transit time is calculated for each service as given by the equation (4.1), assuming a turnaround time in port equal to 18 hrs. Furthermore, given the service transit time  $tt_s$ , it is possible to derive the number of vessels ( $nv_s$ ) or the service frequency ( $\varphi_s$ ), depending on the information needed, as formulated in the (4.2).

### 4.3 The supply model of container shipping services

Following Cascetta (2009), a transport system can be defined as the “*combination of elements and their interactions, which produce the demand for travel within a given area and the supply of transportation services to satisfy this demand*”. Overall, every transport system includes two main elements, namely the transport demand and the transport supply and, thus, modelling a transport system means developing quantitative tools and methods for each of these components and for their interactions.

In general, the proper implementation of a transport supply model should follow a consolidated path encompassing the following steps:

- definition of the study area;
- zoning and selection of relevant supply elements (infrastructures, services, ...);
- implementation of the topological model;
- implementation of the analytical model.

As mentioned in the literature review in the Chapter 2 of the thesis, there are no comprehensive and holistic worldwide container supply models. The most advanced worldwide available model is by Tavasszy et al. 2011, who however implemented a rather simplistic representation of container routes, mainly limited to key trade lanes across continents, whilst a detailed approach is described by Bell et al. 2011 and 2013 however without any implementations to real-size contexts. Many papers circumvent the problem by resorting to AIS data, see e.g. the work by Ducruet (2012), however at the price of not developing any supply models with forecasting capabilities and able to inform network-related analyses. Overall, an original contribution of this thesis is to fill this gap, through the implementation of the supply model described in the next sub-sections, dealing with the study area and its zoning (Section 4.3.1), the implementation of the topological model (Section 4.3.2) and of the analytical model (Section 4.3.2.2).

#### 4.3.1 Definition of the study area and its zoning

In principle, following Cascetta (2009) “*the study area identifies the geographical area including the transportation system under analysis and most of the project effects*”. Since container services span over the entire world, and planning/governance even at port level cannot get rid of general optimization principles pursued by ocean carriers on their entire worldwide network, it is clear that the study area should include the entire world.

The zoning of the study area, that is its partition into traffic analysis zones (TAZ), is usually finalised at simplifying the spatial representation of the mobility, through the identification of centroids proxying the actual origins and/or destinations of trips from/to each TAZ. In fact, the development of a supply freight model for the worldwide container network is usually

finalised at calculating port-to-port skim matrices, either for port-related analyses or in the context of the development of a multimodal freight supply model (see Chapter 3), that will have its own centroids representing landside origins/destinations. Overall, there is no formal zoning, and ports are directly treated as centroids. Many applications and results presented in subsequent chapters of the thesis make use of the aggregation of ports into macro-areas, described in the previous paragraph.

### 4.3.2 Topological model

This task represented the most challenging step for the implementation of the worldwide container model. Preliminary, it is important to recall that the proper implementation of a topological model implies the choice of the approach to adopt for the representation of the transport supply system under analysis. In this respect, a first decision refers to the nature of the concerned transport system: the supply of container services falls intuitively into the category of discontinuous transport systems in space (access/egress is possible only in ports) and in time (vessel calls follow a liner schedule and thus are not always available). That said, a further classification pertains to the magnitude of such discontinuities; in passenger transport, a distinction is introduced between services with high frequency (typically, transit systems in urban and metropolitan areas) and low frequencies (typically, intercity rail and air services). The former are modelled with a so-called *line-based* approach, the latter with a *schedule-based* approach, resulting respectively in a synchronic hypergraph (Nguyen and Pallottino, 1988; Spiess and Florian, 1989; Cascetta, 2009) and in a diachronic graph (Nuzzolo et al, 1996).

Apparently, the average frequency (or, equivalently, headway) of container services would suggest adopting a schedule-based approach. However, it should be underlined that the concepts of “high” and “low” frequencies are conditional upon the nature of the decision-maker and of the context. In this respect, it seems clear that an average frequency of a call per service per week – a typical situation in liner container shipping – is surely low from the viewpoint of a passenger, whilst is high from the perspective of a shipper, that usually faces a throughput time of the magnitude of weeks for most of supply chains. As a result, the network of worldwide container services can be modelled effectively via a synchronic hypergraph. This aspect was also discussed by Bell et al. 2011, already recalled in Chapter 2, who however did not apply this approach to a real-size network. Furthermore, the adoption of a synchronic hypergraph requires knowledge of the schedule of the liner services, that is the database developed in Section 4.2, which is usually unavailable; this opens room for approximate approaches, recalled later in Section 4.3.2, that exhibit shortcomings when applied for policy and planning purposes.

Given these premises, the topological model of worldwide container freight flows should be developed in two steps:

- the former is the implementation of a graph representing all routes between pairs of ports. As explained in Section 4.3.2.1, this activity is particularly challenging because

of the absence of raw topological graph usually available when dealing with landside transport modes, e.g. the *OpenStreetMap* datasets for road and rail infrastructures. The output of this activity enables calculating port-to-port skim matrices, e.g. port distances in nautical miles used also for the implementation of the database illustrated in Section 4.2 and also serves as basis for the implementation of the synchronic hypergraph (see next bullet point);

- the latter is the implementation of the synchronic hypergraph representing all container services sailing between ports in the study area. This activity is also challenging because of the size of the database of the services to model ( Section 4.2), and is normally prevented by the absence of such a database.

#### 4.3.2.1 *Topological model of port-to-port maritime routes*

By definition, the topological model is usually represented by a graph, defined as an order pair of sets,  $G=\{N,L\}$ , given by nodes ( $N$ ) and links ( $L$ ) between pairs of nodes. These links usually represent phases related to the trip and/or activities modelling a physical movement (e.g. shipping on a leg) or not (e.g. waiting for transshipping). Similarly, nodes reproduce significant events between links. In addition, nodes can be represented also by time coordinate to model events occurring at different moments of time. This is the case of diachronic network representation. Conversely, in synchronic networks nodes do not have a time specification and the same node is used to represent events occurring with different timing (Cascetta, 2009).

In many inland transport applications, the development of the topological model starts by modifying raw graph representing the relevant infrastructures which are often available on free online tools. Unlike to these, for maritime applications, there are currently no sources offering such tools. Thus, first information about port locations is collected to reproduce a georeferenced map.

Data on worldwide ports are gathered from free online sources (e.g. World Port Index database, [www.worldportsource.com](http://www.worldportsource.com), [www.marinetraffic.com](http://www.marinetraffic.com)). Using the software TransCad, 1194 ports are represented on a georeferenced map of worldwide landmasses.

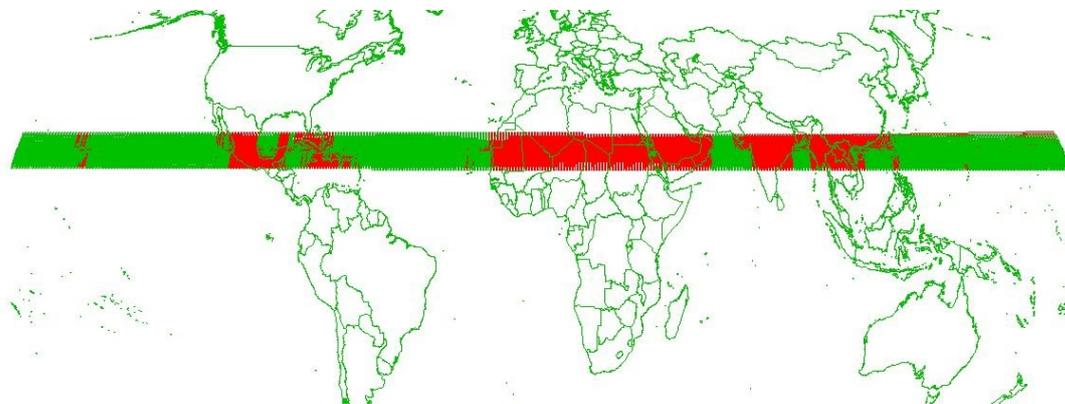
To create connections between them, it is necessary to generate a network which likely represents the actual route followed by liner services. This means that an essential condition to comply with is that sea routes do not pass over landmasses. This goal is achieved implementing an ad hoc procedure which combines VBA coding and TransCad tools, overcoming computational issues without loss of details. This methodology consists of the steps which are briefly described in the following:

- firstly, due to the lack of dedicated GIS tools to design sea routes, a sufficiently dense grid of points has been drawn on the entire water surface of the Earth, i.e., excluding lands, with an average distance of 0.5 nautical miles between adjacent points;

- each point of the grid has been connected, via direct two-way links, to all surrounding points within a radius of 30 nautical miles; links touching/crossing land have been clearly discarded. Further links have been then added to account for cases (e.g., sailing across straits) wherein the above procedure did not yield satisfactory results;
- each link has been associated with a length calculated as great circle on the Earth.

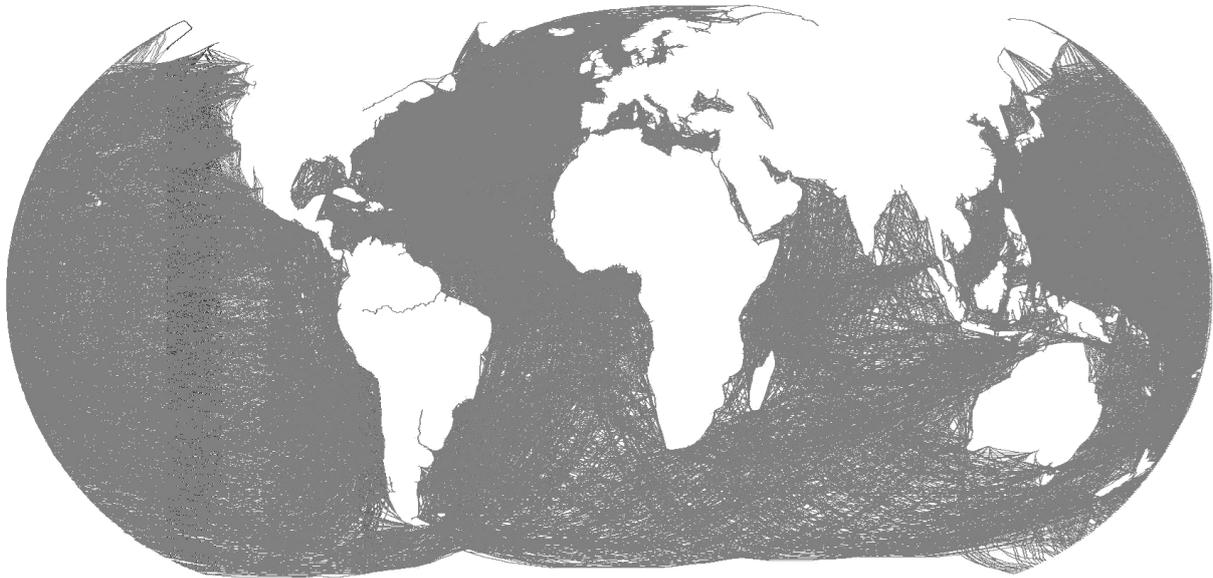
However, due to graph size, the software TransCad failed to generate connections between nodes and to create the network. Thus, the procedure has been implemented by combining a MatLab R2020b code and the GIS features of TransCAD 7.0 by Caliper Inc. In detail, the worldwide graph has been split into 146 main “bands” (Figure 14) and, for each of them, the aforementioned three-step process has been implemented.

**Figure 14: application of splitting method**



At the end of the process, all the bands have been merged (joined) in a single graph including around 7 million of links (Figure 15).

**Figure 15:maritime graph**



The shortest distance paths between all pairs of ports have been calculated on this graph, to select only links effectively used to sail between ports in the dataset. Specifically, recalling the target of the analysis, shortest paths have been calculated twice, once including and once excluding the Panama channel, to account for the presence of new Post-Panamax vessels.

The resulting final topological model, that includes 406.059 links and 127.811 nodes, among which 1,194 are ports included in the network analysis.

This model has been validated by double-checking a sample of sailing distances between ports calculated with the topological model against available shortest distances from various sources (e.g., [www.classic.searoutes.com](http://www.classic.searoutes.com), [www.sea-distances.org](http://www.sea-distances.org)), with very satisfactory results.

Also, this topological model has been used primarily to fill missing information in the dataset illustrated before, mainly transit times between ports. It also served as basis to calculate port-to-port distances and sailing times, to be used as elemental bricks for the implementation of concerned graphs, and for calculating centrality metrics, as shown in the following paragraph.

#### *4.3.2.2 Topological model of container shipping services*

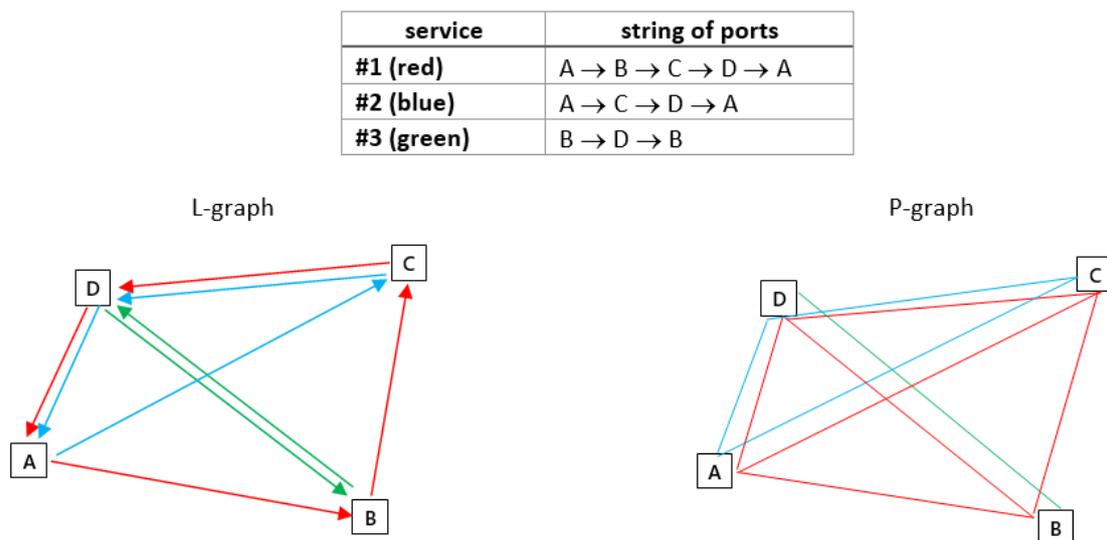
In the literature, maritime transport service networks are usually represented via synchronic graphs (Cascetta, 2009). Mostly, two options are usually available to model links between ports which are namely the *Graph of Direct Links* (GDL) and the *Graph of All Links* (GAL). These two approaches enable different applications based on the network properties modelled. Thus, the GAL approach can be useful to analyse port connectivity, due to the representation of all available connections, both direct and undirect. Conversely, to assess vulnerabilities and

to identify critical links in the system, the GDL approach is preferable.

The GDL and the GAL are based on the classical approach proposed by Hu and Zhu (2009) which defines two different types of topological networks:

- *L-graph*: in this *line graph* links between port nodes are created on the basis of the successive port of calls defined by services' loops. Consequently, only direct connections among ports are represented;
- *P-graph*: in the star graph edges represent both direct and indirect connections between ports. The rationale is that ports belonging to the same loop are indirectly connected. Thus, they are linked even if they are not adjacent calls. The result is a graph including all available connections defined by services' loops.

**Figure 16: L-graph and P-graph to model maritime container service networks (links in the P-graph are two way)**



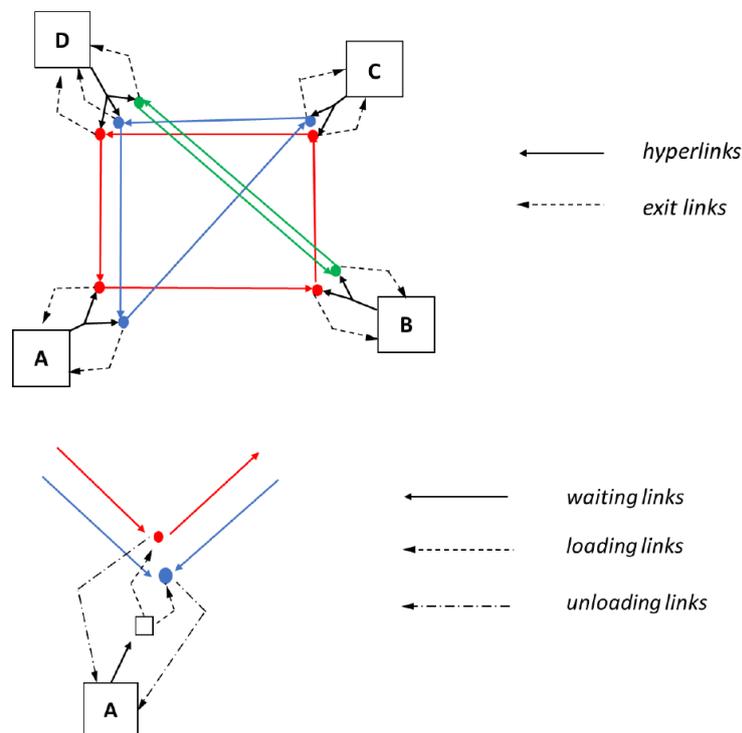
However, both the L-graph and the P-graph exhibit an important limitation, as described by Bell et al. (2011) and Bell et al. (2013), who proposed the extension to container service networks of the approach based on the concept of hyperpath introduced by Spiess and Florian (1989). The graph underlying the hyperpath approach is a hypergraph, wherein each port node is connected, via a hyperlink – that is a link with a single tail and possibly multiple heads – to all services calling at that port. This allows modelling appropriate routing strategies, including multiple paths altogether with possibly multiple transshipment options between the ports of origin and of destination.

Let  $P$  be a set of ports and  $S$  a set of maritime container services; each service  $s \in S$ , characterised by a weekly capacity  $caps$  and a weekly frequency  $\phi_s$ , is associated with an

ordered string  $P_s$  of ports of call. In turn, each port  $p \in P$  is associated with a set  $S_p$  of services calling at  $p$ . The hypergraph includes an exclusive node  $ps \in P_s$  for each called port, and each port  $p \in P$  is the tail of a hyperlink whose heads are all nodes  $ps$  with  $s \in S_p$ . The hypergraph proposed by Spiess and Florian (1989) can be turned into a hyperlink-free graph by introducing appropriate waiting links, as reported by Cascetta (2009). The impedance associated with waiting links is related to the cumulated frequency of all “attractive services” calling at port  $p$ , whose set  $S_{ph}$  is a function of the overall considered hyperpath.

Also, the representation of services in a hypergraph may follow the same rationale illustrated for the L-graph and the P-graph: that is, each service can be modelled either via an ordered sequence of links representing legs (i.e., a sailing between two subsequent ports in the service string), yielding a L-hypergraph, or by a star of direct port-to-port links (irrespective of the possible presence of intermediate ports of call), yielding a P-hypergraph. The two type of hypergraphs are termed respectively HL-graph and HP-graph in the following. An illustrative example of an HL-graph related to the example of Figure 16 is reported in Figure 17; extension to the case of an HP-graph is straightforward.

**Figure 17: Illustration of a HL-graph related to in the example of Figure 1: representation by means of hyperlinks (top) and waiting links (detail for node A, bottom)**



Whatever type of hypergraph (i.e., HL-graph or HP-graph), let  $K_h$  be the set of elemental paths belonging to an hyperpath  $h$  between two ports  $p_1$  and  $p_2$ . Each path  $k \in K_h$  represents a sequence of (one or more) services, and is associated with a choice probability  $\pi[k|h]$  given by the product of the probabilities to choose  $k$  at each of the diversion nodes belonging to  $k$ . A diversion node is any ports  $p$  in the set  $P_k$  of ports belonging to  $k$ , and the choice probability  $\pi[k|p]$  of  $k$  given  $p$  depends on the frequency of all attractive services within  $h$ , that is:

$$\pi[k|h] = \prod_{p \in P_k} \pi[k|p] = \prod_{p \in P_k} \frac{\varphi_{s(k)}}{\sum_{s' \in S_{ph}} \varphi_{s'}} \quad (4.3)$$

being  $s(k)$  the service on path  $k$  from port  $p$ . Equation (4.3) allows also defining the impedance  $i_h$  of the whole hyperpath  $h$ , as the average of the impedances  $i_k$  of the paths within  $K_h$  weighted with the corresponding choice probabilities (4.3), that is:

$$i_h = \sum_{k \in K_h} i_k \cdot \pi[k|h] \quad (4.4)$$

Nguyen et al. (1998) proposed an arborescent shortest path algorithm to find the shortest hyperpath tree in a (not necessarily) transit network, with a backward network exploration from a destination node, setting the basis for several variants and applications. As a result, calculation of shortest hyperpaths and hyperpath-based assignment are very common procedures.

The relevant question is how to specify/particularise the centrality metrics: this aspect is discussed in the next section.

## 4.4 Calculation of centrality metrics in hypergraphs for maritime container service networks

The extension of common centrality metrics to the hypergraphs is not straightforward and deserves attention. This paragraph illustrates how to specify some of the most relevant centrality metrics (e.g. betweenness, closeness and degree centrality) to the proposed container service network representation.

### 4.4.1 Degree centrality

Calculation of degree centrality on HL-graphs and HP-graphs resembles the same calculation on L-graphs and P-graphs respectively because, backward stars and forward stars in L-graphs versus HL-graphs and in P-graphs versus HP-graphs include the same nodes. This means that hypergraphs are not expected to yield appreciable improvements over L-graphs and P-graphs in the calculation of degree centrality metrics.

### 4.4.2 Closeness centrality

Calculation of closeness centrality in hypergraphs leverages a straightforward extension of (2.3), that is replacing the impedance of the shortest path calculated on the HL-graph and on the HP-graph respectively, with the impedance of the shortest hyperpath calculated on the H-graph, that is using (4.4). Again, this calculation is theoretically sounder with respect to P-graph because impedance (4.4) accounts for the whole underlying routing strategy modelled by the hyperpath.

### 4.4.3 Betweenness centrality

Extension of betweenness centrality to hypergraphs has been faced by few authors, see, e.g., Xiao (2013) and Amato et al. (2017), who proposed the following straightforward extension of equation (2.4):

$$BC(n) = \frac{1}{(|N|-1)(|N|-2)} \sum_{i,j \in N - \{n\}} \frac{nh_{ij}(n)}{nh_{ij}} \quad (4.5)$$

wherein  $nh_{ij}$  be the number of (possible multiple) shortest hyperpaths between  $i$  and  $j$  and  $nh_{ij}(n)$  the number of shortest hyperpaths between  $i$  and  $j$  passing through  $n$ .

This paper deviates from this definition, proposing a different definition of the betweenness centrality, inspired by recognising that (2.4) and (4.5) are based on 0/1 values, that is the contribution of each pair of ports  $\{i \neq n, j \neq n\}$  to the betweenness centrality of the port  $n$  equals 1 if the shortest path/hyperpath between  $i$  and  $j$  goes through  $n$  and 0 otherwise. That said, one

might leverage equation (4.3), that is calculating the sum of the occurrence probabilities of the elemental paths in the shortest hyperpath passing through  $n$ , that is:

$$BC(n) = \frac{1}{(|N|-1)(|N|-2)} \sum_{i,j \in N - \{n\}} \frac{\sum_{k \in K_{n,h_{ij}}} \pi[k|h_{ij}]}{nh_{ij}} \quad (4.6)$$

wherein  $K_{n,h_{ij}}$  is the set of paths passing through  $n$  within hyperpath  $h_{ij}$  connecting  $i$  and  $j$ . In other words, the proposed betweenness centrality is a “fuzzy” metric, in the sense that the contribution of each pair of ports  $\{i \neq n, j \neq n\}$  to the betweenness centrality of the port  $n$  in equation (2.4) accounts for the probability that the shortest hyperpath between  $i$  and  $j$  goes through  $h$ , that is in between 0 and 1. Equation (4.6) has never been applied to container service networks. Interestingly, it can be applied to both HL-graphs and HP-graphs, leading to the same thoughts on the interpretation of the betweenness centrality in L-graphs and P-graphs.

From a practical perspective, calculation of (4.6) might appear cumbersome at a first sight. However, it is useful to recall that a transit all-or-nothing (AoN) assignment<sup>13</sup> of a unit square o-d flow matrix with dimension  $|P-1|^2$  to the hypergraph yields all concerned flows for each node  $n$ , that is import/export flows (i.e., unloaded/loaded containers directed to/coming from landside), transshipment flows (i.e., containers unloaded from a maritime service and loaded to another maritime service), and through flows (i.e., containers remaining onboard the ship during the call at the port). Since the assigned o-d flow matrix has unit entries, such flows represent the contribution of all o-d pairs with respect to each node  $n$ , that is the outer summation in equation (4.6). Intuitively, considering the sum of transshipment flows and through flows for each node  $n$  yields the betweenness centrality on an HL-graph, whilst considering only through flows for each node  $n$  yields the betweenness centrality on an HP-graph.

---

<sup>13</sup> See for details, amongst others, Cascetta (2009) and Ortuzar and Willumsen (2011). AoN transit assignment is a very common feature offered by most of the commercial transport software.

#### 4.4.4 Results

This section illustrates the results of the application of the aforementioned centrality metrics to the worldwide maritime container service network.

##### 4.4.4.1 Degree centrality

The calculation of degree centrality is not affected by the type of underlying graph; for the sake of completeness, the following Table 23 reports the degree centrality for the top 20 ports.

**Table 23: top 30 ports Degree Centrality results in P-graph**

<i>Port</i>	<i>P-graph</i>	<i>rank</i>
Shanghai	1.0000	1
Singapore	0.9880	2
Busan	0.7970	3
Hong Kong	0.7727	4
Ningbo	0.7399	5
Kaohsiung	0.5027	6
Shekou	0.4969	7
Port Kelang	0.4860	8
Qingdao	0.4446	9
Antwerp	0.4426	10
Rotterdam	0.4170	11
Yokohama	0.3735	12
Tokyo	0.3141	13
Yantian	0.3110	14
Tanjung Pelepas	0.3025	15
Xiamen	0.2997	16
Kobe	0.2984	17
Hamburg	0.2785	18
Nagoya	0.2782	19
Xingang	0.2714	20

##### 4.4.4.2 Closeness centrality

A similar comparison can be carried out also for the closeness centrality, recalling that closeness centrality should be calculated on P- (and HP-) graphs. Results are reported in Table 24 for the top 20 ports in HP-graphs vs. P-graphs and in Table 25 for the 20 top ports in P-graphs vs. HP-graphs.

Table 24: top 20 ports - closeness centrality ranking results in HP-graph vs. P-graph

<i>Port</i>	<i>HP-graph</i>	<i>rank HP</i>	<i>P- graph</i>	<i>rank P</i>
Tangier	1.0000	1	0.9796	3
Algeciras	0.9796	2	0.9746	4
Antwerp	0.9526	3	0.9474	35
Rotterdam	0.9492	4	0.9490	30
Le Havre	0.9481	5	0.9642	13
Valencia	0.9464	6	0.9671	8
Singapore	0.9461	7	0.9141	90
Port Said	0.9399	8	0.9637	14
Jeddah	0.9359	9	0.9678	7
Tanjung Pelepas	0.9351	10	0.9135	93
Hamburg	0.9325	11	0.9227	75
Barcelona	0.9325	12	0.9665	10
Southampton	0.9321	13	0.9432	44
London-Gateway	0.9314	14	0.9514	27
Colombo	0.9310	15	0.8946	120
Piraeus	0.9307	16	1.0000	1
Casablanca	0.9288	17	0.9481	31
Felixstowe	0.9288	18	0.9472	36
Marsaxlokk	0.9281	19	0.9971	2
Sines	0.9238	20	0.9540	22

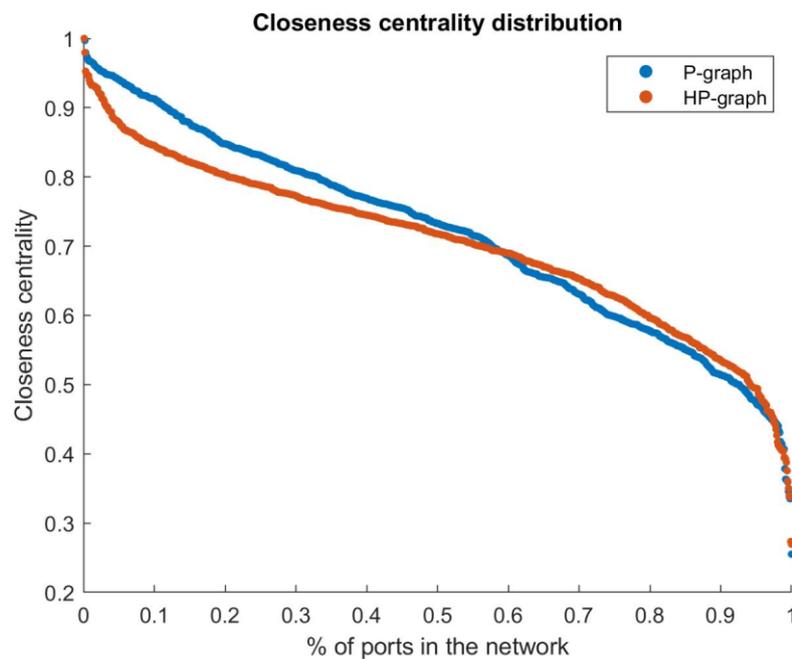
Table 25: top 20 ports – closeness centrality ranking results in P-graph vs. HP-graph

<i>Port</i>	<i>P- graph</i>	<i>rank P</i>	<i>HP-graph</i>	<i>rank HP</i>
Piraeus	1.0000	1	0.9307	16
Marsaxlokk	0.9971	2	0.9281	19
Tangier	0.9796	3	1.0000	1
Algeciras	0.9746	4	0.9796	2
Damietta	0.9731	5	0.9196	23
Genoa	0.9707	6	0.9133	25
Jeddah	0.9678	7	0.9359	9
Valencia	0.9671	8	0.9464	6
Algiers	0.9671	9	0.8475	90
Barcelona	0.9665	10	0.9325	12
Oran	0.9661	11	0.8804	46
Malaga	0.9653	12	0.9008	33
Le Havre	0.9642	13	0.9481	5
Port Said	0.9637	14	0.9399	8
Skikda	0.9611	15	0.8256	137
Fos	0.9604	16	0.9009	32
La Spezia	0.9592	17	0.8952	35
Castellon	0.9581	18	0.8802	47
Bejaia	0.9573	19	0.8222	142
Tarragona	0.9566	20	0.8423	100

The comparison between the results obtained in the two different approaches highlight the strong influence of the service frequency in the evaluation of the closeness measure. Thus, some of the busiest ports belonging to the Europe – Far East trade lane (i.e. Antwerp, Rotterdam and Singapore) exhibit better performances in the HP-graph (Table 24), whereas in the P-graph they do not even rank in the top ten ports (Table 25).

For the sake of completeness, the following figure reports the distribution of the closeness centrality across ports in the study area.

**Figure 18: Distribution of the closeness centrality in HP-graphs vs. P-graphs**



#### 4.4.4.3 4.3.1 Betweenness centrality

Consistent with the proposed formulation, betweenness centrality is calculated on both HL graphs and HP-graphs via equation (4.6) and contrasted respectively with calculations on L-graphs and P-graphs, respectively; notably, betweenness centrality (4.6) has been further normalised with respect to the maximal value to allow direct comparison. For the sake of simplicity, results are reported only with reference to the first 20 ports in each ranking, that is HL-graph in Table 26 and L-graph

Table 27, and HP-graph and P-graph in Table 28 and Table 29.

A comparison between HL-graph and L-graph results reveals that the top 10 ports in each ranking also belong to the top 20 ports of the other ranking, with an overall consistency between the two approaches. However, the betweenness centrality on HL-graphs assigns a much higher ranking to important transshipment ports, e.g., Port Kelang (12.3 MTEU in 2018, of which 7.8 MTEU transhipped) and Tanjung Pelepas (9.0 MTEU in 2018 with 8.2 MTEU transhipped), that compete directly with Singapore. This happens because the combination of the L-graph and of equation (2.4) yields a more “deterministic” behaviour, that is with most of shortest routes across the Indochinese Peninsula/Singapore area going through Singapore, whilst adoption of equation (4.4) on the HL-graph allows also considering the presence of other transshipment ports in that region. The same happens for some terminals in the Northern Range, with Hamburg and Le Havre much better ranked in the HL-graph with respect to the L-graph.

**Table 26: top 20 ports - betweenness centrality ranking results in HL-graph vs. L-graph**

<i>Port</i>	<i>HL -graph</i>	<i>rank HL</i>	<i>L - graph</i>	<i>rank L</i>
Singapore	1.0000	1	1.0000	1
Rotterdam	0.6461	2	0.5779	2
Busan	0.3838	3	0.3944	5
Shanghai	0.3365	4	0.2385	12
Hong Kong	0.3218	5	0.4332	4
Antwerp	0.3100	6	0.2751	8
Ningbo	0.2920	7	0.1784	14
Algeciras	0.2870	8	0.2765	7
TangerMed	0.2323	9	0.2505	11
Jeddah	0.2104	10	0.5698	3
Hamburg	0.2096	11	0.0640	71
Le Havre	0.2087	12	0.1178	31
Tanjung Pelepas	0.2032	13	0.0665	63
Kaohsiung	0.2004	14	0.1219	30
Bremerhaven	0.1948	15	0.2746	9
Yantian	0.1926	16	0.1134	33
Colombo	0.1687	17	0.2686	10
Jebel Ali	0.1586	18	0.1547	16
Piraeus	0.1564	19	0.1352	22
Port Kelang	0.1501	20	0.0938	44

Table 27: top 20 ports - betweenness centrality ranking results in L-graph vs. HL-graph

<i>Port</i>	<i>L - graph</i>	<i>rank L</i>	<i>HL -graph</i>	<i>rank HL</i>
Singapore	1.0000	1	1.0000	1
Rotterdam	0.5779	2	0.6461	2
Jeddah	0.5698	3	0.2104	11
Hong Kong	0.4332	4	0.3218	5
Busan	0.3944	5	0.3838	3
Algeciras	0.2765	6	0.2870	8
Antwerp	0.2751	7	0.3100	6
Bremerhaven	0.2746	8	0.1948	16
Colombo	0.2686	9	0.1687	18
TangerMed	0.2505	10	0.2323	10
Shanghai	0.2385	11	0.3365	4
Durban	0.1894	12	0.0851	38
Ningbo	0.1784	13	0.2920	7
Manzanillo (Pan)	0.1745	14	0.1277	26
Jebel Ali	0.1547	15	0.1586	19
Damietta	0.1535	16	0.0624	61
Santos	0.1391	17	0.1499	22
Jakarta	0.1369	18	0.1268	27
Dakar	0.1357	19	0.0538	72
Barcelona	0.1352	20	0.0948	34

Table 28: top 20 ports - betweenness centrality ranking results in HP-graph vs. P-graph

<i>Port</i>	<i>HP- graph</i>	<i>rank HP</i>	<i>P graph</i>	<i>rank P</i>
Singapore	1.0000	1	1.0000	1
Hong Kong	0.6201	2	0.3332	8
Ningbo	0.4651	3	0.0619	68
Le Havre	0.4571	4	0.3186	9
Yantian	0.4421	5	0.2753	11
Rotterdam	0.4242	6	0.6812	2
Shanghai	0.4035	7	0.1941	20
Jeddah	0.3844	8	0.4035	7
Algeciras	0.3759	9	0.2427	13
Cai Mep	0.3563	10	0.0784	54
Tanjung Pelepas	0.3342	11	0.4289	5
Kaohsiung	0.3102	12	0.1275	32
Port Kelang	0.2902	13	0.2419	14
Colombo	0.2808	14	0.1731	22
Port Said	0.2475	15	0.0532	89
Busan	0.2469	16	0.4627	3
Bremerhaven	0.2384	17	0.2729	12
Aalesund	0.2310	18	0.1173	36
London-Gateway	0.2254	19	0.1159	37
Antwerp	0.2202	20	0.2113	17

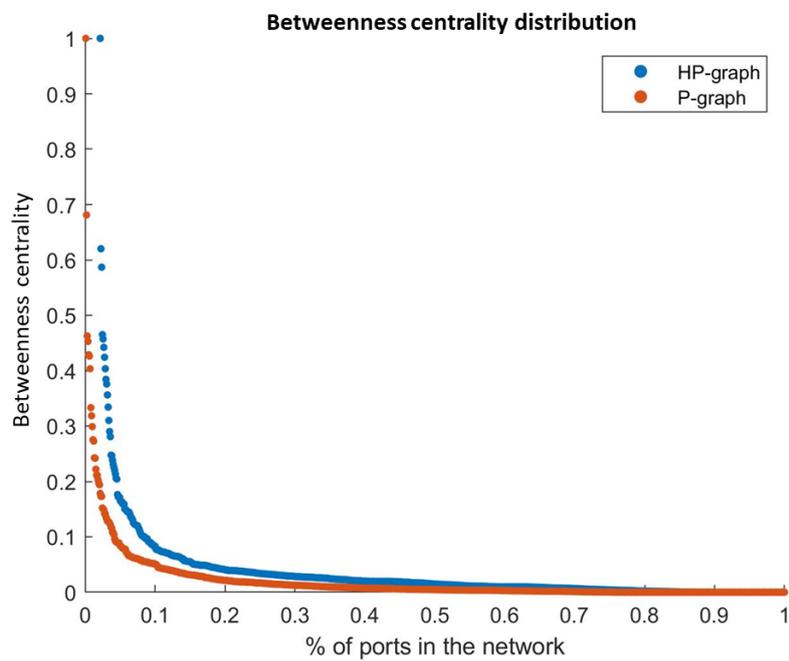
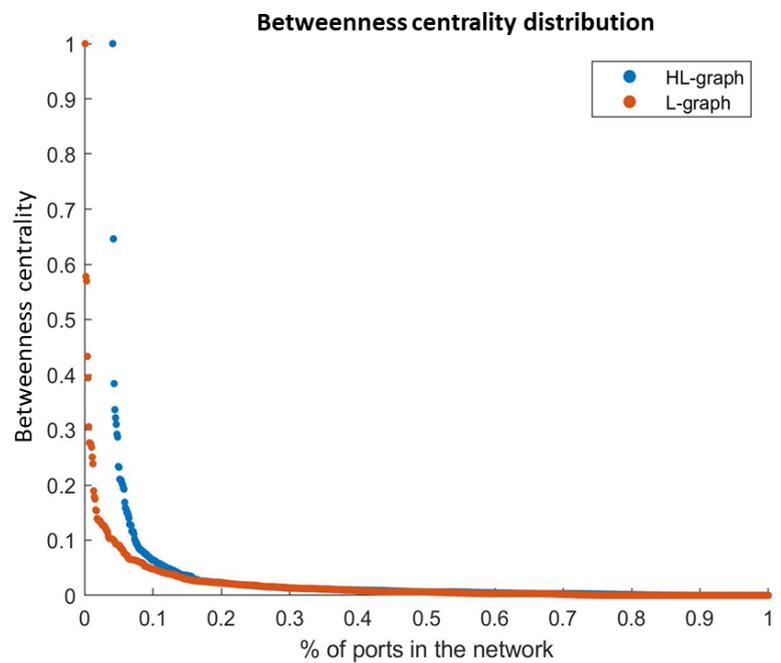
Table 29: top 20 ports - betweenness centrality ranking results in P-graph vs. HP-graph

<i>Port</i>	<i>P graph</i>	<i>rank P</i>	<i>HP- graph</i>	<i>rank HP</i>
Singapore	1.0000	1	1.0000	1
Rotterdam	0.6812	2	0.4242	7
Busan	0.4627	3	0.2469	17
TangerMed	0.4529	4	0.1445	40
Tanjung Pelepas	0.4289	5	0.3342	12
Piraeus	0.4264	6	0.1371	43
Jeddah	0.4035	7	0.3844	9
Hong Kong	0.3332	8	0.6201	2
Le Havre	0.3186	9	0.4571	5
Marsaxlokk	0.2990	10	0.1165	54
Yantian	0.2753	11	0.4421	6
Bremerhaven	0.2729	12	0.2384	18
Algeciras	0.2427	13	0.3759	10
Port Kelang	0.2419	14	0.2902	14
Sines	0.2221	15	0.1233	48
Valencia	0.2113	16	0.1590	34
Antwerp	0.2113	17	0.2202	21
Jebel Ali	0.2027	18	0.0614	113
Damietta	0.1970	19	0.1592	33
Shanghai	0.1941	20	0.4035	8

On the contrary, a comparison between HP-graph and P-graph yields more appreciable differences: two out of the 10 top ports in the HP-graph ranking (Ningbo and Cai Mep) do not belong to the top 20 ports in P-graph and, vice versa, three out of the 10 top ports in the P-graph ranking (TangerMed, Piraeus, and Marsaxlokk) do not belong to the top 20 ports according to the HP-graph.

In general, it is important to remark that hypergraphs allow accounting also for the effects of the frequency of services, that impact on the transfer waiting time at transshipment nodes. For the sake of completeness, Figure 19 illustrates the distribution of the betweenness centrality across ports for the different contrasted approaches: as expected, the distribution on HL- and HP-graphs is smoother than for the corresponding L- and P-graphs, thanks to the “fuzziness” of the indicator (4.4).

Figure 19: Distribution of the betweenness centrality in HL-graphs vs. L-graphs (top) and in HP-graphs vs. P-graphs (bottom)



## 4.5 Multimodal integration of port-to-port and landside supply models

The final “brick” to get the supply model completed, filling all gaps identified in the concerned literature, is to make a step further from the doctoral thesis by Vitillo (2011) and the paper by Cascetta et al. (2013) by inserting the proposed port-to-port supply models for containers (Chapter 4) and railway freight transport (Chapter 3) into a multimodal freight supply model. For this aim, this section describes the proposed approach for the calculation of shortest paths in multimodal freight networks with respect to a non-additive impedance, that is with respect to the main aspect non covered by existing approaches in the literature<sup>14</sup>.

Freight supply performances highly depend upon a wide range of factors, related e.g. to the nature of good, the type of vehicle, the business relationships, and so on. Thus, a proper segmentation of the supply model should be adopted, developing a specific multimodal freight model for each freight segment: in the following, reference is made to a single segment, i.e. the notation will not include any specific segment-related superscripts/subscripts.

In principle, the proposed approach can be coupled with the NODUS approach recalled in Chapter 2, that consists of exploding mode-change nodes (e.g. port container terminals connected with rail and road networks) such that multimodal links, that is change of mode within the node, are represented by specific links, as it will be recalled later.

Since non-additive impedances can be treated with explicit path enumeration and calculation of shortest path amongst enumerated paths (as recalled in Section 2.2), the key point for linking the proposed approach in Chapter 4 for maritime container networks and in Chapter 3 for railway freight networks is how to make this explicit path enumeration more tractable and less computationally demanding. For this aim, the proposed approach starts from the monomodal networks for each concerned freight mode, that allows creation of proper *macrolinks*, each representing a direct connection between an origin-destination pair of that mode: this allows associating macrolinks with non-additive impedances related to that origin-destination pair. As a result, non-additive impedances in the initial monomodal network are turned into additive impedances in the macrolink-based monomodal network, obviously at the price of increasing the number of links in the network. Practically speaking, an example of macrolink is a generic link of the P-graph introduced in Section 4.3.2.2

A formal presentation of the proposed approach is reported in the following. Let be  $M$  a set of available freight modes indexed by  $m \in 1 \dots n_m$ , comprising the road mode and other non-road modes. Each mode  $m$  is represented by a “real” monomodal graph  $\Gamma_m \equiv \{N_m, L_m\}$  being  $N_m$  and  $L_m$  the sets of nodes and links, respectively. Without loss of generality, the set of nodes associated with the road mode encompasses a subset  $C$  of centroids representing traffic analysis zones (TAZs). Furthermore, within the study area, a set  $T$  of specific loading/unloading points, called terminals, with cardinality  $n_\tau$ , allows transfers across modes. Each terminal  $\tau \in T$  is associated with a set  $M_\tau \subseteq M$  of  $n_{m\tau}$  modes, and for each mode  $m \in M_\tau$  a

---

<sup>14</sup> Most of the content of this section is reprised by a conference paper presented at the 96th Annual Meeting of the Transportation Research Board, Washington (D.C.), USA, January 2017.

corresponding node  $\tau_m \in N_m$  exists. The union of all terminal nodes  $\tau_m$  for the mode  $m$  leads to the set  $T_m$  of all terminals of mode  $m$ , with cardinality  $n_m$ .

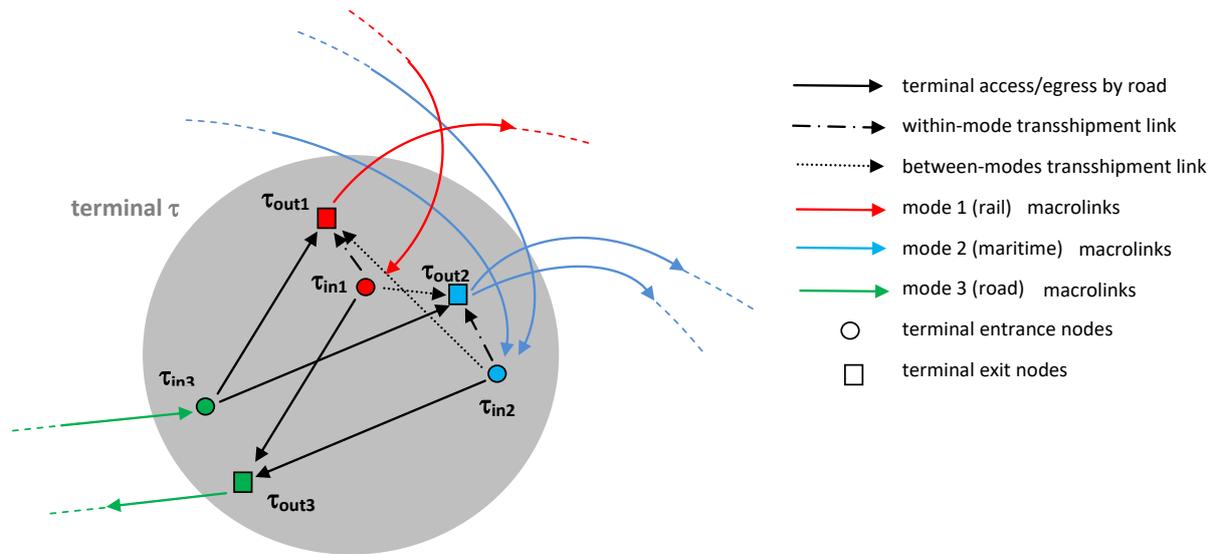
In the following, without loss of generality, attention will be focused on a specific impedance. If the impedance is additive, then there is no need to build a macrolink-based monomodal network, and the “real” monomodal graph can be used directly in the multimodal network. If the impedance is sub-additive, in accordance with the above, the original “real” graph for the mode  $m$  should be replaced with another graph  $\Gamma_m^* \equiv \{N_m^*, L_m^*\}$  in which  $N_m^*$  contains only origins and destinations and  $L_m^*$  all related macrolinks, being a macrolink  $l^* \in L_m^*$  a direct connection between  $o$  and  $d$ , with  $o, d \in N_m^*$ . Specifically, the set of nodes  $N_m^*$  includes the set of all nodes belonging to relevant o-d pairs for the mode  $m$ . In general, for the road mode, o-d pairs will include both terminals and centroids as origins and destinations, i.e.  $N_{road}^* \equiv C \cup T_{road}$ , whilst for non-road modes normally only terminals will be origins and/or destinations, i.e.  $N_m^* \equiv T_m \forall m \neq road$ . However, this is not a strict requirement.

In terms of macrolinks  $L_m^*$ , a distinction between road and non-road modes should be introduced. For the road mode, the connection between o-d pairs depends only on the presence of road infrastructures. For the other modes, the connection between o-d pairs (normally terminals, in accordance with the above) depends not only on the presence of infrastructures, but also on the availability of concerned freight services. Thus, building the macrolinks means basically building all connections between terminals depending on the available services, exactly as described for maritime container networks in Section 4.3.2.

In the case of direct services, a macrolink directly connecting the two concerned terminals suffices. In the case of services calling at various terminals (e.g. a maritime container service calling at various ports), let  $\tau_1 \dots \tau_n \in T_s$  be the sequence of those terminals, where  $T_s$  is the set of terminals of service  $s$ , each terminal  $\tau_i \in T_s$  should be connected with every other terminal  $\tau_j \in T_s$  with  $j \neq i$  so as to associate all possible connections amongst terminals within  $T_s$  with the appropriate sub-additive impedances, i.e. leading to  $|T_s| \cdot (|T_s| - 1) / 2$  two-way macrolinks.

Obviously, this implies that a pair of terminals will be connected potentially by many macrolinks, which is of course not a problem. Overall, each macrolink connecting two nodes  $o$  and  $d$  with mode  $m$  – either road or non-road – will be associated with an impedance calculated on the “real” network of mode  $m$ . It is worth noticing that it always occurs  $|N_m| > |N_m^*|$  and  $|L_m| \ll |L_m^*|$ , that is turning non-additive impedances in the real graph  $\Gamma_m$  into additive impedances in  $\Gamma_m^*$  through macrolinks comes at the price of a substantial increase in the number of links, equal at most to  $n_{od}^m(n_{od}^m - 1)$  being  $n_{od}^m$  the number of origin/destination nodes for the mode  $m$ . In this respect, the maximum number of macrolinks can be limited, at least landside, by considering for instance upper thresholds on the total travel time by road, that is excluding all macrolinks yielding a drivers’ rest period, usually after 10 hours of net driving time.

**Figure 20 : Illustrative example of the approach pursued by the NODUS model for a three-modes example (road, sea, maritime)**



In formal terms, each mode  $m$  using the terminal  $\tau$  is associated with two nodes, an “entering node”  $\tau_{in,m}$  and an “exiting node”  $\tau_{out,m}$ . Therefore, each terminal  $\tau \in T$  is modelled by means of a number of nodes equal to  $2 * n_{m\tau}$ , possibly connected by the following types of links for each generic mode  $m$ :

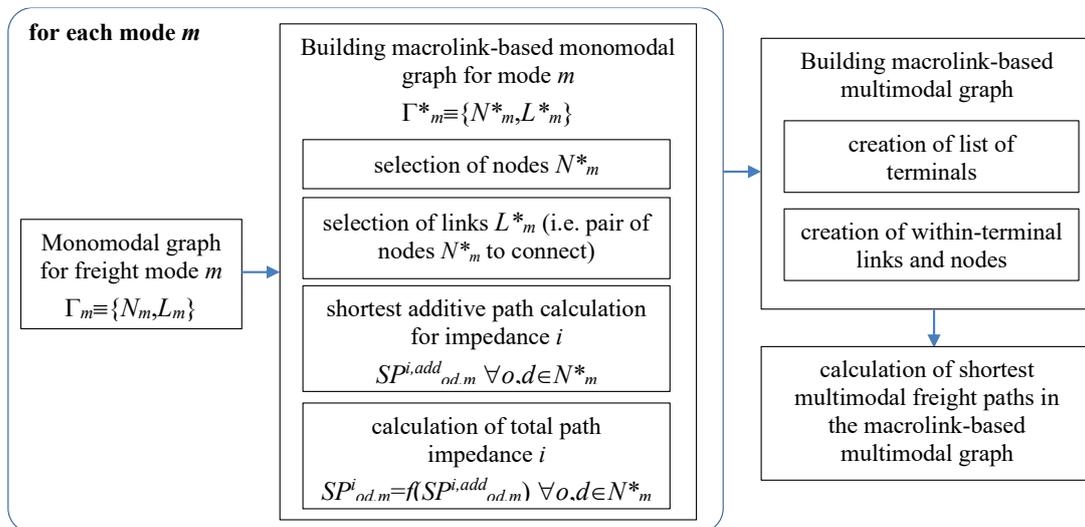
- the backward star of the entering node  $\tau_{in,m}$  includes all macrolinks of mode  $m$  going to the terminal  $\tau$ . Specifically, if the mode is road, those macrolinks are either of type  $o-\tau_{in,road}$ , i.e. they represent the overall shortest paths from each origin  $o \in C$  to the terminal  $\tau$ , or of type  $\tau'_{out,road}-\tau_{in,road}$ , being  $\tau'$  another terminal connected by road. If the mode is not road, those macrolinks represent all services of mode  $m$  arriving at terminal  $\tau$  from other terminals of the same mode, i.e. links of type  $\tau'_{out,m}-\tau_{in,m}$  with  $\tau' \neq \tau \in T_i$ , in accordance with the definition of macrolink reported above. Notably, in the approach by Jourquin et al. (1996) and Beuthe et al. (2001) the backward star of the entering node  $\tau_{in,m}$  includes simply links of the monomodal network of mode  $m$ , possibly duplicated into “virtual links” to account for different freight demand services;
- the forward star of the exiting node  $\tau_{out,m}$  is made by all the concerned links of mode  $m$  leaving from terminal  $\tau$ . In accordance with the above, if the mode is road, those macrolinks are either of type  $\tau_{out,road}-d$ , i.e. they represent the overall shortest paths from the terminal  $\tau$  to a given destination  $d \in C$ , or of type  $\tau_{out,road}-\tau'_{in,road}$ , with  $\tau' \neq \tau \in T_{road}$ , so as to represent road land bridge. If the mode is not road, those macrolinks represent all services of mode  $m$  leaving from terminal  $\tau$ , i.e. of type  $\tau_{out,m}-\tau'_{in,m}$  with  $\tau' \neq \tau \in T_m$ .
- the forward star of the entering node  $\tau_{in,m}$  is defined as follows:
  - if the mode is road, it is made up by all links representing connections with non-road other modes, i.e. by all links of type  $\tau_{in,road}-\tau_{out,m} \forall m \neq road \in T\tau$ . Each

of such links represents access from the road mode to other modes, and therefore all related first access terminal impedances should be associated with those links. Notably, the connection  $\tau_{in,road} - \tau_{out,road}$  should be prohibited, so as to avoid undesired effects. Recalling the above discussion, and restricting attention to the travel time as impedance, each macrolink is associated with the shortest total path, inclusive of non-additive impedances.

- If the mode is not road, it is made up by three types of links, exactly reprising the rationale underlying Jourquin et al. (1996) and Beuthe et al. (2001). The first is a possible link representing the connection with the road mode, i.e. a link  $\tau_{in,m} - \tau_{out,road}$  which represents egress from the mode  $m$  towards the road mode, therefore it will be associated with all related impedances (e.g. final custom clearance times and costs in the case of a terminal container). The second is a within-mode  $m$  transshipment link of type  $\tau_{in,m} - \tau_{out,m}$  which is associated with all within-mode transshipment impedances (e.g. in the previous example of the container terminal, only transshipment times and costs will be taken into account, without customs impedances). The third is a set of links of type  $\tau_{in,m} - \tau_{out,m'}$  with  $m' \neq road$ ,  $m \in T_\tau$  representing the between-modes transshipment not involving road (e.g. from sea to rail in a container terminal), again with specifically designed impedances.
- the backward star of the  $\tau_{out,m}$  node includes all links from the entering nodes of all modes connected with  $m$ , either with a first access link, a within-mode transshipment link and/or a between-modes transshipment link. Obviously, if the mode is road, the backward star of the node  $\tau_{out,road}$  cannot encompass the link  $\tau_{in,road} - \tau_{out,road}$  for the same reason explained above.

Overall, the proposed approach is depicted at a glance in the flowchart in Figure 21.

**Figure 21 : Flowchart for the implementation of the proposed approach**



As a result, it is important to underline that the proposed modified network with macrolinks contains only additive costs, which however include all non-additive impedances in each monomodal network, and therefore standard shortest path algorithms can be applied for calculation of shortest paths. In addition, in principle, this network can also be used for standard assignment procedures applied in transport engineering. In that respect, in accordance with the above, it should be only noted that once performed the assignment on the macrolink, the resulting flows should be reported back on the single monomodal links of each single monomodal network. Notice also that one might not consider in the macrolink-based multimodal network the direct connection between centroids via the road mode: this will force the shortest path algorithm to look necessarily for a multimodal route.

Marzano et al. (2017) proposed a real-size application of the approach on a huge Euro-Mediterranean multimodal network, showing the viability of the proposed approach and illustrating its effectiveness in detecting shortest multimodal paths.

## 5 Container shipping services: descriptive statistics and concentration analysis

The implementation of the worldwide container services' network enabled interesting applications to understand the shipping market structure at different levels. In this respect, an in-depth analysis of the worldwide market of liner container liner services through the implementation of statistical analysis is provided in this chapter.

In detail:

- section 5.1 provides a complete analysis of the container shipping services supply, deployed between 2018 and 2020, focusing on the fleet composition and the carrying capacity of both main shipping carriers and alliances, highlighting the extent of their market by trade region;
- section 5.2 focuses on the results of the Herfindahl-Hirschman Index applied to the liner service network, assessing the market concentration of main trade regions as well as single ports.

### 5.1 Container shipping services supply

The container transport supply consists of several liner operating on different worldwide trade lanes. These services feature many characteristics such as geographical coverage, carrying capacity and fleet composition, operating shipping carriers and ocean alliances. According to these features, the descriptive statistics presented in the following propose an in-depth analysis of the worldwide market of liner container liner services. In addition, to highlight market dynamics, the analyses adopt a twofold perspective, namely individual shipping companies and ocean alliances.

Generally, the transport of containerized freight in global trade lanes is operated by different types of services (e. g. container & breakbulk, container & roro, full container...). Referring to all the available types of container services, Table 30 reports the yearly liner supply amount between 2018 and 2020 (data refers to the autumn season).

**Table 30: Number of active container services and their yearly variation**

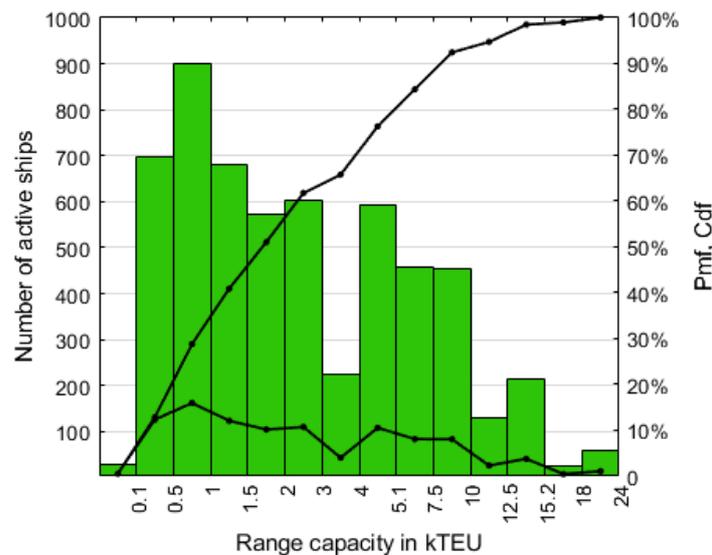
# active container services			variation		
2018	2019	2020	2018-2019	2019-2020	2018-2020
1777	1769	1686	-0.5%	-4.9%	-5.4%

Table 30 highlights a 5.4% reduction of the number of operating services from the 2018 to the 2020. It is worth citing that this analysis does not account for blank sailings (mostly occurred

in the 2020) but refers to the official scheduled liner services. The purpose, indeed, is to reveal basic adjustments in the liner services in terms of long terms planning.

Also, shipping services provide vessel capacity on which containers are transported to their destination ports. The fleet of liner shipping services consists of several types of vessels. The bulk of containers are usually transported by large cellular vessels which are classified based on their TEU size range. Figures 22 to 24 show the capacity breakdown of the total container fleet between 2018 and 2020, according to a standard vessel grouping.

**Figure 22: active ships by TEU size range – 2018**



**Figure 23: active ships by TEU size range – 2019**

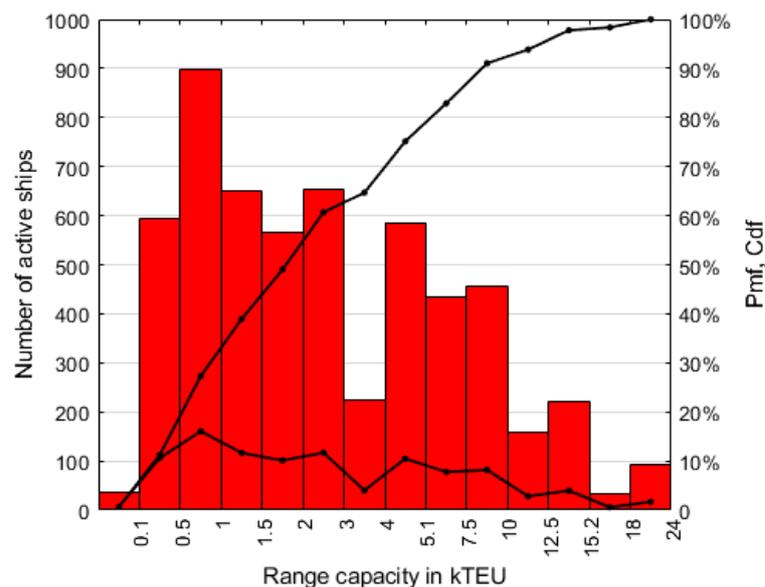
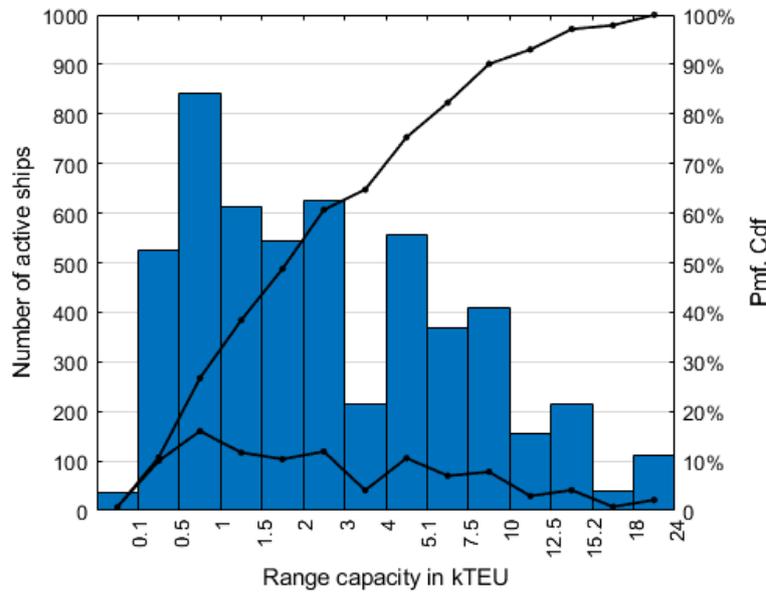
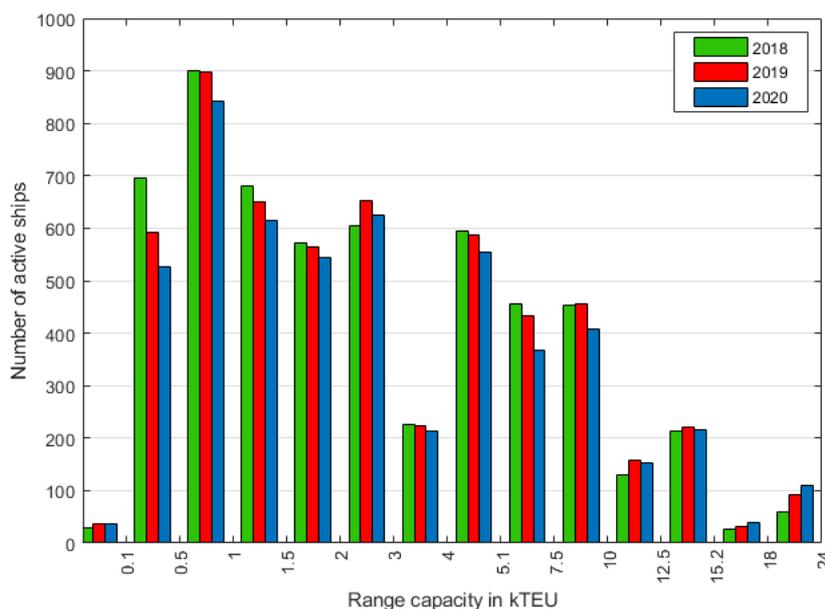


Figure 24: active ships by TEU size range – 2020



During the last years, due to the increasing container trade, the size of container ships has grown reaching the threshold of 24.000 TEUs. The reallocation of container ships among shipping services led to an increase in the number of vessels from 10.000 TEUs onwards. In detail, from 2018 to 2020, the number of vessels sized between 18.000 and 24000 TEUs has increased of 51 units, providing an additional 95% of fleet capacity. Whereas the other size ranges reveal a decrease in the number of ships, especially the 100-499 TEUs class, which shows a loss of 24% of units and a -23% of total capacity.

Figure 25: active ships by TEU size range between 2018 and 2020



Moreover, taking a close look at the fleet capacity, it is possible to see how the worldwide container services have been affected by the coronavirus crisis. Indeed, from 2018 to 2019 the total fleet capacity grew from 20.32 million TEUs to 21.4 million TEUs. However, despite the increased number of *megamax* cellular vessels reported in 2020, the total fleet capacity decreased to 20.74 million of TEUs, performing a loss of 3% compared to the total amount of 2019 (Table 31).

**Table 31: Fleet capacity breakdown by TEU size range**

TEU range	2018		2019		2020	
	Units	fleet capacity [kTEU]	Units	fleet capacity [kTEU]	Units	fleet capacity [kTEU]
18000-24000	60	1154	92	1812	111	2255
15200-17999	26	431	32	538	40	671
12500-15199	213	2911	220	3022	216	2977
10000-12499	130	1369	157	1675	154	1657
7500-9999	453	3992	457	4027	409	3611
5100-7499	456	2830	433	2690	368	2292
4000-5099	594	2692	586	2652	555	2515
3000-3999	226	787	224	780	214	745
2000-2999	604	1526	653	1660	625	1593
1500-1999	573	982	565	971	545	942
1000-1499	681	786	651	748	614	706
500-999	899	658	898	657	842	615
100-499	696	203	593	177	526	157
0-100	30	1	36	2	37	2
<b>Total</b>	5641	20322	5597	21409	5256	20737

It is worth noting that *post panamax plus* vessels accounted for largest share of fleet capacity for all the three years (19.64% in the 2018, 18.81% in the 2019 and 17.42% in the 2020), followed by the *new panamax* vessels (holding at 14%).

### 5.1.1 Services and fleet composition of top ten carriers

From the perspective of shipping companies, the market is dominated by the ten largest ocean carriers. The number of active services for each of the top ten carriers and their shares among total services of top ten are reported in the following table:

**Table 32: overview of container services operated by top ten shipping carriers**

Carriers	# active services			share among top ten		
	2018	2019	2020	2018	2019	2020
<b>Maersk Line</b>	177	175	149	20.2%	17.4%	15.7%
<b>MSC</b>	176	186	158	20.1%	18.5%	16.7%
<b>COSCO</b>	131	147	163	15.0%	14.7%	17.2%
<b>CMA CGM</b>	166	184	163	19.0%	18.3%	17.2%
<b>Hapag-Lloyd</b>	63	79	81	7.2%	7.9%	8.5%
<b>ONE</b>	0	64	68	0.0%	6.4%	7.2%
<b>Evergreen Line</b>	71	77	69	8.1%	7.7%	7.3%
<b>Yang Ming</b>	42	41	42	4.8%	4.1%	4.4%
<b>HMM</b>	24	24	26	2.7%	2.4%	2.7%
<b>Zim</b>	25	26	29	2.9%	2.6%	3.1%
<b>total</b>	875	1003	948			

It is possible to spot that main carriers such as Maersk and MSC experienced a decrease in the number of services deployed from 2018 to 2020. Conversely, COSCO, Hapag-Lloyd exhibit a different trend. The ONE shipping company follows an increasing trend over the three years since it came to market in 2018: in 2019 it gained 6.4% among the top ten carriers and this percentage increased in 2020 (7.2%).

Moving on the analysis of fleet deployed on liner services, Table 33 ranks top ten carriers according to their total fleet capacity in million TEUs. The percentage of the aggregate carrying capacity of the top ten accounted for 66% of the total worldwide fleet capacity in the 2018. This percentage even increased in 2019 and 2020, reaching 73%. Among these bigger players, the top three carriers, namely Maersk, MSC and COSCO supplied 40% of the total worldwide container ship capacity.

Furthermore, most of the top ten shipping carriers are involved in the three main ocean alliances, as set out in the next paragraph.

**Table 33: active vessels and fleet capacity of top 10 ocean carriers**

Ocean Carriers	# active ships			fleet capacity [MTEU]		
	2018	2019	2020	2018	2019	2020
<b>Maersk</b>	529	527	477	3.2	3.3	3.2
<b>MSC</b>	469	484	477	3.0	3.2	3.1
<b>COSCO</b>	305	321	359	1.8	2.0	2.1
<b>CMA CGM</b>	295	324	308	1.7	1.8	1.7
<b>Hapag-Lloyd</b>	186	231	211	1.4	1.6	1.5
<b>ONE</b>	0	204	191	0	1.4	1.4
<b>Evergreen Line</b>	189	195	179	1.0	1.2	1.1
<b>Yang Ming</b>	96	91	79	0.6	0.6	0.5
<b>HMM</b>	60	72	52	0.4	0.4	0.4
<b>Zim</b>	58	48	37	0.3	0.2	0.2

Analysing the number of active ships, it is possible to underline a negative trend which affects the majority of shipping companies between the 2018 and the 2020. Indeed, all carriers but COSCO, show a reduction of the number of active vessels, especially the ZIM company, a niche player with an overall reduction of 36%. A possible cause is the worldwide pandemic crisis due to Covid-19 virus. Indeed, to find alternative strategies to face potential loss, shipping companies resorted to blank sailings to discipline the capacity. Moreover, it is a matter of fact that the entry into force of the IMO 2020 regulation affected the vessels supply. Indeed, to comply with the new limit of sulphur for marine fuel oil, ships are periodically taken out of service (especially the biggest ones) to prepare for new low sulphur fuel and for the scrubber installation.

Moving on the fleet composition, the next figures show top ten carriers' fleet breakdown in 2018, 2019 and 2020 respectively. Based on their carrying capacity, the vessels are classified by 8 different size, namely: <2000 TEU, 2000-2999 TEU, 3000-5099 TEU, 5100-7499 TEU, 7500-9999 TEU, 10000-15099 TEU, 15100-17999 TEU, 18000-24000 TEU.

Figure 26: Breakdown of operated capacity by TEU size range for Top 10 ocean carriers in 2018

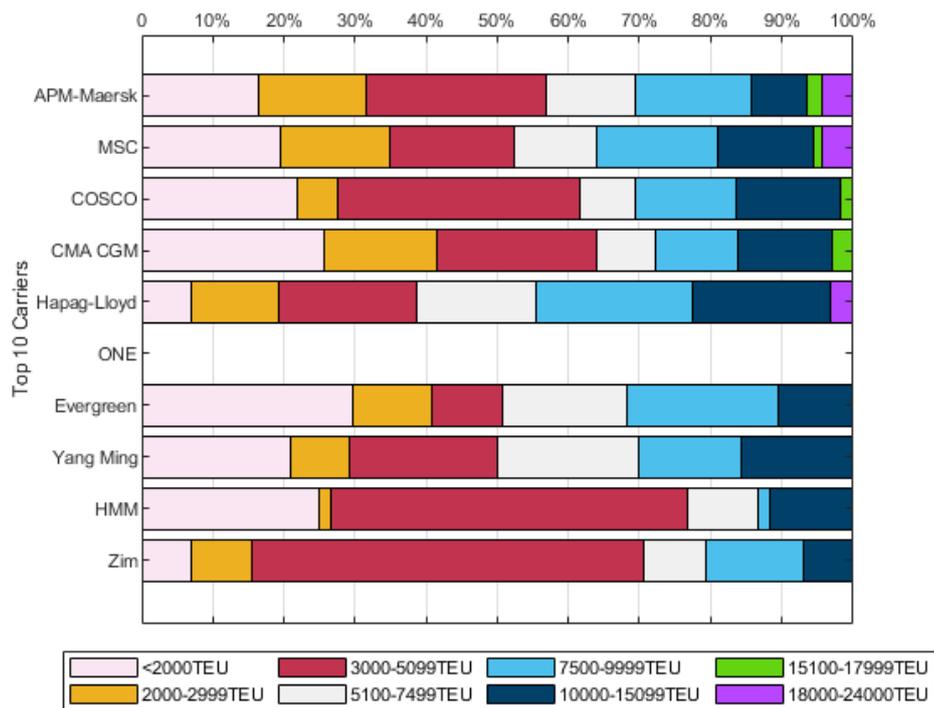
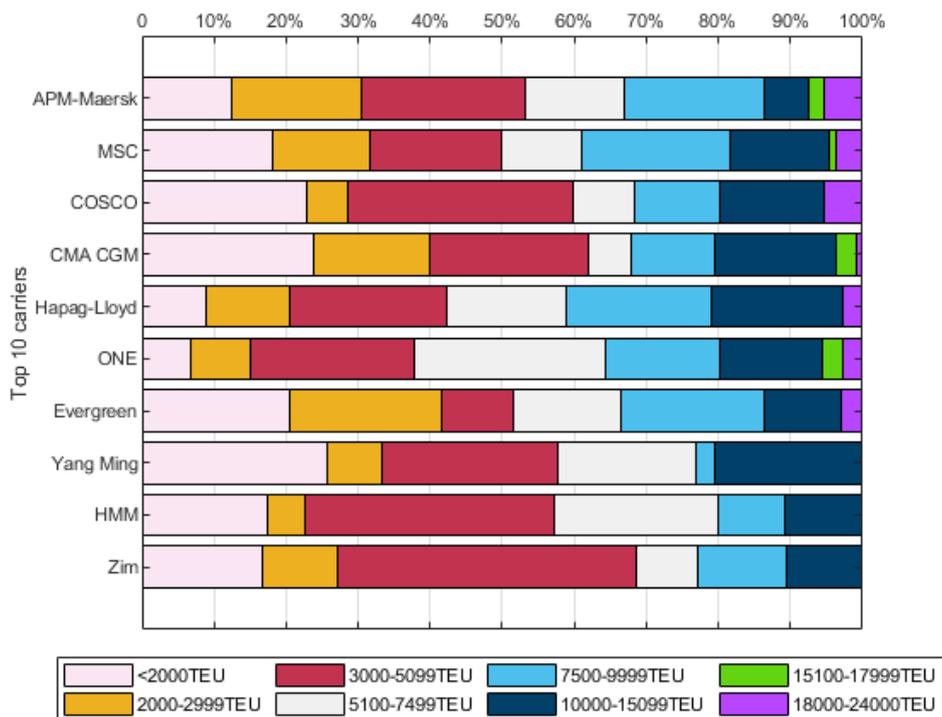
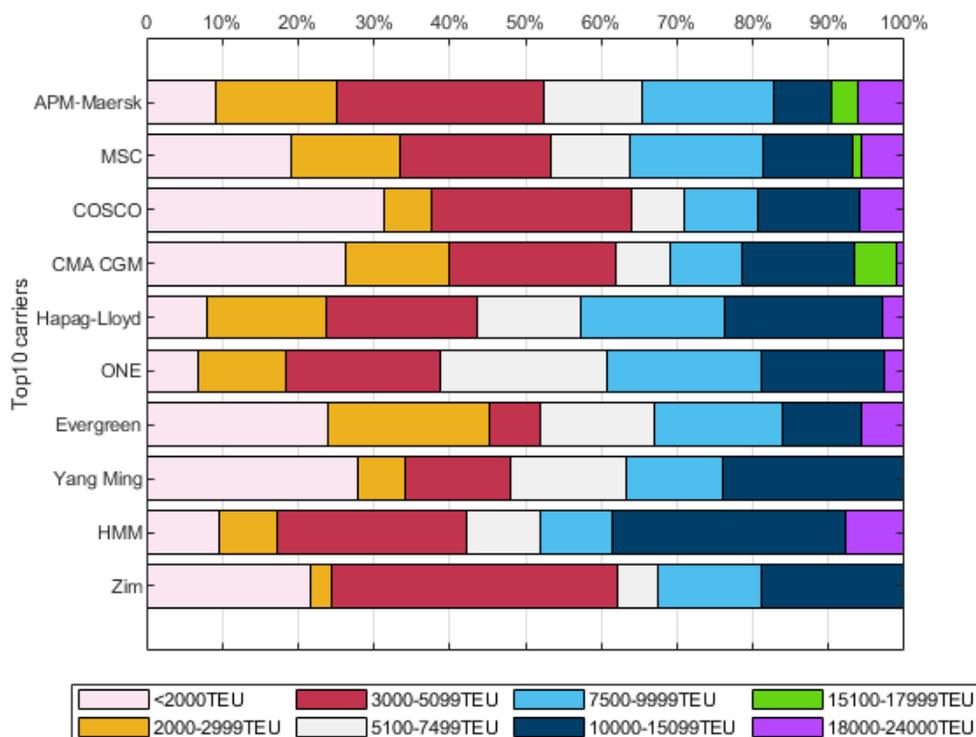


Figure 27: Breakdown of operated capacity by TEU size range for Top 10 ocean carriers in 2019



**Figure 28: Breakdown of operated capacity by TEU size range for Top 10 ocean carriers in 2020**



The top two carriers, namely Maersk and MSC, have a similar fleet composition which remained stable over the period 2018-2020. This fleet ranges among all available vessel sizes and shows a considerable share of over 18,000 TEU ships (6% for both carriers), with 11 *megamax* vessels (24,000 TEUs) belonging to MSC. Similarly, COSCO and CMA-CGM share a comparable fleet composition between 2018 and 2019. From 2019, COSCO increased the number of vessels bigger than 18,000 TEU and it is expected to change more in the next future thanks to the recent order of five 23,000 TEU container vessels. Conversely, Hyundai Merchant Marine (HMM) and ZIM, exhibit a greater number of middle-sized ships (between 3,000 and 5,100 TEU) which accounted for the 50% of the total fleet in 2018. This percentage has dwindled during the last two years while the number of vessels between 10,000-15,099 TEU has increased.

### 5.1.2 Services and liner fleet of ocean carriers' alliances

In liner services' industry, shipping lines often engage cooperative arrangements to increase their competitive advantages. There are different cooperative approaches such as slots and vessels sharing, joint operated services (2M + ZIM) and shipping alliances.

Nowadays, the shipping market is dominated by three main alliances, namely 2M, formed by Maersk and MSC, the Ocean Alliance which includes CMA CGM, COSCO Shipping, Evergreen Line and OOCL, and the Alliance made up of Hapag-Lloyd, K Line, NYK, Yang Ming Marine Transport Corp. It can be seen that, most of the top ten carriers are involved in

these alliances.

Compared to the other, Ocean Alliance operates more services, followed by 2M and The Alliance, as shown in Table 34.

**Table 34: overview of operated active services by ocean alliance**

Alliance	# active services			Δ		
	2018	2019	2020	% 2020-2019	% 2019-2018	% 2020-2018
<b>2M</b>	24	28	23	-22%	14%	-4%
<b>Ocean Alliance</b>	41	39	38	-3%	-5%	-8%
<b>The Alliance</b>	32	33	35	6%	3%	9%

In 2020 both the Ocean Alliance and the 2M experienced a reduction of liner services compared to the 2019. Conversely, The Alliance shows an opposite trend with a 6% increase of operated services from 2019 to 2020.

As previously stated, top ten shipping carriers are involved in three main alliances. The following tables show the extent of the involvement of each carrier in the respective alliance in terms of number of co-operated services.

**Figure 29: number of shipping carriers' liner services addressed to ocean alliances – 2018**

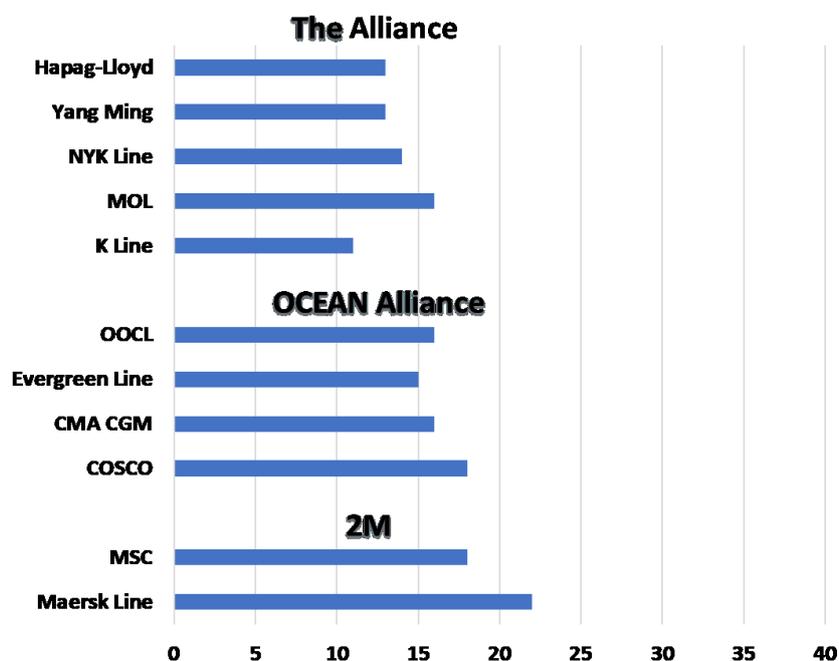


Figure 30: number of shipping carriers' liner services addressed to ocean alliances - 2019

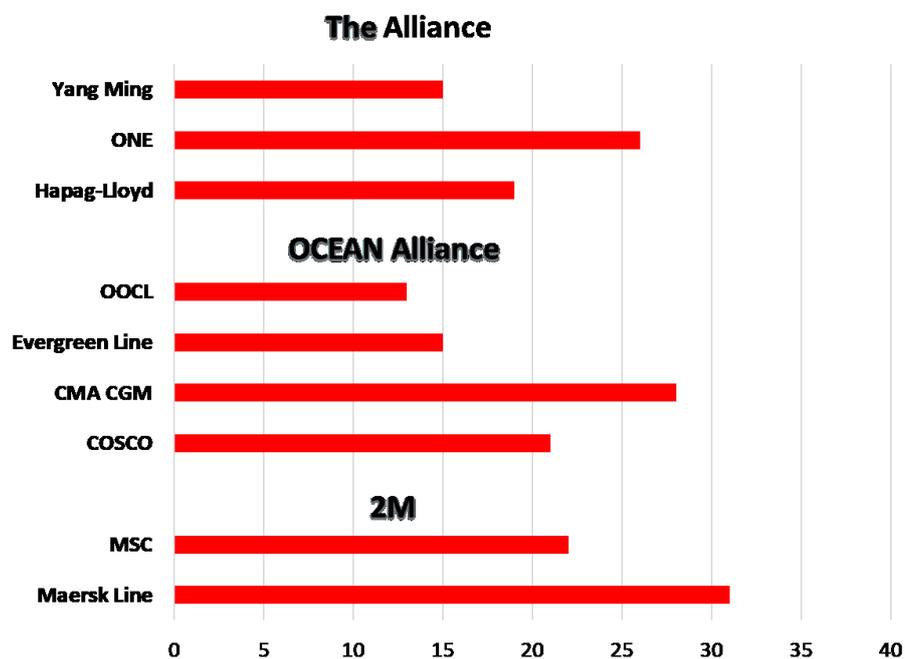
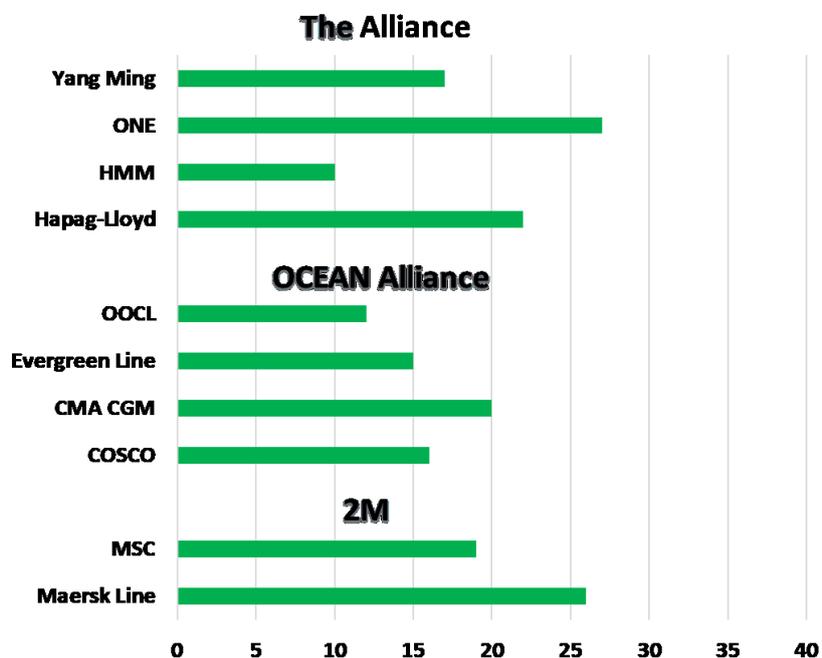


Figure 31: number of shipping carriers' liner services addressed to ocean alliances - 2020



Moving on the analysis of deployed fleet, Table 35 shows the total active vessel capacity operated by the three main shipping alliances in 2018, 2019 and 2020 and their share of the worldwide fleet. The total active fleet followed a positive trend from the 2018 to the 2019 (+ 5.3%) reaching the amount of 21.4 million TEUs in 2019. However, the losses experienced

during the last year, led to a drop of 3.1% of the global active capacity which currently amounts to 20.7 million TEUs.

From the ocean alliances perspective, both 2M and the Ocean Alliance showed a fairly stable trend, accounting respectively for 13% and 14% of the total active fleet capacity. Whereas, in the last year, The Alliance shows a slightly increase (+2.5%) of its share at the expense of single carriers (-2.7%).

**Table 35: Total active fleet capacity of ocean alliances**

	total active fleet [kTEU]			share of global total active fleet (%)			Δ		
	2018	2019	2020	2018	2019	2020	2020-2019	2019-2018	2020-2018
<b>2M</b>	2701	2707	2700	13.3%	12.6%	13.0%	0.4%	-0.6%	-0.3%
<b>OCEAN ALLIANCE</b>	2933	3119	2992	14.4%	14.6%	14.4%	-0.1%	0.1%	0.0%
<b>THE ALLIANCE</b>	1835	1935	2376	9.0%	9.0%	11.5%	2.4%	0.0%	2.4%
<b>single carriers</b>	12853	13648	12669	63.2%	63.7%	61.1%	-2.7%	0.5%	-2.2%
<b>total active capacity</b>	<b>20322</b>	<b>21409</b>	<b>20737</b>				<b>-3.1%</b>	<b>5.3%</b>	<b>2.0%</b>

The following tables report a more detailed analysis of the fleet capacity share of each participant involved in the main shipping alliances.

**Table 36: fleet capacity share by carrier in the 2M alliance**

	2M		
	2018	2019	2020
<b>Maersk Line</b>	57%	58%	58%
<b>MSC</b>	43%	42%	42%

**Table 37: fleet capacity share by carrier in the Ocean Alliance**

	OCEAN ALLIANCE		
	2018	2019	2020
<b>CMA CGM</b>	28%	29%	25%
<b>COSCO</b>	33%	32%	36%
<b>EVERGREEN LINE</b>	24%	25%	26%
<b>OOCL</b>	15%	13%	13%

**Table 38: fleet capacity share by carrier in The Alliance**

	THE ALLIANCE		
	2018	2019	2020
<b>HAPAG - LLOYD</b>	40%	41%	28%
<b>HMM</b>	-	-	14%
<b>K LINE</b>	16%	-	-
<b>Mitsui-OSK L. (MOL)</b>	8%	-	-
<b>ONE</b>	-	35%	41%
<b>NYK LINE</b>	13%	-	-
<b>YANG MING</b>	23%	24%	17%

Referring to the 2M alliance, the fleet shares of both MSC and Maersk remained stable over the investigated period. The involvement of COSCO in the Ocean Alliance' operated capacity increased during last year, mostly due to the addition of 21 units of 15000 TEU-sized vessels to the total fleet.

The capacity breakdown of The Alliance is featured by different changes. Indeed, during the last year, HMM joined the alliance contributing for 14% of the total fleet. Similarly, ONE came to market in the 2019 and, during the 2020, increased its capacity share of a 6%. Conversely, both Hapag-Lloyd and YANG MING experienced a decrease of their capacity share from 2018 to 2020.

Moving to a deeper analysis, the following tables define for each carrier the share of total capacity addressed to joined services and the share within the worldwide fleet capacity.

**Table 39: carrying capacity of each carrier participating in the 2M and percentage addressed to alliance services**

	carrying capacity [kTEU]			% addressed to the alliance			% within the global fleet		
	2018	2019	2020	2018	2019	2020	2018	2019	2020
<b>Maersk Line</b>	3105	3457	3195	49.9	45.7	49.1	15.3	16	15.4
<b>MSC</b>	2994	3432	3141	38.5	32.8	36	14.7	16	15.1
	<b>6099</b>	<b>6889</b>	<b>6336</b>						

Maersk Line and MSC company addressed to the alliance services a quite stable share of capacity over the three years.

**Table 40: carrying capacity of each carrier participating in the Ocean Alliance and percentage addressed to alliance services**

	carrying capacity [kTEU]			% addressed to the alliance			% within the global fleet		
	2018	2019	2020	2018	2019	2020	2018	2019	2020
<b>CMA CGM</b>	1670	1862	1738	49.7	48.8	42.5	8.2	8.7	8.4
<b>COSCO</b>	1762	2054	2099	54.2	49	50.6	8.7	9.6	10.1
<b>EVERGREEN LINE</b>	1030	1194	1122	68.1	65.8	70.6	5.1	5.6	5.4
<b>OOCL</b>	656	657	649	68.2	63.7	61.3	3.2	3	3.1
	<b>5188</b>	<b>5767</b>	<b>5608</b>						

Looking at the situation of the Ocean Alliances, the carriers providing the most part of the co-operated capacity are CMA CGM e COSCO, with 1,738 kTEU and 2,099 kTEU in 2020 respectively.

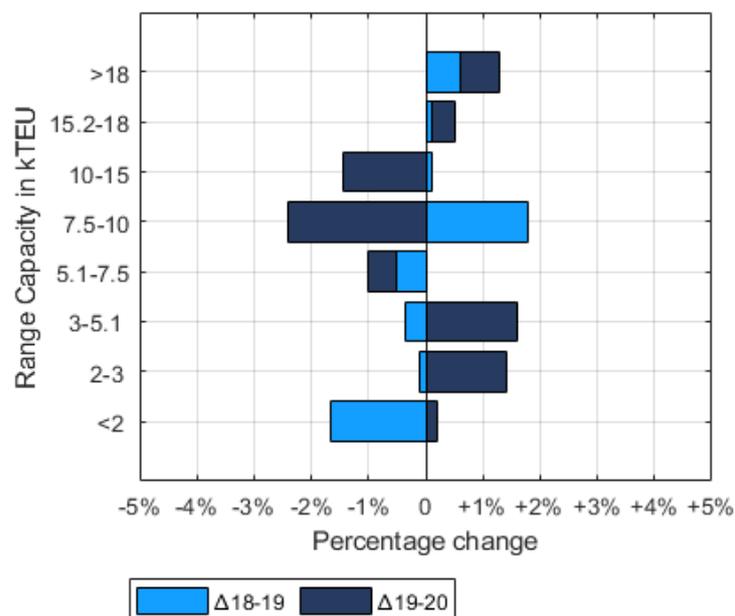
The Alliance shows different trends over the three years, mainly due to merges and changes that affected the alliance participants (Table 41). However, Hapag Lloyd and Yang Ming operated a stable share of capacity on the alliance services over the years.

**Table 41: carrying capacity of each carrier participating in The Alliance and percentage addressed to alliance services**

	carrying capacity [kTEU]			% addressed to the alliance			% within the global fleet		
	2018	2019	2020	2018	2019	2020	2018	2019	2020
<b>HAPAG - LLOYD</b>	1360	1644	1513	54.4	47.9	44.3	6.7	7.7	7.3
<b>HMM</b>	-	-	438	-	-	73.6	-	-	2.1
<b>K LINE</b>	342	-	-	85.2	-	-	1.7	-	-
<b>MOL</b>	534	-	-	27.7	-	-	2.6	-	-
<b>ONE</b>	-	1491	1365	-	45.6	72.2	-	7	6.6
<b>NYK LINE</b>	532	-	-	44.7	-	-	2.6	-	-
<b>YANG MING</b>	584	601	515	71.7	78	77.3	2.9	2.8	2.5
	<b>3352</b>	<b>3736</b>	<b>3831</b>						

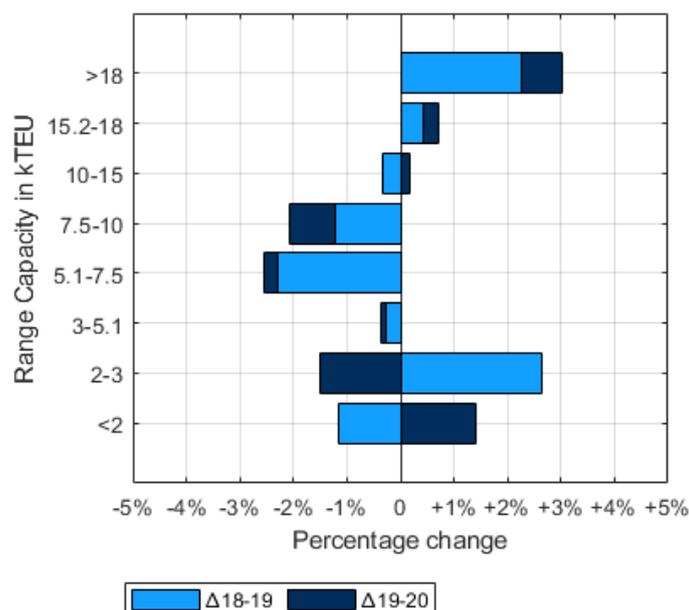
Moreover, it is interesting to analyse how the fleet breakdown changed between 2018 and 2019 to understand possible strategical choices in the service structure of the three alliances.

**Figure 32: changes in the fleet composition (2018-2019) and (2019-2020) – 2M**



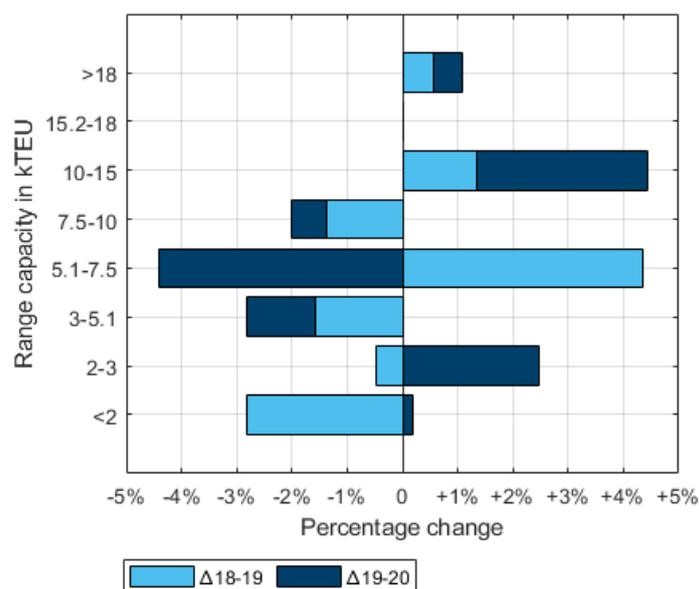
The 2M shows opposite trends between 2018-2019 and 2019-2020 for all the size ranges but the biggest ones. This is mostly due to the entry into service in 2019 of the largest worldwide vessel (e.g. the *MSC Gulsun* with a carrying capacity of 23.756 TEU). Of course, this strategic choice led to the consequent decrease of the *new panamax* and the *ULCV* in the 2020.

**Figure 33: changes in the fleet composition (2018-2019) and (2019-2020) – Ocean Alliance**



Conversely, The Ocean Alliance shows a different trend over the three years. Indeed, ships sized over 3,000 TEUs and up to 10,000 TEUs had a negative trend, while ships over 15000 TEUs constantly increased. This reflect the choice of some of the alliance participants (OOCL and CMA CGM) investing on Megamax vessels.

Figure 34: changes in the fleet composition (2018-2019) and (2019-2020) – The Alliance



Changes in the fleet composition of The Alliance led to an increase in number of *megamax* vessels, given by the new joined members ONE in 2019 and the HMM in the 2020.

Moreover, the net figure reports the breakdown of the alliances’ fleet between 2018 and 2020.

Figure 35: fleet breakdowns of ocean alliances

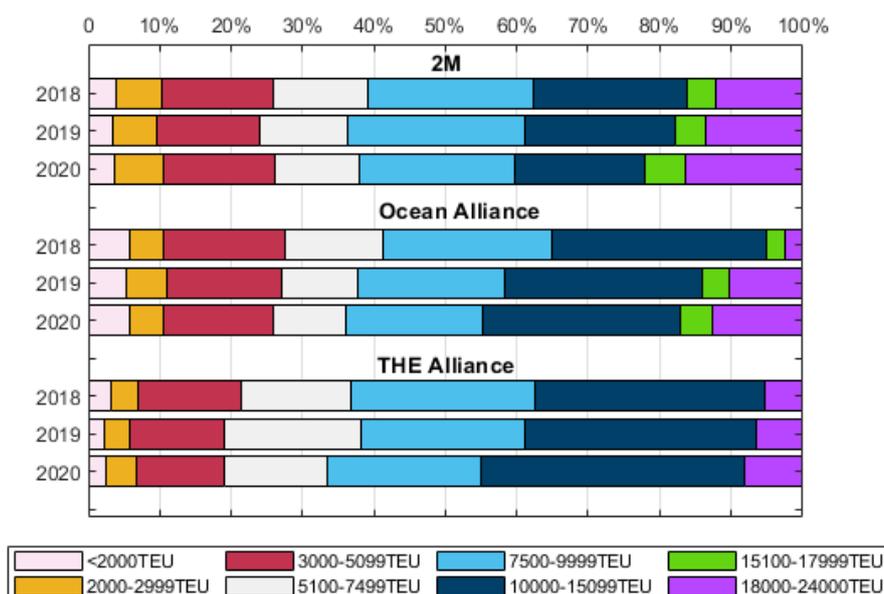


Figure 40 shows that the fleet of the Ocean Alliance and The Alliance is mostly composed by vessels sized between 10,000 and 15,000 TEUs, as this type of vessels still ensures more flexibility (e.g. in terms of maximum draught allowed, turnaround time and port congestion). Differently, despite the great number of *new panamax* vessels, the majority of the 2M fleet sized between 7,500 and 10,000 TEUs still accounting for 22% in 2020 (

Table 42).

The megamax vessels exhibit a positive trend, confirming the increasing run-up to mega ships over the last years. In particular, the 2M invested more than the others in the *megavessels* (+ 4.5%) from 2018 to 2020 (

Table 42).

**Table 42: Share of each size range on the total fleet capacity by ocean alliance**

		<2000 TEU	2000- 2999 TEU	3000- 5100 TEU	5100- 7499 TEU	7500- 9999 TEU	10000- 15099 TEU	15100- 17999 TEU	18000- 24000 TEU
2M	Pmf 2018	3.95	6.32	15.54	13.34	23.27	21.47	4.08	12.03
	Pmf 2019	3.46	6.11	14.37	12.29	24.91	20.97	4.36	13.53
	Pmf 2020	3.65	6.78	15.78	11.79	21.69	18.29	5.53	16.49
	Δpmf 20-19	0.19	0.67	1.41	-0.5	-3.22	-2.68	1.17	2.96
	Δpmf19-18	-0.49	-0.21	-1.17	-1.05	1.64	-0.5	0.28	1.5
	Δpmf 20-18	-0.3	0.46	0.24	-1.55	-1.58	-3.18	1.45	4.46
OCEAN	Pmf 2018	5.87	4.66	17.04	13.63	23.65	30.1	2.67	2.38
	Pmf 2019	5.41	5.64	16	10.59	20.74	27.66	3.79	10.17
	Pmf 2020	5.71	4.83	15.41	10.2	19.12	27.57	4.54	12.62
	Δpmf 20-19	0.3	-0.81	-0.59	-0.39	-1.62	-0.09	0.75	2.45
	Δpmf 19-18	-0.46	0.98	-1.04	-3.04	-2.91	-2.44	1.12	7.79
	Δpmf 20-18	-0.16	0.17	-1.63	-3.43	-4.53	-2.53	1.87	10.2
THE ALLIANCE	Pmf 2018	3.11	3.85	14.39	15.42	25.76	32.3	0	5.17
	Pmf 2019	2.32	3.49	13.17	19.11	23.01	32.35	0	6.55
	Pmf 2020	2.42	4.32	12.19	14.61	21.37	36.87	0	8.22
	Δpmf 20-19	0.1	0.83	-0.98	-4.5	-1.64	4.52	0	1.67
	Δpmf 19-18	-0.79	-0.36	-1.22	3.69	-2.75	0.05	0	1.38
	Δpmf 20-18	-0.69	0.47	-2.2	-0.81	-4.39	4.57	0	3.05

Beside the aforementioned top alliances, some of the main carriers engaged strategic cooperation to enhance their competitive advantages across certain trade lanes. This is the case of Mediterranean Shipping Company (MSC) and its 2M partner Maersk Line that in 2019 arranged a strategic cooperation with ZIM on Asia-US West Coast (USWC) and Asia-Mediterranean trades. The new operational agreement improved efficiency of services enabling a faster and better geographical coverage and extending direct port calls. Also, it is

based on a combination of vessel sharing and slot exchange and purchase accounting for an additional 3% of 2M total fleet capacity (Table 43).

**Table 43: fleet capacity of 2M+ZIM agreement**

<i>Carrier</i>	<b>2M fleet capacity [kTEU]</b>		
	<b>2018</b>	<b>2019</b>	<b>2020</b>
<b>Maersk</b>	3105	3457	3195
<b>MSC</b>	2994	3432	3141
<b>Zim</b>	0	242	202
<b>total</b>	<b>6099</b>	<b>7131</b>	<b>6336</b>

### 5.1.1 Container services supply: regional market and port-based analyses

The information collected in the database of container shipping services enables the evaluation of relevant statistical analyses to define the accessibility to liner services for both countries and ports in the proposed network.

Thus, to analyse the main regional markets, usually connected by the *Main Streets*, the macro-areas - previously defined in the previous Chapter - are grouped in 9 main trade regions, namely *Africa*, *Asia*, *Europe Mediterranean*, *Europe Northern Range*, *Middle East*, *North and Central America EC*, *North and Central America WC*, *South America EC*, *South America WC*. The following table shows the composition of main trade regions.

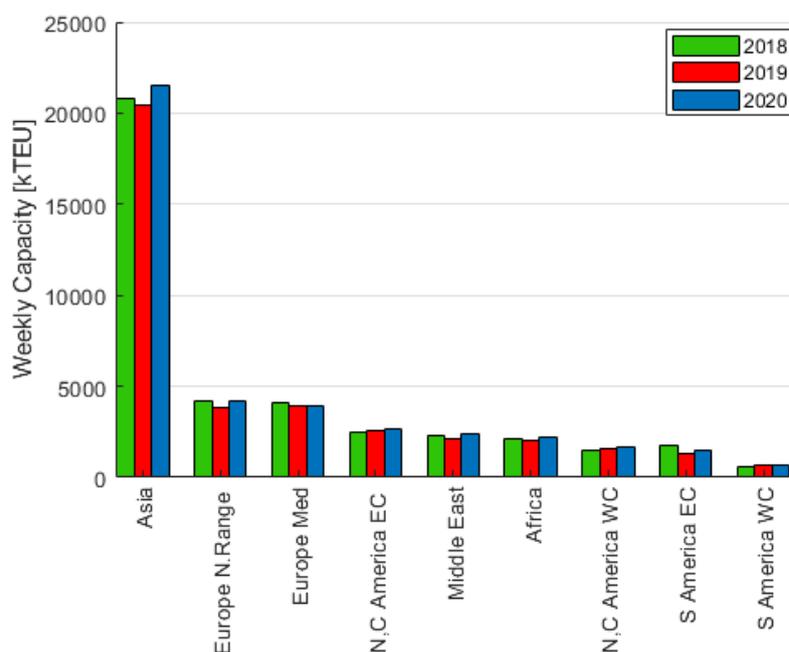
**Table 44: trade regions' compositions**

<b>Main Trade Region</b>	<b>Macro-areas included</b>
<b><i>Africa</i></b>	East Med Africa, West Med Africa, North West Africa, South West Africa, South East Africa, Indian Ocean
<b><i>Asia</i></b>	Far East, India Sri Lanka Bangladesh, Indochina, Korea & Japan, Russia East, Oceania, South East Asia
<b><i>Europe Mediterranean</i></b>	West Med Eu, South West Eu, East Med Adriatic, East Med Aegean Black Sea, EastMed, RussiaWest
<b><i>Europe Northern Range</i></b>	Greenland & Iceland, Scandinavia & Baltic Sea, Northern Range, UK
<b><i>Middle East</i></b>	Red Sea, Persian Gulf
<b><i>North and Central America EC</i></b>	North America East, Central America East, Caribbean
<b><i>North and Central America WC</i></b>	North America West, Central America West

<b>South America EC</b>	South America East
<b>South America WC</b>	South America West

Firstly, the total capacity supplied by liner container services is evaluated and the results are shown in Figure 36. Also, the yearly variation rates are reported in Table 45.

**Figure 36: weekly capacity provided by liner container services**



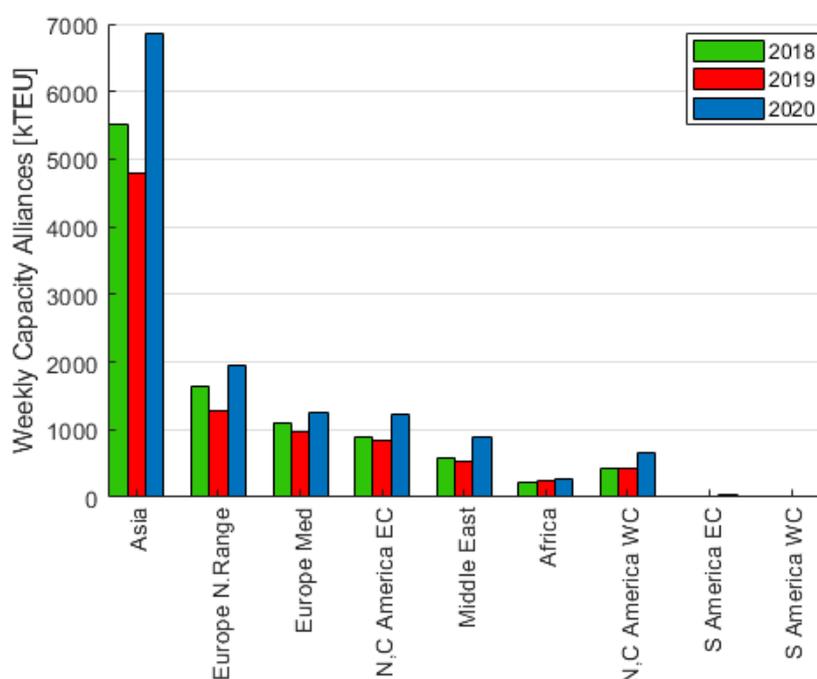
**Table 45: variations in weekly capacity supplied by liner container services by main trade region**

Main Trade Region	$\Delta$		
	2020-2019	2019-2018	2020-2018
<b>Africa</b>	+6%	+10%	+17%
<b>Asia</b>	+2%	-6%	-4%
<b>Europe Mediterranean</b>	-3%	+4%	+1%
<b>Europe Northern Range</b>	+6%	-11%	-6%
<b>Middle East</b>	+11%	-12%	-3%
<b>North and Central America EC</b>	+4%	+1%	+5%
<b>North and Central America WC</b>	-3%	+13%	+10%
<b>South America EC</b>	+12%	-19%	-9%
<b>South America WC</b>	-2%	+5%	+3%

Results exhibit a strong dominance of the Asian region in the global market due to the massive presence of domestic services operating in the region. Moreover, regional markets show quite stable trends over the three years, highlighting the fact that changes occurred between 2019 and 2020 are mainly due to evolving phenomena unrelated to COVID-19 pandemic crisis, which however worsen this negative trend.

Referring to the capacity supplied by ocean alliances, results follow the geographical pattern exhibited in Figure 36. The increase in alliances' operated capacity shows a positive trend in 2020, highlighting the fact that the additional capacity deployed on the main trade market has been provided by co-operated services ( Figure 37 ).

**Figure 37: weekly capacity provided by e shipping Alliances**



In detail, the Asian region exhibits the higher increase of weekly capacity, mostly due to the HMM joining The Alliance and the deployment of mega vessels on the Far East-related services. The additional carrying capacity amounts to (more than) 4 million of TEUs in the *Far East* and 1.5 million of TEUs in the South East Asia.

The Northern Range leads the growth of the shipping alliances' market in the North European area, with an increase of 500.000 TEUs of services' capacity per week. In the *Europe Mediterranean* market, the majority of alliances' services are deployed in the *Western Europe*, showing an increase of 721000 TEUs (47%) from 2019 to 2020. Whereas in the *EastMed – Aegean – BlackSea* area the increase amounts to more than 320.00 TEUs per week, providing a 43% of additional capacity. Finally, the *Sud West Eu e East Med* areas grows of 15% and 18% respectively in the 2020. By contrast, the *East Med Adriatic* is the

showing a decrease in the operated weekly capacity which falls down of 56% from 2019 to 2020, and of 22% from 2018 to 2020.

The growth of the operated weekly capacity in the *Middle East* region is mainly due to the increased number of services operated by shipping alliances in the Red Sea and in the Persian Gulf areas, with an additional weekly capacity respectively of 72% and 60%.

The African region shows quite stable values over the three years: between 2019 and 2020 the total increase of weekly capacity amounts to 8% in the West Med Africa and to 12% in the East Med Africa.

Moving on the analysis of the North and Central America East Coast, over the three years, the shipping alliances increased the number of services deployed. In detail, the major growth refers to the Caribbean area (+ 35%) and to the *Alaska- Canada Usa-East* (+34%). While in the North and Central America West Coast exhibits an even more positive trend, with an increase of 41% in the *Alaska- Canada Usa-West* area and of 85% in the Central America West Coast between 2018 and 2020.

This analysis has been carried out also taking into account the weekly capacity deployed by container liner services on each port. Table 46 shows the results for the twenty major ports.

**Table 46: weekly capacity by port**

Port	Weekly Capacity			Percentage Variation		
	2018	2019	2020	Δ 2020-2019	Δ 2019-2018	Δ 2020-2018
Shanghai	1783127	1680379	1949969	14%	-6%	9%
Singapore	1889843	1745161	1935028	10%	-8%	2%
Ningbo	1367718	1305030	1406272	7%	-5%	3%
Busan	1094484	1071366	1202401	11%	-2%	9%
Qingdao	848784,8	773742,2	1095928	29%	-10%	23%
Hong Kong	1094614	1008045	959289,8	-5%	-9%	-14%
Rotterdam	816800,1	815961,9	931300	12%	0%	12%
Yantian	804887,8	655262,4	744012,4	12%	-23%	-8%
Port Klang	703471,5	759037	724285,5	-5%	7%	3%
Xiamen	607487,1	641450,2	700644,1	8%	5%	13%
Jebel Ali	533111,8	518198,5	626042,5	17%	-3%	15%
Antwerp	588672,2	601796,2	625813,4	4%	2%	6%
Xingang	638591,6	757699,9	580242,2	-31%	16%	-10%
Tanjung Pelepas	443345,6	483365,5	520984,9	7%	8%	15%
Nansha	499108,5	406429,2	511311,5	21%	-23%	2%
Hamburg	378272,2	345038,8	412356,2	16%	-10%	8%
Laem Chabang	278442,8	294509	334380	12%	5%	17%
Dalian	377966,9	318008,1	321477,9	1%	-19%	-18%
New York	241831,2	264545,9	297980,1	11%	9%	19%
Bremerhaven	348334,3	283341,8	293108,3	3%	-23%	-19%

Results show that the major Asian container ports perform the best rank for weekly capacity supplied by liner services. The Hong Kong port exhibits a decrease in services supply, following the dropping trend of the port throughput mainly caused by the political crisis.

In the northern Range ports exhibit a positive trend boosted by mega vessels start operating Europe-Far East services of HHM and MSC. The port of to Hamburg shows the major increase mainly due to the Alliance shifting Atlantic services from Bremerhaven port.

However, in the 2020 some ports experienced a loss of the weekly capacity deployed due to earlier impacts of the pandemic crisis. The port of Xingang (Tianjin) has been hit by test procedures related to import cargo which slowed down the operations in port terminal leading to unavailable storage capacity for cargo (i.e. Port Congestion charge adopted to avoid the shipping of referees).

## 5.2 Shipping market concentration analysis

According to the literature findings outlined in Section 2.4.3, the analysis of the maritime network includes the assessment of container shipping market concentration at different geographical levels (i.e. macroareas and ports).

Usually, there are several indicators to measure market concentration. For the purpose of this study, for the number of shipping companies playing in the market, as well as their market share, the concentration level has been assessed through the evaluation of the Herfindahl-Hirschman Index.

### 5.2.1 Herfindahl-Hirschman index

The maritime freight transport, especially the container sector, is usually considered as an oligopolistic market. During the last decades, the development of hub-and-spoke networks led to the concentration of global container traffic in few major ports. The ambition to become hubs serving busiest trade lanes pushed the port sector to invest in infrastructure upgrading (additional terminal capacity, dredging channels, etc) and equipment to accommodate large container vessels.

As a result of the strong inter ports competition, some of the busiest trade regions (i.e. Far East) experienced a decrease of concentration levels in the container transport market (Wang et al. ,2004).

This paragraph proposes to enhance the assessment of the container market concentration, adapting the formulation of the well-known HHI to different perspectives

As introduced in Section 2.4 the HHI is formulated as

$$HHI = \sum_{f=1}^{n_f} ms_f^2 \quad (5.1)$$

wherein  $n_f$  is the number of firms in the market and  $ms_f$  the market share of the  $f$ -th firm.

A value of  $HHI=1$  implies a monopolistic market whilst  $HHI \rightarrow 0$  corresponds to perfect competition; in particular, usually  $HHI < 0.15$  indicates a competitive market,  $0.15 < HHI < 0.25$  indicates a moderately concentrated market,  $HHI > 0.25$  indicates a highly concentrated market.

In the framework of the container market, the HHI evaluation allows to analyse the liner services market both for shipping companies and ocean alliances (US Department of Justice, 2010).

For the purpose of the present research, the HHI is reformulated and applied to two different backgrounds, namely ports and trade regions connected to ports. Moreover, depending on the analysis target, different formulations of the HHI are proposed for each of the aforementioned backgrounds (i.e. assuming as *market operators* the container ports, shipping carriers and ocean alliances respectively).

Due to lack of information on port throughput, the weekly capacity deployed on each container service is evaluated and assumed as a proxy of market share.

Consequently, the HHI can be calculated as:

$$HHI_m = \sum_n \left( \frac{Capw_n}{\sum_n Capw_n} \right)^2 \quad (5.2)$$

where  $m$  is the selected market,  $n$  is the number of firms operating in the market and  $Capw_n$  is the weekly capacity supplied on services by the company  $n$ .

#### 5.2.1.1 Macro areas

The level of market concentration in each Macro area has been assessed based on the weekly capacity of active liner container services belonging to three different *operators*, namely the ports included in the macro area, the shipping carriers and the ocean alliances playing in the market.

Table 47 shows the HHI evaluated for the investigated period from 2018 to 2020), highlighting values based on the different HHI thresholds described above.

During the last three years, the levels of traffic concentration remained quite stable, although some fluctuations can be observed. The Far East region as well as the Indochina exhibit the less concentrated markets, mainly due to the inter port competition (which led to the expansion of the container port sector) and the growth of the domestic traffic.

The Northern Range and the Southern Europe show a moderately concentrated market, with the former exhibiting increasing value of HHI, generated by the less number of port called by liner services together with the increase of the vessels capacity calling North European ports (i.e. HMM and MSC mega vessels calling Rotterdam, Hamburg and Antwerp).

Highly concentrated markets are the ones referred to West and East Med Africa, which include in their basin few major ports dominating the market (i.e. Tangier in the West Med Africa and Damietta and Port Said in the East Med Africa).

Table 47: HHI based on weekly capacity of liner container services calling ports

Macro area	HHI - PORTS		
	2020	2019	2018
Greenlan & Iceland	0,291	0,299	0,230
North America West	0,157	0,158	0,161
North America East	0,096	0,088	0,087
Central America East	0,129	0,126	0,127
Central America West	0,202	0,193	0,190
South America West	0,126	0,141	0,130
South America East	0,077	0,076	0,076
West Med Africa	0,554	0,387	0,518
East Med Africa	0,368	0,395	0,346
North West Africa	0,091	0,075	0,082
South West Africa	0,147	0,163	0,166
South East Africa	0,206	0,178	0,197
Red Sea	0,337	0,294	0,292
West Med Eu	0,123	0,104	0,101
Scandinavia & Baltic Sea	0,057	0,056	0,043
South West Europe	0,216	0,215	0,233
East Med Adriatic	0,157	0,189	0,158
East Med Egeo Black Sea	0,119	0,113	0,124
Northern Range	0,207	0,205	0,184
Caraibi	0,145	0,140	0,137
East Med	0,248	0,262	0,261
India Sri Lanka Bangladesh	0,182	0,168	0,170
Indochina	0,061	0,060	0,105
Far East	0,078	0,075	0,077
Korea Japan	0,184	0,165	0,167
Russia East	0,285	0,327	0,293
Russia West	0,202	0,202	0,157
Oceania	0,083	0,096	0,069
Persian Gulf	0,173	0,161	0,146
UK	0,224	0,227	0,256
Indian Ocean	0,282	0,282	0,252
South East Asia	0,250	0,247	0,274

Moving to the analysis of the shipping carriers in each Macro area, the results are shown in Table 48.

Table 48: : HHI based on weekly capacity of shipping carriers

	HHI - CARRIERS		
Macro area	2020	2019	2018
Greenlan & Iceland	0,324	0,373	0,404
North America West	0,093	0,090	0,077
North America East	0,108	0,113	0,103
Central America East	0,128	0,127	0,161
Central America West	0,148	0,140	0,159
South America West	0,150	0,177	0,156
South America East	0,176	0,192	0,191
West Med Africa	0,238	0,204	0,200
East Med Africa	0,125	0,159	0,131
North West Africa	0,164	0,171	0,155
South West Africa	0,166	0,146	0,127
South East Africa	0,179	0,169	0,178
Red Sea	0,105	0,112	0,109
West Med Eu	0,161	0,171	0,161
Scandinavia & Baltic Sea	0,097	0,107	0,088
South West Europe	0,215	0,235	0,266
East Med Adriatic	0,188	0,174	0,159
East Med Egeo Black Sea	0,140	0,129	0,130
Northern Range	0,117	0,132	0,135
Caraibi	0,110	0,119	0,108
East Med	0,109	0,127	0,111
India Sri Lanka Bangladesh	0,066	0,060	0,068
Indochina	0,076	0,077	0,070
Far East	0,076	0,077	0,071
Korea Japan	0,051	0,049	0,038
Russia East	0,167	0,195	0,145
Russia West	0,883	0,896	0,564
Oceania	0,113	0,124	0,098
Persian Gulf	0,091	0,103	0,085
UK	0,108	0,114	0,112
Indian Ocean	0,263	0,274	0,191

The results show a very competitive market for all the considered Macro areas but the Russia West, mainly related to the presence of few container feeder services ranging from 50 to 500 TEUs per week. Even if looking at the carriers shares the market seems to be not much concentrated, if we look at the ocean alliances market share the results show a different

condition.

**Table 49: : HHI based on weekly capacity of ocean alliances**

Macro area	HHI - ALLIANCES		
	2020	2019	2018
Greenlan & Iceland			
North America West	0,362	0,364	0,394
North America East	0,336	0,339	0,346
Central America East	0,454	0,406	0,398
Central America West	0,682	0,684	0,603
South America West	-	-	-
South America East	1,000	-	-
West Med Africa	0,749	0,741	0,547
East Med Africa	0,421	0,595	0,418
North West Africa	-	-	-
South West Africa	1,000	1,000	1,000
South East Africa	1,000	1,000	1,000
Red Sea	0,335	0,356	0,423
West Med Eu	0,365	0,397	0,377
Scandinavia & Baltic Sea	0,623	0,535	0,525
South West Europe	1,000	1,000	1,000
East Med Adriatic	0,542	0,597	0,502
East Med Egeo Black Sea	0,374	0,374	0,406
Northern Range	0,339	0,358	0,366
Caraibi	0,877	0,829	0,846
East Med	0,365	0,341	0,370
India Sri Lanka Bangladesh	0,345	0,347	0,361
Indochina	-	1,000	1,000
Far East	0,345	0,379	0,384
Korea Japan	0,401	0,343	0,334
Russia East	-	-	-
Russia West	-	-	-
Oceania	-	-	-
Persian Gulf	0,334	0,380	0,393
UK	0,341	0,353	0,371
Indian Ocean	-	-	-

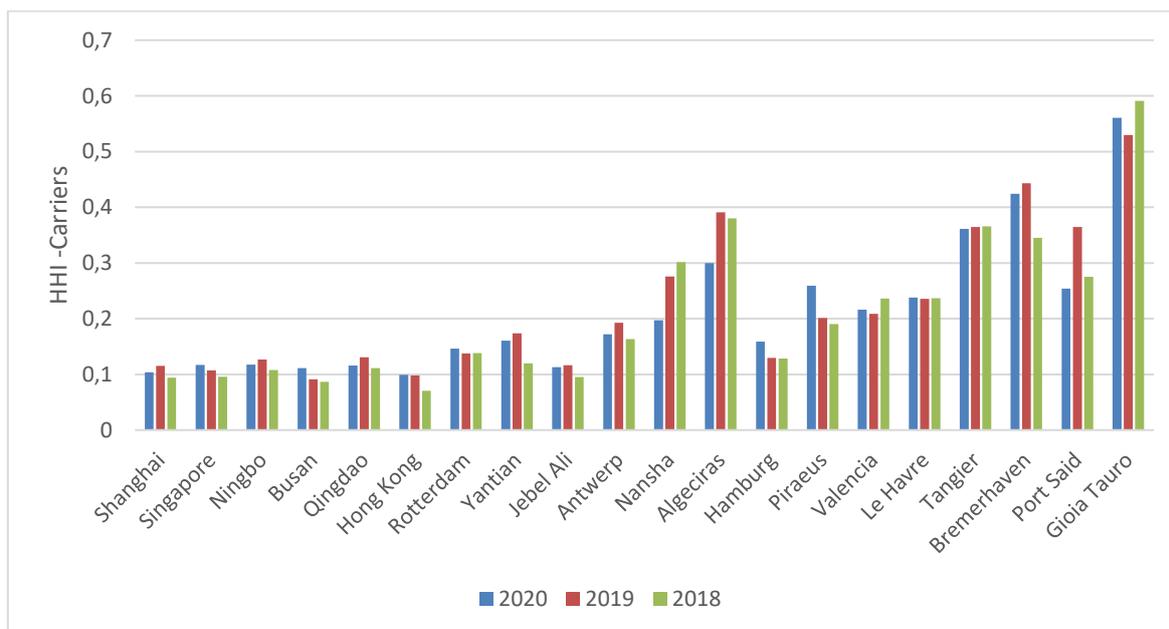
Table 49 reveal that the three ocean alliances (2M, The Alliance, Ocean Alliance) currently control the majority of the container market in almost all the Macro area, with HHI reaching the threshold of highly concentrated markets.

### 5.2.1.2 Ports

This section shows the results of the evaluation of the Herfindahl-Hirschman index based on the Port market. In detail, the HHI has been evaluated taking into account the share of weekly capacity operated by each carrier in the selected ports.

Figure 38 shows the results for some of the most representative ports in the network.

**Figure 38: HHI based on weekly capacity of shipping carriers in ports**



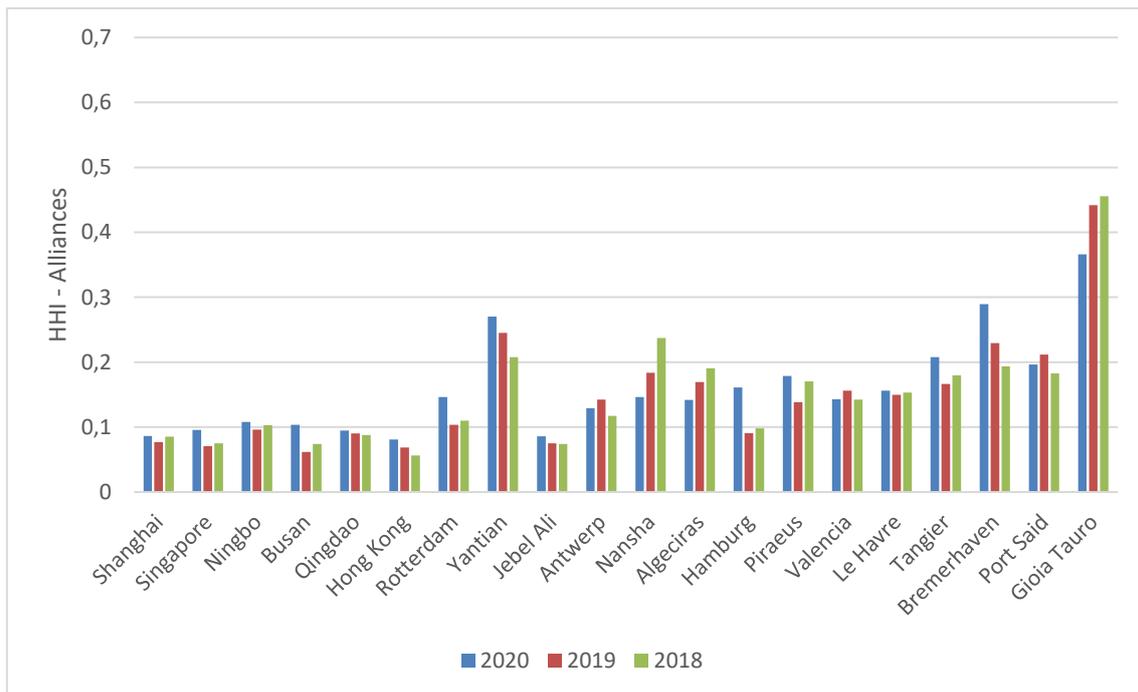
Asian ports show a very competitive market, with low values of the HHI. Nansha (Guangzhou) exhibits more concentrated levels: this large comprehensive port in the South China is called by liner services connecting the North America and the Europe-Mediterranean region, operated by some of the main shipping carriers (MSC, MAERSK, COSCO, CMA CGM, APL).

Also, the ports belonging to the Northern Range (Rotterdam, Antwerp, Hamburg) reveal an intermediate level of market concentration whereas the ports affected by the most concentrated markets are the Euro Med ports. Among these, Gioia Tauro port exhibits the higher value of HHI, highlighting the strong dominance of MSC.

Figure 39 shows the results of the HHI evaluation based on the market share of ocean alliances: as before, lower values of the index are exhibited by Asian ports, confirming the strong competitiveness of the maritime shipping market in this area.

Yantian port is the only one performing higher values of the HHI: currently it is called by liner services belonging to several carriers (35 shipping operators) but many of them (more than the 30%) belong to the three main alliances and own the majority of the weekly capacity calling the port.

**Figure 39: HHI based on weekly capacity of ocean alliances in ports**



A very monopolistic market is the one related to the Gioia Tauro port: this pure transshipment container hub is totally dominated by the MSC which deploy only a smallest part (13%) of the total weekly capacity on services co-operated with Maersk (2M alliance).

## 6 A GLS-based procedure to estimate/update freight flows between ports

The advances achieved in the estimation of freight flows between ports are illustrated in this chapter. In detail, this section proposes a procedure to estimate Ro-Ro/Ro-Pax freight flows between ports in the Western Mediterranean area, given the total inbound/outbound port throughput and the total weekly capacity of port-to-port services, based on the well-known GLS-based procedure that updates/estimates o-d flows from traffic measurements. The application of the procedure to both a laboratory experiment and to a real case study yields very effective results. The proposed methodology is part of a research carried out in the framework of this PhD project, and published in a top transport journal<sup>15</sup>.

Accordingly, the remainder of this chapter is organized as follows: after a briefly introduction Section 6.2 illustrates the Ro-Ro/Ro-Pax supply model which includes 495 ports belonging to the Western Mediterranean basin, with a focus on the Italian port system. Whereas Section 6.3 reports on the proposed GLS-based procedure to infer port-to-port Ro-Ro freight flows based on port throughputs and service capacities.

### 6.1 Introduction

A key pillar of the European transport policy is represented by the so-called *Motorways of the Sea* (MoS), defined by the Article 21 of the European Union (EU) Regulation 1315/2013 on the Trans-European Transport Network (TEN-T; European Parliament, 2013) as “... *the maritime dimension of the trans-European transport network [...] They shall consist of short-sea routes, ports, associated maritime infrastructure and equipment, and facilities as well as simplified administrative formalities enabling short-sea shipping or sea-river services to operate between at least two ports, including hinterland connections*”. MoS are thus aimed at concentrating freight flows on sea-based logistical routes, increasing cohesion within UE territory, reducing road congestion through modal shift, and promoting intermodal sea-road-rail connections. Many authors in the concerned scientific literature provided in-depth review of MoS-oriented policies and identified critical factors to get MoS projects effectively implemented. However, most of available Ro-Ro/Ro-Pax data come from aggregated official statistical sources, and little is known from often commercially sensitive shipping company-based or port-based information, being freight surveys usually cumbersome (Ben-Akiva et al., 2016). Therefore, this circumstance prevents informing properly policymaking of the sector. This study aims at contributing to this goal, by providing a transport supply model of maritime routes in the Mediterranean and a Generalised Least Squares (*GLS*) procedure to estimate Ro-Ro freight flows between ports based on port throughputs and on the capacity of

---

<sup>15</sup> MARZANO, Vittorio, et al. Ro-Ro/Ro-Pax maritime transport in Italy: A policy-oriented market analysis. Case Studies on Transport Policy, 2020.

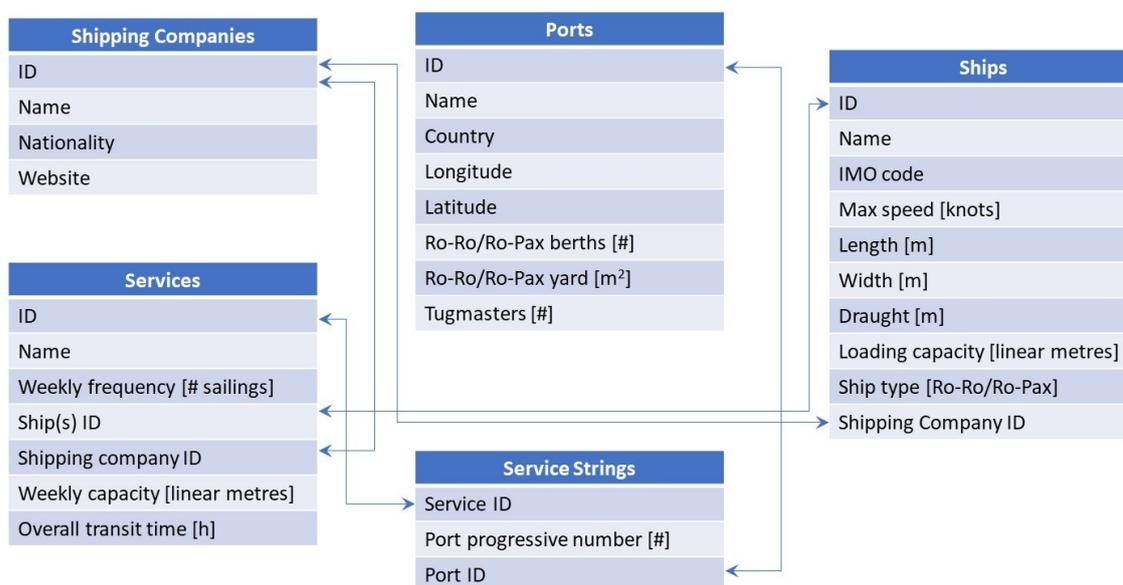
the services.

## 6.2 Development of a mathematical tool for Ro-Ro/Ro-Pax supply

The mathematical tool developed to support the procedure to estimate Ro-Ro freight flows in Italy encompasses a database and a topological model of all concerned maritime routes: both have been implemented adopting the first semester of year 2018 as temporal reference, and the European ports of Western Mediterranean (i.e. Italy, France, Spain) as geographical study area.

The structure of the implemented database is depicted in Figure 40. The database encompasses five tables with concerned relations and has been populated based on a fusion of various data sources, including shipping company websites, port authorities' websites, and other available resources online, such as the IMO database (International Maritime Organization (IMO), n.d.) and the Ro-Ro & Ferry Atlas - Harbours review (Ro-Ro & Ferry Atlas, 2017). Geographical location of ports has been obtained by merging Google Maps data with ports' locations within the World Port Index. Furthermore, port infrastructural data were complemented on the Italian side with official information from the Italian national transport account. Notably, maritime services to/from smaller islands (e.g. to/from islands in the Gulf of Naples) have been not considered.

Figure 40: Structure of the database of Ro-Ro/Ro-Pax services under analysis



The topological supply model representing concerned maritime routes has been obtained by adapting the supply model already developed by Arvis et al. (2016), Wilmsmeier and Hoffmann (2008), covering 495 ports in the entire Euro-Mediterranean basin. Notably, the database in Figure 40 allows obtaining the overall weekly capacity between each pair of ports in the dataset, by summing up the weekly capacity of all concerned services calling at both

ports of the pair; in turn, this capacity can be assigned to the supply model, yielding the overall available weekly capacity on each maritime link.

### 6.3 Estimation of the Ro-Ro/Ro-Pax freight flows between Italian ports given the port throughput: application to Italian ports

As previously mentioned, total throughput by port – disaggregated by inbound and outbound directions – is usually the only available information on Ro-Ro demand, whilst it would be of great interest to get information also on the flows between pair of ports. For this aim, this study proposes an adaptation of the well-known estimation/ updating o-d flows procedure based on traffic measurements.

From a general perspective, this issue falls within the more general problem of estimating system-wide origin-destination freight flows, based on mathematical models and decision support systems. The related literature is wide and covers more than thirty years of relevant contributions: Jong et al. (2016) provided, amongst others, comprehensive reviews of the state-of-the-art, mainly focussing on the estimation of production-consumption and origin-destination freight flows, and on mode choice modelling. In this respect, a recent contribution which attempts to ameliorate route choice modelling was discussed by Papola et al. (2018), who extended a previous work by Papola (2016). A promising opportunity that will be explored in this paper leverages the well-known problem of estimating/ updating o-d flows based on observed link flows and/or other traffic measurements (see for instance Marzano et al., 2018a for a recent review).

Specifically, the target of the procedure is the estimation of the unknown yearly Ro-Ro flow  $x_{od}$  expressed in tons between each pair of ports  $o$  and  $d$ . Clearly, all  $x_{od} \forall o, d$  should be consistent with the known total inbound  $d_d$  and outbound  $d_o$  throughputs of all ports  $d$  and  $o$  respectively. In addition, in synergy with the analysis presented in Section 3, all  $x_{od} \forall o, d$  should be consistent with observed weekly capacity  $c^{ml/w}_{od}$  in linear meters between each pair of ports  $o$  and  $d$ , in turn given by the sum of capacities of concerned services between  $o$  and  $d$ , belonging to the set  $S_{od}$ :

$$c_{od}^{ml/w} = \sum_{s \in S_{od}} c_s^{ml/w} \quad (6.1)$$

This consistency should be expressed by considering that the weekly capacity (6.1) should be converted into an equivalent yearly capacity  $c_{od}^{t/y}$  in tons/year (i.e. the measurement unit of  $x_{od}$ ) through the following formula:

$$c_{od}^{t/y} = c_{od}^{ml/w} \cdot \frac{w_{av}}{l_{av}} \cdot n_w \quad (6.2)$$

wherein  $w_{av}$  and  $l_{av}$  are respectively the average weight and average length of a Ro-Ro embarked unit and  $n_w$  the number of working weeks in a year. In the following it has been

assumed  $l_{av}=15$  m,  $w_{av}=20$  t and  $n_w=40$  weeks/year. As a result, each  $x_{od} \forall o, d$  should comply with the following constraint:

$$\alpha c_{od}^{t/y} \leq x_{od} \leq c_{od}^{t/y} \quad \text{with} \quad \alpha \in [0,1] \quad (6.3)$$

This allows formulating the problem of estimating onboard flows as:

$$\mathbf{x}_{od}^* = \min_{\mathbf{x}_{od} \in F_{od}} \left\{ \sum_p \left[ (d_{p\cdot} - \sum_i x_{pi})^2 + (d_{\cdot p} - \sum_i x_{ip})^2 \right] \right\} \quad (6.4)$$

wherein  $F_{od}$  is the feasibility set of port flows  $x_{od}$  defined by (6.3),  $p$  indexes ports, and  $d_{p\cdot}$  and  $d_{\cdot p}$  represent respectively the total outbound and inbound flow of port  $p$ . Notably, equation (6.4) does not impose respecting row and column marginals (i.e. generated/attracted port flows) as a hard constraint, because port throughput and service capacity come from different datasets and because of the subjectivity in the definition of  $\alpha$ . In this respect, a possibility is to choose  $\alpha=0$  straightforwardly; alternatively, one might impose  $\alpha$  such that to be consistent with a reasonable lower bound threshold on the loading factor.

In general, estimator (6.4) is known to yield poor results because of the unbalance between unknowns (# of o-d flows to estimate) and equations (# of available link counts and/or other traffic measurements), see for instance Marzano et al. (2009), Toledo et al. (2014) and Antoniou et al. (2016). However, its instance presented in this paper is characterized by two specific features. First, the structure of the network is peculiar, being each link representative of a maritime services connecting directly an o-d pair. This means that, if there is a single maritime service between a pair of ports, the o-d flow between those ports equals the flow on that link.

Of course, this circumstance is consistent with the nature of the Ro-Ro/Ro-Pax freight market, usually characterized by point-to-point flows without any transshipments. Second, information on the weekly capacity deployed between pair of ports allows setting a strongly reliable upper/lower bound on the likely flows between those ports: this circumstance, in particular, shrinks significantly the space of feasible solutions to the problem (6.4), thus ameliorating its performance.

To show the effectiveness of the proposed approach, a laboratory experiment has been carried out, following Marzano et al. (2009) and Antoniou et al. (2016), with the following rationale:

- a set of ports and a corresponding set of maritime services with a given weekly capacity has been randomly generated. In particular, the number of ports ranged between 10 and 90 and the number of services between 10 and a maximum depending upon the number of ports, however not larger than 1000;

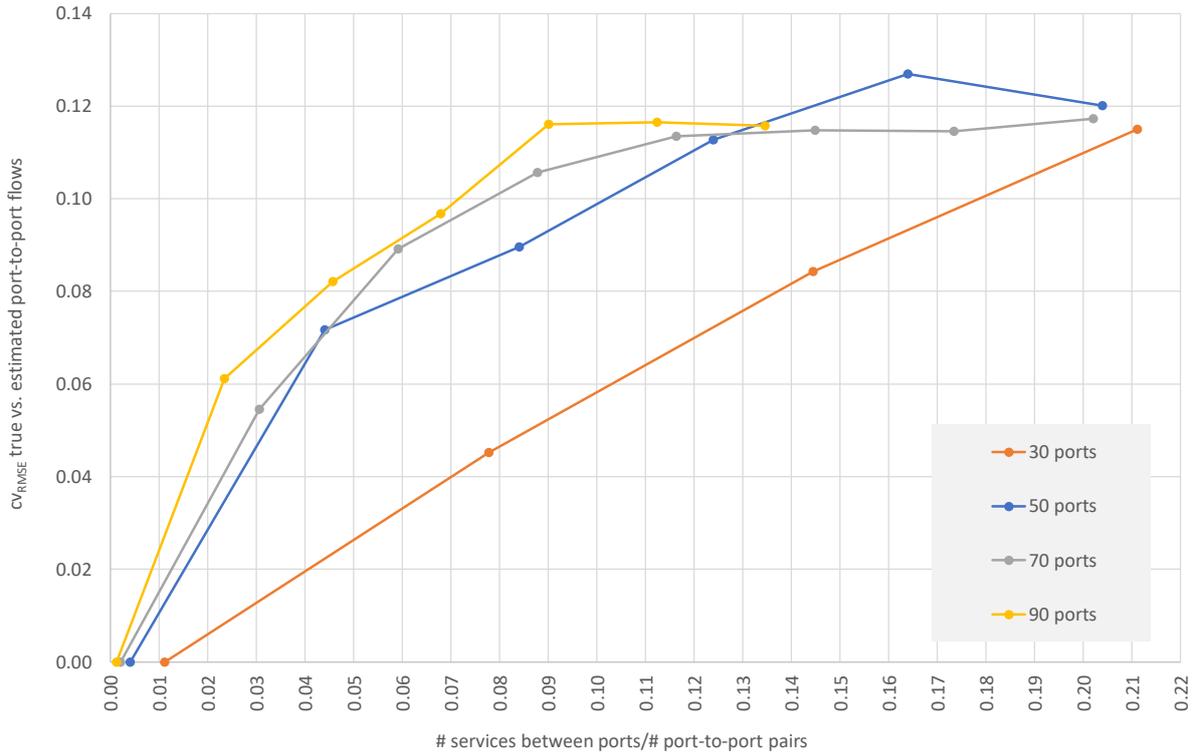
- each service was associated with a weekly capacity randomly generated between 3000 and 14000 linear meters, and with a true (unknown in the real world) load factor also randomly generated in between 0.50 and 1.00;
- for each set of ports and related services, true port-to-port flows and port inbound/outbound flows were calculated based on the above randomly generated ground truth;
- the estimator (6.4) was applied and estimated port-to-port flows  $\mathbf{x}_{od}^*$  were contrasted against true underlying flows by means of the  $cv_{RMSE}$ , i.e. the coefficient of variation of the root mean square error between true and estimated port-to-port flows.

Results are reported in the following Figure 41, that shows a  $cv_{RMSE}$  between true and estimated port-to-port flows not higher than around 0.12 at worse, and a remarkably stable error pattern. Notably, experiments are harmonised by drawing on the x-axis the ratio between the number of port-to port services and the number of o-d pairs of ports. Overall, the proposed approach is shown to be effective, and thus it is worth applying to the real case study discussed in the paper.

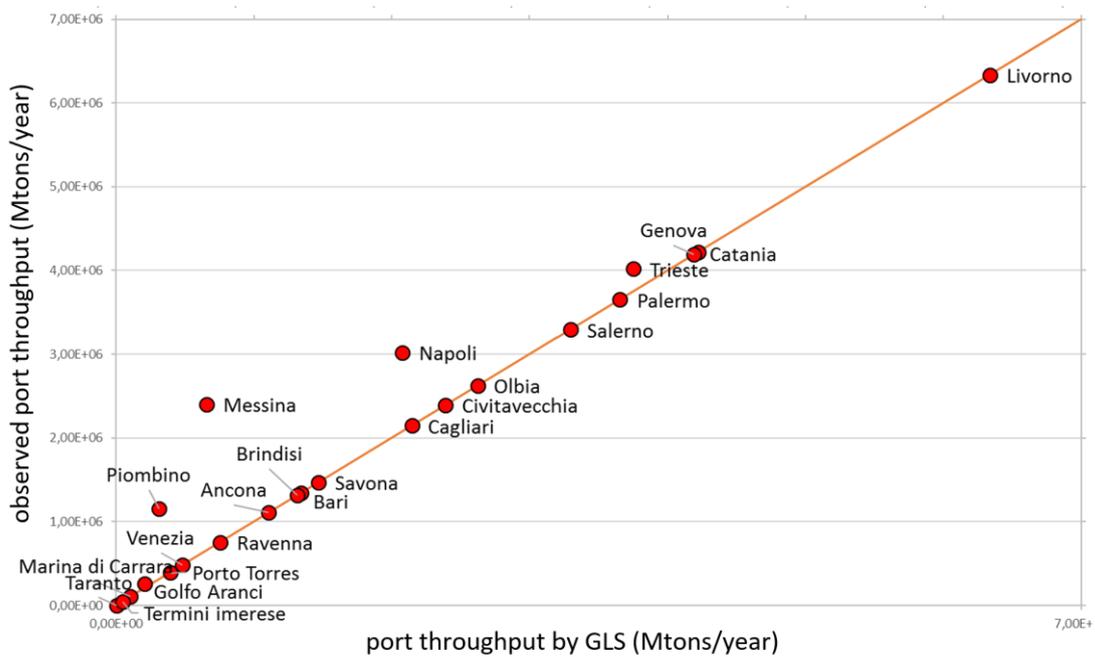
For this aim, the problem (6.4) has been solved with a very simple code in Matlab, assuming all Italian ports as study area, running four different scenarios resulting from the combination of two different initial points, respectively  $x_{od}=c_{od}^{t/y} \forall o, d$  and  $x_{od}$  randomly selected between  $\alpha c_{od}^{t/y}$  and  $c_{od}^{t/y}$ , and two values for  $\alpha$ , respectively equal to 0 and 0.70. The best performance has been observed for  $\alpha=0$  and  $x_{od}$  randomly selected between 0 and  $c_{od}^{t/y}$ , yielding an objective function value at optimum of  $1.54 \cdot 10^7$ .

Figure 42 contrasts true values and estimated values of attracted Ro-Ro freight flow by port, expressed in tons/year, showing that a perfect match is obtained for almost all ports: notably, the ones not correctly reproduced are those (Messina, Napoli and Piombino) with a significant share of short-range ferry services, contributing to the port throughput statistics but not present in the supply database illustrated previously illustrated.

**Figure 41: cvRMSE between true and estimated port-to-port flows in the laboratory experiment to check performance of the proposed estimator (6.4)**



**Figure 42: Comparison between model estimates (x-axis) and true values (y-axis) of attracted Ro-Ro freight flow by port.**



## 7 Summary and conclusions

The modelling of ocean container transport is a very wide line of research and it involves several interdisciplinary subjects, ranging from economics to transport engineering and geography. This study has proposed an innovative approach to modelling properly worldwide container transport networks. In this respect, the most relevant current issues concern:

- the simplifications adopted in the available models which prevent adequate realism for many applications;
- the missing representation of non-additive key performance variables in the rail freight supply model;
- the fact that the hypergraph-based approach is not extended to the worldwide maritime container network;
- the estimation of port-to-port container flows on a global scale.

In this respect, the present study addressed three main research questions:

- 1 *How to cope with the non-additivity of key performance variables in the rail freight supply model?*
- 2 *How to adapt the hypergraph-based approach to worldwide container service network modelling?*
- 3 *Is it possible to define a methodology to estimate port-to-port container flows on a worldwide scale?*

To answer these questions this doctoral thesis has proposed an innovative approach, which consists in the following steps:

- 1 implementation of a methodology to deal with non-additivity of key performance variables in the rail freight supply model;
- 2 definition of a hyperpath-based container maritime transport model;
- 3 analysis of liner container shipping services:
  - assessing effectiveness of the hyperpath-based centrality measures in container transport patterns;
  - investigating the shipping market structure through the evaluation of the HHI
- 4 definition of a practical approach to estimate port-to-port container flows.

The improvements achieved through this doctoral thesis are briefly outlined in the following.

First, in the light of narrowing down the research questions, an in-depth literature review related to the container transport modelling has been carried out. Several contributions have been exhaustively analyzed and grouped according to different subjects (i.e. national and international freight transport models, shortest paths in multimodal freight networks, maritime container service network modelling, network analysis and port connectivity measures).

The review process revealed that the models currently available resort of several simplifications preventing adequate realistic results. To properly evaluate the performances of a freight transport system, it is necessary to take into account that the costs associated to each path are not additive (e.g. dwell times of freight road vehicles, non-linear decreasing distance fees). In the current literature, this issue has been faced by some researchers, who tried to propose some contributions. However, none of the existing researches deal with the non-additivity of key performance variables in the rail freight supply mode.

Thus, this study has provided a methodology to deal with on-additivity of key performance variables in the rail freight supply model. Indeed, both the unitary cost of transport and the total capacity of a freight train depend upon best/worst rail freight link variables across a path, a circumstance that prevents standard shortest path algorithms from being applied. For this aim, a new approach has been proposed, very effective and optimal, and easily applicable also to large-scale networks.

To highlight the effectiveness of the proposed methodology, an application to a new scheme of incentives to railway undertakings has been presented.

From a methodological standpoint, the main challenge has been to quantify the infrastructural gap, and thus the proposed incentive, defined as the difference of the unit transport costs in the current and in the optimal scenario. For this aim, diverse network characteristics (length, slope, loading gauge, weight) and train characteristics (number of locomotives, weight of locomotives and unladen weight of railcars, length of locomotives and railcars, height of railcars, ...) have been considered. Mainly, the unitary costs under analysis are non-additive, and thus a simple and operational procedure to calculate shortest paths has been proposed. An application to the railway intermodal transport in Italy has been presented, to show the feasibility of the approach and illustrate the differences with respect to the current incentive scheme.

The proposed methodology has relevant policy implications, thanks also to the feasibility of its practical implementation to real-size contexts, as demonstrated in Section 4. Indeed, the rail network infrastructure manager can calculate easily the o-d based incentive for each year and/or semester, thus providing railway undertakings with equitable market conditions. Also, linking the quantification of the incentive with the infrastructural performance of the network allows setting integrated planning strategies, capable to handle effectively also the transition period towards the completion of infrastructural upgrading. The main research prospect deals with the possibility of including demand-side considerations into the proposed incentive scheme. An optimization problem could be set, that is looking for an o-d based incentive capable to maximize the modal shift from alternative modes to rail: this approach will be

explored in a subsequent research step. A further research prospect, which is worth mentioning, concerns the analysis of the impacts of the proposed incentive scheme on the costs and on the operations of railway undertakings.

Second, the review of the literature revealed that usually maritime container transport models are represented according to two common approaches: the *L-graphs* (with links representing legs of each service) or the *P-graphs* (with links representing direct port of loading-port of unloading connections enabled by each service).

However, to represent routing strategies in global networks, a hypergraph-based approach could be more appropriate. Despite this type of approach has already been proposed in the literature, the application to maritime container transport modelling has not been sufficiently investigated.

In this respect, this study has developed an innovative representation of the maritime container service network based on the hyperpath approach, proving the viability of a worldwide application.

In detail, HL-graphs and HP-graphs have been introduced by combining the definition of hyperpath with the well-known concepts of L-graphs and P-graphs, already extensively applied in the literature. Then, the calculation of centrality metrics in HL-graphs and HP-graphs has been discussed, and a new betweenness centrality measure has been presented. Finally, this approach has been applied to a worldwide maritime container service network, comparing the results of calculations of various centrality metrics on different types of graphs. In general, degree centrality metrics do not change appreciably over type of graph. On the contrary, betweenness centrality appears theoretically sounder on hypergraphs, particularly on HP-graphs, rather than on P-graphs. This is also evidenced by the practical worldwide example, wherein the centrality of important transshipment ports is better captured by the proposed betweenness centrality measure. In particular, the calculation of the betweenness centrality measure on the HL-graph triggers a lesser deterministic calculation with respect to the betweenness centrality measure on the L-graph: this allows a better interpretation, and thus a more realistic ranking, of terminals belonging to port clusters.

A noteworthy research prospect, currently under way for the inherent difficulty in a comprehensive data collection, concerns the investigation of the linkage between centrality metrics in HP-graphs and container terminal throughput – differentiated by import/export and transshipment – to analyse possible correlation patterns.

Furthermore, the implementation of the worldwide container services' database has enabled interesting analysis to understand the shipping market structure at different levels. In detail, an in-depth assessment of liner container services on a global scale has been carried out, yielding interesting insights on market positioning of shipping companies and ocean alliances across ports and trade regions.

Also, by evaluating the Herfindahl-Hirschman Index (HHI) based on the weekly service capacity deployed on liner services, the competition between carriers/alliances in each port as well as the competition between ports belonging to the same macro-area have been assessed.

On the whole, the analyses show that naval gigantism is still a significant trend which affects the structure of liner services as well as the cooperation between carriers. The applications have been carried out focusing on different market dimensions and accounted for two main perspectives, namely the shipping carriers and the ocean alliances.

The results revealed that the container shipping market exhibits a very different trend not only among different trade regions, but also among different ports belonging to the same macro area. In this respect, relevant policy implications can be highlighted.

Indeed, in the ports which exhibit a strong concentration, few players control most of the traffic and this strengthens their position with respect to ports.

This circumstance may trigger various policy implications, related to the market concentration of shipping companies and to how competition amongst port is managed.

A straightforward research step regards the analysis of the impact generated by the Covid- 19 pandemic on the container market concentration on a global scale.

Finally, this thesis has investigated the issue of the estimation of port-to-port container flows on a worldwide scale. Indeed, whilst supply data (e.g., schedules of liner container shipping services and related vessel characteristics) are easily available, demand data on container flows on routes and port-to-port container flows matrices are commercially sensitive and difficult to collect. As a result, there is lack of demand data to inform freight demand models to assess both so-called matrix production-consumption (p-c) and primarily underlying origin-destination (o-d) flows. In this respect, a Generalized Least Squares (GLS) approach has been developed; unfortunately, due to the unavailability of adequate input data on worldwide port container throughputs, this approach has been applied to within-Mediterranean Ro-Ro freight flows. In detail, the well-known GLS-based procedure that updates/estimates o-d flows from traffic measurements has been adapted to estimate Ro-Ro/Ro-Pax freight flows between ports, given the total inbound/outbound port throughput and the total weekly capacity of port-to-port services.

Applications to both a laboratory experiment and to a real case study yielded very effective results. Overall, relevant policy implications can be highlighted:

1. Ro-Ro/Ro-Pax freight flows represent a significant share of total port throughput in Italy and in many Mediterranean countries. At least for Italy, the Ro-Ro/Ro-Pax network ensures accessibility to strategic import/export markets in the Mediterranean area, which is known to represent a significant market share of the overall Italian international trade (OECD, 2020). Likely, the lower interest towards Ro-Ro planning is partly due to the circumstance that the operation of Ro-Ro services does not require costly maritime and port-related infrastructures, thus attracting less media attention. In this respect, it is advisable Ro-Ro/Ro-Pax traffic to play a major role in the agenda of freight policymakers.
2. the mathematical procedure proposed in the paper allows circumventing the sensitivity

issues on commercial data related to freight flows onboard Ro-Ro/Ro-Pax services. Indeed, knowledge of port-to-port Ro-Ro freight flows allows informing planning decisions and budget allocations for concerned infrastructural developments.

Three straightforward research steps are worth mentioning:

- the first deals with an in-depth validation of the proposed approach based on port-to-port observed data, currently unavailable;
- the second concerns the specification of a dynamic version of the estimator, i.e., introducing a temporal dimension, in the attempt to capture seasonal trends in Ro-Ro/Ro-Pax traffic;
- the further research step is about the collection of data on container port throughput to apply the GLS approach also to this case.

## 8 References

1. VIII Rapporto sull'Economia del Mare, UNIONCAMERE, 2019.
2. AGARWAL, Richa; ERGUN, Özlem. Mechanism design for a multicommodity flow game in service network alliances. *Operations Research Letters*, 2008a, 36.5: 520-524.
3. AGARWAL, Richa; ERGUN, Özlem. Ship scheduling and network design for cargo routing in liner shipping. *Transportation Science*, 2008b, 42.2: 175-196.
4. ALLAN, C., et al. Future potential for modal shift in the UK rail freight market. AECOM Ltd., commissioned by the UK Department of Transport (DfT). London, 2016.
5. ALVAREZ, Jose Fernando. Joint routing and deployment of a fleet of container vessels. *Maritime Economics & Logistics*, 2009, 11.2: 186-208.
6. AMATO, Flora, et al. Centrality in heterogeneous social networks for lurkers detection: An approach based on hypergraphs. *Concurrency and Computation: Practice and Experience*, 2018, 30.3: e4188.
7. ANTONIOU, Constantinos, et al. "Towards a generic benchmarking platform for origin–destination flows estimation/updating algorithms: Design, demonstration and validation." *Transportation Research Part C: Emerging Technologies* 66 (2016): 79-98.
8. ARVIS, Jean-François, et al. Trade costs in the developing world: 1995–2010. The World Bank, 2013.
9. ARVIS, J.F., Duval, Y., Shepherd, B., Utoktham, C., Raj, A., 2016. Trade costs in the developing world: 1996-2010. *World Trade Rev.* <https://doi.org/10.1017/S147474561500052X>
10. BELL, Michael GH, et al. A frequency-based maritime container assignment model. *Transportation Research Part B: Methodological*, 2011, 45.8: 1152-1161.
11. BELL, Michael GH, et al. A cost-based maritime container assignment model. *Transportation Research Part B: Methodological*, 2013, 58: 58-70.
12. Ben-Akiva ME, Toledo T, Santos J, Cox N, Zhao F, Lee YJ, Marzano V. Freight data collection using GPS and web-based surveys: Insights from US truck drivers' survey and perspectives for urban freight. *Case studies on transport policy*. 2016 Mar 1;4(1):38-44.

13. BEUTHE, Michel, et al. Freight transportation demand elasticities: a geographic multimodal transportation network analysis. *Transportation Research Part E: Logistics and Transportation Review*, 2001, 37.4: 253-266.
14. Blauwens, G., & De Baere, P. (2008). en Van de Voorde E., 2001. *Transport Economics, Third Edition*, De Boeck, Antwerpen.
15. BOLOGNA, Sergio. *Le multinazionali del mare: Letture sul sistema marittimo-portuale*. EGEA spa, 2011.
16. Boston Consulting Group, 2015. *The 2015 European Rail Performance Index. Exploring the link between performance and public cost*. Report by Duranton S., Audier A., Hazan J., Gauche V.
17. BOVENKERK, M. SMILE+: the new and improved Dutch national freight model system. *Proceedings of ETC 2005, Strasbourg, France 18-20 September 2005-Transport Policy and Operations-Freight and Logistics-Freight Modelling I*, 2005.
18. BROOKS, M., et al. Regulation in the liner shipping industry: pathways to a balance of interests. to be quoted as: *Technical Report: Regulation in the liner shipping industry: pathways to a balance of interests*. Online available at <https://www.uantwerpen.be/en/staff/christa-sys/my-website/research-interest/>, 2019, 3-3.
19. CASCETTA, Ennio, et al. Passenger and Freight Demand Models for the Italian Transportation System. Volume 2: Modelling Transport Systems. In: *World Transport Research. Proceedings of the 7th World Conference on Transport Research World Conference on Transport Research Society*. 1996.
20. CASCETTA, Ennio; MARZANO, Vittorio; PAPOLA, Andrea. Multi-regional input-output models for freight demand simulation at a national level. *Recent developments in transport modelling: Lessons for the freight sector*, 2008, 93-116.
21. CASCETTA, Ennio. *Transportation systems analysis: models and applications*. Springer Science & Business Media, 2009a.
22. CASCETTA, Ennio; MARZANO, Vittorio; PAPOLA, Andrea. Schedule-based passenger and freight mode choice models for ex-urban trips. In: *Schedule-Based Modeling of Transportation Networks*. Springer, Boston, MA, 2009b. p. 1-10.
23. CASCETTA, Ennio, et al. A multimodal elastic trade coefficients MRIO model for freight demand in Europe. In: *Freight Transport Modelling*. Emerald Group Publishing Limited, 2013. p. 45-68.
24. CHRISTIANSEN, Marielle, et al. Maritime transportation. *Handbooks in operations research and management science*, 2007, 14: 189-284.

25. CODD, Edgar F. A relational model of data for large shared data banks. In: *Software pioneers*. Springer, Berlin, Heidelberg, 2002. p. 263-294.
26. *Connettere l'Italia: Strategie per le infrastrutture di trasporto e logistica*, Nuova Struttura Tecnica di Missione, Ministero delle Infrastrutture e dei Trasporti, 2016.
27. CROZET, Yves; NASH, Chris; PRESTON, John. Beyond the quiet life of a natural monopoly: Regulatory challenges ahead for Europe's rail sector. Policy paper, CERRE, Brussels, December, 2012, 24.
28. CROZET, Yves. Rail freight development in Europe: how to deal with a doubly-imperfect competition?. *Transportation research procedia*, 2017, 25: 425-442.
29. DE DIOS ORTÚZAR, Juan; WILLUMSEN, Luis G. *Modelling transport*. John Wiley & sons, 2011.
30. de JONG, G.; GUNN, H.; WALKER, W.. National and international freight transport models: An overview and ideas for future development. *Transport Reviews*, 2004, 24.1: 103-124.
31. de JONG, G., VIERTH I., TAVASSZY L., and BENAKIVA Moshe. "Recent Developments in National and International Freight Transport Models Within Europe." *Transportation* 40, no.2 (June 13, 2012): 347-371
32. de JONG G, TAVASSZY L, Bates J, Grønland SE, Huber S, Kleven O, Lange P, Ottemöller O, Schmorak N. The issues in modelling freight transport at the national level. *Case Studies on Transport Policy*. 2016 Mar 1;4(1):13-21.
33. de LANGEN, Peter; NIDJAM, Michiel; VAN DER HORST, Martijn. New indicators to measure port performance. *Journal of Maritime Research*, 2007, 4.1: 23-36.
34. DEHOUSSE, Franklin; MARSICOLA, Benedetta. The EU's fourth railway package: a new stop in a long regulatory journey. *Egmont Paper 76*, April 2015. 2015.
35. Department for Transport (DfT), 2017a. *Guidance Grant funding to transport freight by rail and water, Mode Shift Revenue Support and Waterborne Freight Grant applications and background information*. Report, UK, July 2017.
36. DIJKSTRA, Edsger W., et al. A note on two problems in connexion with graphs. *Numerische mathematik*, 1959, 1.1: 269-271.
37. DUCRUET, César; NOTTEBOOM, Theo. The worldwide maritime network of container shipping: spatial structure and regional dynamics. *Global networks*, 2012, 12.3: 395-423.

38. European Court of Auditors, 2016. Rail freight transport in the EU: still not on the right track. Special report, doi:10.2865/53961.
39. European Parliament, 2013. Regulation (EU) No 1315/2013 of the European Parliament and of the Council of 11 December 2013 on Union guidelines for the development
40. European Union, 2011. White paper: Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system. European Commission, Brussels.
41. FINGER, Matthias; HOLTERMAN, Martin. Incentive-based Governance of the Swiss railway sector: a study of the performance incentives created by the financing of the railways. 2013.
42. FOWKES, A. S.; JOHNSON, D. H.; WHITEING, A. E. Modelling the effects of policies to reduce the environmental impact of freight traffic in Great Britain. In: ETC Annual Conference. 2010.
43. FREMONT, Antoine. Global maritime networks: The case of Maersk. *Journal of Transport Geography*, 2007, 15.6: 431-442.
44. GLEAVE, Steer Davies. Study on the Cost and Contribution of the Rail Sector. Final Report prepared for the European Commission, 2015.
45. GUGLIELMINETTI, P., et al. Single wagonload traffic in Europe: Challenges, prospects and policy options. *Ingegneria Ferroviaria*, 2015, 70.11: 927-948.
46. HARALAMBIDES, Hercules E. Gigantism in container shipping, ports and global logistics: a time-lapse into the future. 2019.
47. HEAVER, Trevor, et al. Do mergers and alliances influence European shipping and port competition? *Maritime Policy & Management*, 2000, 27.4: 363-373.
48. HEAVER, Trevor; MEERSMAN, Hilde; VAN DE VOORDE, Eddy. Co-operation and competition in international container transport: strategies for ports. *Maritime Policy & Management*, 2001, 28.3: 293-305.
49. HOFFMANN, Jan. Concentration in liner shipping: its causes and impacts for ports and shipping services in developing regions. 1998.
50. IBM Global Business Services, Kirchner C., 2011. Rail Liberalisation Index 2011. Market Opening: Rail Markets of the Member States of the European Union, Switzerland and Norway in comparison. Report (online available).

51. ISLAM, D. Barriers to and enablers for European rail freight transport for integrated door-to-door logistics service. Part 2: Enablers for multimodal rail freight transport. *Transport Problems*, 2014, 9.
52. ISLAM, Dewan Md Zahurul; ZUNDER, Thomas H. The necessity for a new quality standard for freight transport and logistics in Europe. *European Transport Research Review*, 2014, 6.4: 397-410.
53. ISLAM, Dewan Md Zahurul. Barriers to and enablers for European rail freight transport for integrated door-to-door logistics service. Part 1: Barriers to multimodal rail freight transport. *Transport problems*, 2014, 9.3: 43--56.
54. ISLAM, Dewan Md Zahurul; RICCI, Stefano; NELLDAL, Bo-Lennart. How to make modal shift from road to rail possible in the European transport market, as aspired to in the EU Transport White Paper 2011. *European transport research review*, 2016, 8.3: 18.
55. ISLAM, Dewan Md Zahurul; MORTIMER, Phil N. Longer, faster and heavier freight trains. *Benchmarking: An International Journal*, 2017.
56. Italian Ministry for Transport and Infrastructure, 2015. Proposal for policies to relaunch rail freight transport. ([www.mit.gov.it](http://www.mit.gov.it))
57. Italian Ministry for infrastructures and transport, 2016. Connecting Italy: Strategies for the transport and logistics infrastructures. ([www.mit.gov.it](http://www.mit.gov.it))
58. JANIC, Milan. Modelling the full costs of an intermodal and road freight transport network. *Transportation Research Part D: Transport and Environment*, 2007, 12.1: 33-44.
59. JANIC, Milan. An assessment of the performance of the European long intermodal freight trains (LIFTS). *Transportation Research Part A: Policy and Practice*, 2008, 42.10: 1326-1339.
60. JOURQUIN, Bart; BEUTHE, Michel. Transportation policy analysis with a geographic information system: the virtual network of freight transportation in Europe. *Transportation research part c: emerging technologies*, 1996, 4.6: 359-371.
61. JOURQUIN B, 2007. *Integration of discontinuity in road freight transport costs: a transportation network model*. Nectar Euroconference 2007, Porto, Portugal.
62. KEPAPTSOGLU, Konstantinos, et al. Free trade agreement effects in the Mediterranean region: An analytic approach based on SURE gravity model. *Transportation research record*, 2009, 2097.1: 88-96.

63. KHALED, Abdullah A., et al. Train design and routing optimization for evaluating criticality of freight railroad infrastructures. *Transportation Research Part B: Methodological*, 2015, 71: 71-84.
64. Lange, A., & Bier, T. (2019). Airline business models and their network structures. *Logist. Res.*, 12(1), 6.
65. LAROCHE, Florent, et al. Imperfect competition in a network industry: The case of the European rail freight market. *Transport Policy*, 2017, 58: 53-61.
66. LEE, Chung-Yee; SONG, Dong-Ping. Ocean container transport in global supply chains: Overview and research opportunities. *Transportation Research Part B: Methodological*, 2017, 95: 442-474.
67. LEI, Lei, et al. Collaborative vs. non-collaborative container-vessel scheduling. *Transportation Research Part E: Logistics and Transportation Review*, 2008, 44.3: 504-520.
68. MARZANO, Vittorio, PAPOLA, Andrea: Modeling freight demand at a national level: theoretical developments and application to Italian demand. Paper presented at the European transport conference 2004, Strasbourg 2004.
69. MARZANO, V., PAPOLA, A., & SIMONELLI, F. (2009). Limits and perspectives of effective O-D matrix correction using traffic counts. *Transportation Research Part C: Emerging Technologies*, 17(2), 120-132.
70. MARZANO V., PAPOLA A., SIMONELLI F., VITILLO R., TOCCHI D. Shortest paths in freight multimodal networks with non-additive impedances: a practical approach. Proceedings of the 96th Annual Meeting of the Transportation Research Board, Washington (D.C.), USA, January 2017.
71. MARZANO, Vittorio, et al. Incentives to freight railway undertakings compensating for infrastructural gaps: Methodology and practical application to Italy. *Transportation Research Part A: Policy and Practice*, 2018, 110: 177-188.
72. MEERSMAN, H., et al. L'histoire se répète? Current Competition Issues in Container Liner Shipping. 2015.
73. MEERSMAN, Hilde, et al. Competition and the container liner shipping industry. In: *Handbook of International Trade and Transportation*. Edward Elgar Publishing, 2018.
74. MERK, Olaf. The competitiveness of global port-cities: synthesis report. 2013.
75. MERK, Olaf. Container ship size and port relocation.
76. MIN, H., 2011. Building a Model-Based Decision Support System for Solving

Vehicle Routing and Driver Scheduling Problems under Hours of Service Regulations. Final Report, Bowling Green State University.

77. MVA and Kessel + Partner: MODEV—Modèle Marchandises, Note Méthodologique. MVA and Kessel + Partner, Paris/Freiburg (2006).
78. Network Rail, 2015. Network Rail leading Europe - Challenging assumptions about our relative performance. Report, UK, August 2015.
79. NGUYEN, Sang; PALLOTTINO, Stefano. Equilibrium traffic assignment for large scale transit networks. *European journal of operational research*, 1988, 37.2: 176-186.
80. NOTTEBOOM, Theo E. Container shipping and ports: an overview. *Review of network economics*, 2004, 3.2.
81. NOTTEBOOM, Theo E.; DE LANGEN, Peter W. Container port competition in Europe. In: *Handbook of ocean container transport logistics*. Springer, Cham, 2015. p. 75-95.
82. NUZZOLO, Agostino; RUSSO, Francesco. Stochastic assignment models for transit low frequency services: Some theoretical and operative aspects. In: *Advanced methods in transportation analysis*. Springer, Berlin, Heidelberg, 1996. p. 321-339.
83. Office of Rail and Road (ORR), 2016a. Annex A: Summary of control period 5 charges and contractual incentives framework. Report, UK, December 2016.
84. Office of Rail and Road (ORR), 2016b. Improving incentives on Network Rail and train operators: A consultation on changes to charges and contractual incentives. Report, UK, December 2016.
85. PANAYIDES, Photis M.; CULLINANE, Kevin. Competitive advantage in liner shipping: a review and research agenda. 2002.
86. PAPOLA A. A new random utility model with flexible correlation pattern and closed-form covariance expression: The CoRUM. *Transportation Research Part B: Methodological*. 2016 Dec 1;94:80-96.
87. PAPOLA A, TINESSA F, MARZANO V. Application of the Combination of Random Utility Models (CoRUM) to route choice. *Transportation Research Part B: Methodological*. 2018 May 1;111:304-26.
88. PEKIN, Ethem, et al. Integrated decision support tool for intermodal freight transport. In: *Nectar Cluster Meeting on Freight Transport and Intermodality*. 2008. p. 1-13.
89. REIS, Vasco. Analysis of mode choice variables in short-distance intermodal freight

- transport using an agent-based model. *Transportation research part A: Policy and practice*, 2014, 61: 100-120.
90. RIMMER, Peter J., et al. Global Flows, Local Hubs, Platforms and Corridors: Regional and Economic Integration in Northeast Asia. *Journal of International Logistics and Trade*, 2004, 1.2: 1-24.
  91. Ro-Ro & Ferry Atlas, 2017. Harbours review (<http://harboursreview.com/ro-ro-i-ferry-atlas-europe-2016/17.pdf>).
  92. RODRIGUE, Jean-Paul; NOTTEBOOM, Theo. Foreland-based regionalization: Integrating intermediate hubs with port hinterlands. *Research in Transportation Economics*, 2010, 27.1: 19-29.
  93. SHEFFI, Yosef. *Urban transportation networks*. Prentice-Hall, Englewood Cliffs, NJ, 1985.
  94. SONG, Dong-Wook. Port co-opetition in concept and practice. *Maritime Policy & Management*, 2003, 30.1: 29-44.
  95. SOUTHWORTH, Frank; PETERSON, Bruce E. Intermodal and international freight network modeling. *Transportation Research Part C: Emerging Technologies*, 2000, 8.1-6: 147-166.
  96. SPIESS, Heinz; FLORIAN, Michael. Optimal strategies: a new assignment model for transit networks. *Transportation research part b: methodological*, 1989, 23.2: 83-102.
  97. STEADIESEIFI, Maryam, et al. Multimodal freight transportation planning: A literature review. *European journal of operational research*, 2014, 233.1: 1-15.
  98. STOPFORD Martin, *Maritime Economics*, 2009.
  99. SYS, Christa, et al. In search of the link between ship size and operations. *Transportation Planning and Technology*, 2008, 31.4: 435-463.
  100. SYS, Christa. Is the container liner shipping industry an oligopoly?. *Transport policy*, 2009, 16.5: 259-270.
  101. SYS, Christa. *Inside the box: assessing the competitive conditions, the concentration and the market structure of the container liner shipping industry*. Ghent University, 2010.
  102. TAVASSZY, L. A., Smeenk, B., & Ruijgrok, C. J. (1998). A DSS for modelling logistic chains in freight transport policy analysis. *International Transactions in Operational Research*, 5(6), 447-459.

103. TAVASSZY, L. A. Freight modelling and policy analysis in The Netherlands. In: Presentation at the CTS-seminar on European and national freight demand models, Stockholm. 2011.
104. TAVASSZY, Lóránt, et al. A strategic network choice model for global container flows: specification, estimation and application. *Journal of Transport Geography*, 2011, 19.6: 1163-1172.
105. TAVASSZY, Lóránt A.; RUIJGROK, Kees; DAVYDENKO, Igor. Incorporating logistics in freight transport demand models: state-of-the-art and research opportunities. *Transport Reviews*, 2012, 32.2: 203-219.
106. TAYLOR, M.A.P., SEKHAR, S.V.C. & D'ESTE, G.M. Application of Accessibility Based Methods for Vulnerability Analysis of Strategic Road Networks. *Netw Spat Econ* 6, 267–291 (2006). <https://doi.org/10.1007/s11067-006-9284-9>.
107. LEVINSON, Marc *The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger*. Princeton, NJ: Princeton University Press, 2006. 376 pp.
108. TROCH, F. et al. Brain-Trains: Scenarios for Intermodal Rail Freight Transport and Hinterland Connections: From a SWOT Analysis of the Belgian Rail Practice to Scenario Development. In: *European Transport Conference 2015* Association for European Transport (AET). 2015.
109. TURNQUIST, M.A., 2006, September. Characteristics of effective freight models. In *Transportation Research Board Conference Proceedings (Vol. 40)*.
110. UIC, 2006. Conveyance of road vehicles on wagons – Technical organisation - Conditions for coding combined transport load units and combined transport lines. UIC 596-6 5th edition.
111. vAN HASSEL, E., et al. The Influence of Internalizing the External Cost on the Competitiveness of Sea Ports in the Same Container Loop. *Safety of Marine Transport: Marine Navigation and Safety of Sea Transportation*, 2015, 131.
112. vAN HASSEL, Edwin, et al. Impact of scale increase of container ships on the generalised chain cost. *Maritime Policy & Management*, 2016, 43.2: 192-208.
113. VAN HASSEL, Edwin, et al. The impact of the expanded Panama Canal on port range choice for cargo flows from the US to Europe. *Maritime Policy & Management*, 2020, 1-19.
114. VITILLO, Roberta. *Potenzialità e impatti dell'intermodalità nel bacino Euro-Mediterraneo: sviluppi teorici e prospettive operative*. PhD Thesis, 2011.

115. VLAANDEREN, Verkeerscentrum. Ontwikkeling van een multimodaal goederenvervoermodel voor vlaanderen [Development of a multimodal freight transport model for Flanders]. Tech. rep., K+ P transport consultants and TRITEL, 2006.
116. WANG, Hua; MENG, Qiang; ZHANG, Xiaoning. Game-theoretical models for competition analysis in a new emerging liner container shipping market. *Transportation Research Part B: Methodological*, 2014a, 70: 201-227.
117. WANG, Shuaian; MENG, Qiang. Liner shipping network design with deadlines. *Computers & Operations Research*, 2014, 41: 140-149.
118. WANG, Xinchang; MENG, Qiang. Discrete intermodal freight transportation network design with route choice behavior of intermodal operators. *Transportation Research Part B: Methodological*, 2017, 95: 76-104.
119. WARDMAN, Mark, et al. European wide meta-analysis of values of travel time. ITS, University of Leeds, Paper prepared for EIB, 2012.
120. WILMSMEIER, Gordon; HOFFMANN, Jan. Liner shipping connectivity and port infrastructure as determinants of freight rates in the Caribbean. *Maritime Economics & Logistics*, 2008, 10.1-2: 130-151.
121. WILLIAMS, I. Fundamentals of modelling goods vehicles in urban areas. WSP, Cambridge, 2011.
122. WOODBURN, Allan G. A logistical perspective on the potential for modal shift of freight from road to rail in Great Britain. *International Journal of Transport Management*, 2003, 1.4: 237-245.
123. WOODBURN, Allan. An investigation of container train service provision and load factors in Great Britain. *European Journal of Transport and Infrastructure Research*, 2011, 11.2.
124. XIAO, Quan. A method for measuring node importance in hypernetwork model. *Research Journal of Applied Sciences, Engineering and Technology*, 2013, 5.2: 568-573.
125. ZHOU, Wei-Hua; LEE, Chung-Yee. Pricing and competition in a transportation market with empty equipment repositioning. *Transportation Research Part B: Methodological*, 2009, 43.6: 677-691.
126. ZUNDER, Thomas Hagen, et al. How far has open access enabled the growth of cross

border pan European rail freight? A case study. *Research in Transportation Business & Management*, 2013, 6: 71-80.