Network Structure Changes of Container Transportation by Barge and Rail in the European Hinterland

Master Thesis

Author: *Berjan J. Waanders*
Student-number: *487107*
Programme: *MSc Supply Chain Management*
Date: *15 July 2018*

Coach: *Prof. Rob Zuidwijk*
Co-reader: *Dr. Otto Koppius*

Company coach: *Jeroen Bozuwa*
Position: *Senior Consultant*

Network Structure Changes of Container Transportation by Barge and Rail in the European Hinterland

Preface

Before you lies the master thesis ‘Network Structure Changes of Container Transportation by Barge and Rail in the European Hinterland’, a study on effects of service throughput time on the network structure for container transportation. This thesis is the final work for fulfilling the graduation requirements of the Master in Supply Chain Management at the Rotterdam School of Management, Erasmus University Rotterdam. I was engaged in researching and writing this thesis from December 2017 to June 2018.

My interest in the subject started with a friendly chat on the rowing course located near Rotterdam. By coincidence Dr. Otto Koppius and I crossed paths and he told me he would have an interesting subject for me to research. Two weeks later we were in a Skype call with Mitchell van Balen from Ecorys. During the call we formulated the research question on which this master thesis is build. The research was challenging as I had no prior knowledge on network theory, but thanks to the articles provided by Dr. Otto Koppius, I was well informed after the Christmas break and had a sound list of literature on which I could continue my research.

I would like to thank both of my supervisors, Prof. Rob Zuidwijk and Dr. Otto Koppius, for their excellent guidance and encouraging support over the last few months. Prof. Rob Zuidwijk has stirred my interest for the ports since the first semester and has continued to provide me with valuable insights during the research process. Meetings with Dr. Otto Koppius were often half an hour longer as planned since talking about the rowing crews we coached was always agenda point number one. I would like to thank him for the relaxing yet intellectually challenging atmosphere in which we could discuss my thesis. I would also like to thank Camill Harter, PhD candidate, for our meetings in which we could discuss our ideas regarding the dataset and your helpful feedback on my draft.

This research could not have been conducted without the support from Ecorys, the owner of the dataset. Specials thanks go to Mitchell van Balen for helping me formulating the research question and giving me insights in the dataset. Jeroen Bozuwa stepped in half of March 2018 and has provided me with valuable feedback on my drafts and significant insights based on his years of experience in the transport and infrastructure sector. I sincerely hope this research can be a valuable contribution for you and for Ecorys.

My parents deserve a particular note of thanks: you have unconditionally supported me during the last six years of university and encouraged me to pursuit my interests and gave room to develop myself. I could not be more grateful to them and my sister for their support.

I hope you enjoy your reading.

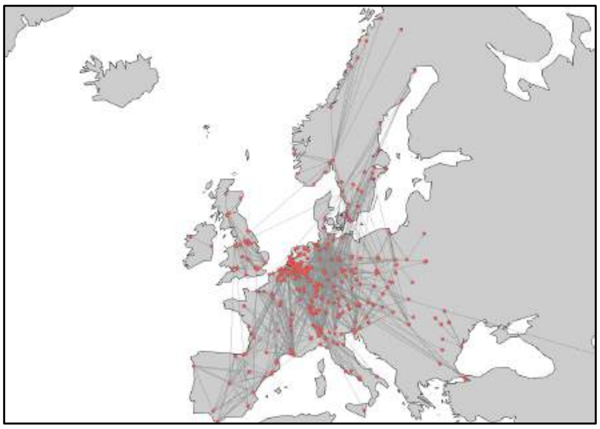
Berjan Waanders

Rotterdam, 15 July 2018

Abstract

Europe's hinterland for container transport is a unique mix of ports of all sorts of sizes and large economic areas served by a competitive mix of truck, rail, and barge operators. The ports establish a certain hierarchy among themselves by means of the number and type of ships or trains leaving or departing to certain destinations, the type of goods they handle, and the relative proximity of demand. This hierarchy can be classified as a network structure which can be measured with a vast number of complex network measures. The network structure can be changing over time because of, among others, competition between service operators, investments in infrastructure by (port) authorities, regulations and policies, and trade patterns.

This thesis researches one of the variables which can be of influence on the network structure: the throughput time of services on specific routes. The service throughput time is the average of the travel time and expected waiting time of all services operating between two cities. The travel time and the frequency between two cities are two main decision factors, next to price, for service operators and shippers to choose a certain route. As there is an absence of research on the influence of time on the network structure, this thesis aims to fill this gap by establishing a link between the change in service throughput time and the change in network structure of container transport in the European hinterland. As the characteristics of the two modalities which are used in the hinterland, rail and barge transport, are different, the network is also split up in order to observe differences between them.



The network for barge and rail transport in 2018.

For this purpose, a longitudinal dataset of three years (2016-2018) covering over 90 percent of all scheduled train and barge services in the European hinterland is used. Complex network measures used in this research are: hierarchy, assortativity, shortest paths, gamma index, centralities, rich-club index, Gini-index, clustering, nearest neighbour degree, and modularity. The network measures are extracted with several (statistical) programs: Gephi, Tulip, R (brainGraph and tnet packages), and Python (NetworkX package). In order to establish the link between the changes in service throughput time and network measures, IBM SPSS is used for calculating correlations and performing linear regression analyses.

Over the course of three years the complete hinterland network with both rail and barge connections shows a trend towards an increasing importance around high degree nodes. These high degree nodes act as hubs in a hub-and-spoke structured network for connecting communities consisting of smaller degree nodes. These communities are geographically dispersed over Europe and connected by certain corridors, linkable to the TEN-T corridors. The use of service throughput time as link weight has showed faster shortest paths exist than

the one crossing the least number of nodes. Correlations between the network measures and the service throughput time are strong and cohere with the line of thought for why measures are moving in a certain direction. None of the correlations or squared R values are however significant on a 95 percent confidence level for the complete network. So, while the relations between the service throughput time and the network measures are seemingly logical, it cannot be statistically established whether the service throughput time is a cause for the network structure changes.

Rail transport in the European hinterland has a relativity stable average service throughput time, but changes in the network structure do however take place. There seems to be a development in the use of intermediary cities with a significant number of connections; yet not as much as the large hubs have. This can be cities which are positioned on the corridors and act as entry points for hinterland destinations receiving and sending containers through a corridor. For the barge network the corridor structure was already in place and is dominated by primarily Antwerp and Rotterdam. The extra services included in the database mainly go from these two hubs to cities along the Rhine river in Germany. These relative long routes for the new services explain the significant changes in average service throughput time. Correlations show service throughput time is correlating with network structure changes around Antwerp, Rotterdam, and some other small cities. The offering of extra services here, not necessarily fast or frequent, on already existing routes increases the centrality.

It has proved to be useful to split the network for the modalities used in the European hinterland since the average service throughput time is showing different network structure developments for barge and rail. Eventually, service throughput time is however not able to statistically prove it is of influence in changing the network structure. The main reason for the lack of evidence is probably the limited timespan of the data, as the literature has showed a relationship between time and network structure changes is highly likely.

The influence of time of the network structure is of practical relevance for ports designing new (trans)port policies, initiating large infrastructure projects, and attracting service providers seeking access to certain hinterlands. A service provider could also benefit from the information how a transport network changes if a faster or more frequent service is offered. For rail operators the practical relevance is more specific on what throughput time is needed to establish a corridor, and for barge operators it is mainly how they can be more attractive than using a truck on a short distance range to cities with a limited number of connections.

Further research attention should be directed to transferring the observation made in the European hinterland to other geographical areas, extending the longitude of the observations in order to increase the statistical significance of service throughput time, and developing more specific measures for hinterland transport.

Table of Contents

List of figures:

List of tables:

List of equations:

List of appendices:

1 Introduction

1.1 Research Context

As hubs for around 80 percent of traded goods in volume worldwide, and over 70 percent of value (United Nations, 2015), deep-sea ports have received considerable attention from researchers over the past decades (see Pallis, Vitsounis & De Langen (2010) for an overview). Over the course of these decades, researchers have made more use of mathematical modelling and sophisticated analytics tools, resulting in the capability to analyse larger transportation networks (Woo, Pettit, Kwak & Beresford, 2011). This is essential with the growing complexity of the port community (Martin & Thomas, 2001) and their increasing role in facilitating global trade (Ng & Lui, 2014). In Europe alone, the main ports are handling close to 15 percent more tons of goods in 2015 compared to 2002 (Eurostat, 2017a). And although the transport sector experienced a large set-back in the aftermath of the financial crisis in 2008, the growth expectations remain positive.

In the eight biggest ports in Europe, over 55 percent of all containers measured in twenty-foot equivalent units are transported to the hinterland by road, rail or barge (although Bremerhaven is an exception with 60.8 percent sea-sea transhipment) (Notteboom, 2010). Rotterdam reaches 74.6 percent of hinterland transport, and Antwerp even the 81.1 percent. Moving goods by road is still the most common way of hinterland transport in the European Union with 75.8 percent of the modal split (Eurostat, 2015). Rail transport reached 17.9 percent in 2015, and barge transportation over inland waterways accounted for 6.3 percent of hinterland transport.

The development of hinterland facilities has been of particular interest for some researchers over the past years (Ng et al., 2014). From the period 1980-1994, there has been a trend of concentration and formation of multi-port gateways and hinterland development (Notteboom, 1997). The prediction is port concentration will eventually reach a limit, and future development in European container transport will predominantly be determined by changes in hinterland dynamics and government (trans)port policies. The prediction was right, the limit is reached and there is increasing emphasis on the role of effective hinterland distribution (Notteboom, 2010). The large and highly concentrated ports enjoy the benefit of economies of scale; larger container volumes simplify the establishment of a good network for the hinterland, which results in even more inflow. However, with the growing volume of containers, large ports become congested and unable to handle the constantly increasing amount of flow. New hinterland transportation links are developed because of an increasing number of containers to these ports. This can be new services from transport providers, but also new canals, railways, and terminals resulting from government policies and cities trying to establish themselves as an inland hub. An example is the focus on rail transport from the Port of Rotterdam. Since shipping by barge has a limited market reach, and roads are congested with polluting trucks, the port authority is assisting railway operators in establishing new services to for example Germany (Betuwelijn) and Genoa, Northern-Italy (Van Klink & Van den Berg, 1998).

The transportation possibilities offered by service operators result in a network for container transport throughout the hinterland of Europe. Just like the Atlantic shipping network (Ducruet, Rozenblat, & Zaidi, 2010), this network has the characteristics of a hub and spoke structure (De Langen, Lases Figueroa, Van Donselaar, & Bozuwa, 2017). As indicated by the research of Notteboom (2010), the European hinterland network is under change over time. Actions taken by operators and other actors in the network will have an impact on how the network evolves. An example is the deepening of the Scheldt river, which links the port of Antwerp to the North Sea. Dredging the river did result in a greater market share for the port of Antwerp (Veldman, Bückmann, & Saitua, 2005). This eventually had an impact on the number of containers transported to the hinterland. Other examples are the promotion of Venlo as a hub for inland transport (Raimbault, Jacobs, & Van Dongen, 2016), resulting in an increase in rail connections from Rotterdam to Venlo and from Venlo onwards, and an European-wide study proving the effect of developing new services in a hub is of positive influence on the hub's strength (intermodal connectivity) in the network (De Langen & Sharypova, 2013).

The three examples above are showing changes by actors do indeed have an effect on the network over time. They influence the services offered and the way containers flow to their destinations. Furthermore, a network having a hub and spoke structure such as the European container network can be described by certain measures. These are interesting because they help understand the strengths and weaknesses of a network and are able to track changes over time in a structured way. This is an extension to merely looking at the effects of a change such as the examples in the previous paragraph. These measures for network structures have primarily been studied at a higher level than the European hinterland. On a more global scale, Xu, Li, Shi, Zhang & Jiang (2015) looked at the traffic development, dominance, centrality, and vulnerability of 17 global shipping regions. They concluded the node attributes of a shipping region are not necessarily changing with what would be expected for the increase or decrease in traffic amount. This however differs per region. For the Southern African port system, the economic growth, increased political stability, and regional trade co-operation resulted in an increase of traffic (Fraser, Notteboom, & Ducruet, 2016). In turn, this resulted in a more central position of Southern African ports in the global shipping network. Though not all developments or measures taken by port authorities are resulting in a change in the port's or region's node attributes. Port specialisation and diversification have little influence on node properties such as betweenness centrality, degree centrality, and the clustering coefficient, but do have a big impact on link properties, such as traffic type and tonnage, due to the absence of a physical infrastructure on the sea (Ducruet, 2017).

On a European level, research on the network structure of container transport over time is limited. One of the only examples is the study by De Langen & Sharypova (2013) mentioned earlier, but their only measure is the intermodal connectivity of a node. Hence, it is of interest to define the European hinterland network for container transport with a set of measures and see its development over time. Additionally, developments over time are interesting if they can be explained. In other studies, researchers have looked at the effects of capacity (Laxe, Seoane, & Montes, 2012), distance (Guerrero, 2014), and volume (De Langen, Nijdam, & Van der Horst, 2007) between nodes and discovered relations between these link properties and the network structure. Two interesting attributes in the dataset which is used for this research, are the travel time and the travel frequency between nodes. Both of them are to date not used to explain changes in network structures for both global and hinterland networks.

Travel time and travel frequency are however, in addition to the geographical location of demand regions, of importance to the structure of transportation networks (Ducruet & Lugo, 2013). Transport time accounts for 10 percent, and frequency for 4 percent, in the decision making for hinterland container transport (the major decision factor is cost with 71 percent) (Meers, Macharis, Vermeiren, & Van Lier, 2017). Furthermore, cost are influenced by travel frequency and travel time through the exploitation of economies of scale (Van der Horst & De Langen, 2008), thus making the travel time and travel frequency of even greater importance in the decision-making process for hinterland container transport.

1.2 Problem Definition

It is of interest for researchers on transport networks to know how a network is shaped when travel time and travel frequency are varying, because both attributes are of influence on why operators choose for a certain hinterland route. So, it helps to explain how the decisions of operators to create a service between two nodes is of influence of the change of the network structure over time. Next to operators, other actors in the network are probably to some extent informed about why operators choose a certain route. These actors, like port authorities, might have the ability to alter their environment in order to attract more or faster services, and thus also have an influence on the shaping of the network.

Research on what the impact of travel time and travel frequency is on the network structure over time is however absent. In order to combine the two attributes in the dataset, the measure of service throughput time is used. It adds the travel time to the average waiting time a customer would have based on the number of departures per week. The average of all links between two nodes is the average throughput time of services. This measure is a link weight and can be used to establish a level of connectivity in the network, and the connectivity of a specific node. The idea of using service throughput time as a link weight in the network is obtained from Ducruet et al. (2010). One of the suggestions they make for further improvements in using complex network analysis tools for transport systems, is using traffic frequency as a link weight instead of the vessel capacity they have used. In this way, the influence of a change in travel time or the travel frequency on a certain link can be measured. As the characteristics between barge and rail transport are different and have their influence on the modal choice preference (Meers et al., 2017), the analysis on these two modalities must be split up in order to sketch a more precise image of why network structures are changing over time.

1.3 Research Objective

With research on weighing edges in networks with variables such as capacity, distance, and volume, the absence of time as a link weight measure poses a new opportunity to explore how network structure changes in the European hinterland can be explained. This is of importance because travel time and travel frequency are decision factors for actors in the network, which will influence the shaping of the network. The master thesis aims to fill part of this research gap, by applying complex network analysis tools to longitudinal data of rail and barge services for containers in the European hinterland. Afterwards there is explored if differences in the network can be explained by a change in service throughput time and if the mode of transport is of influence on the service throughput time. The conclusions from this thesis can be useful for explaining why a hinterland transportation network is changing if travel time or travel frequency are varying.

So, the objective is:

To explain if network structure changes of container transport in the hinterland of Europe can be attributed to changes in service throughput time and the influence of the modality on the service throughput, by applying complex network measures to longitudinal data on rail and barge services.

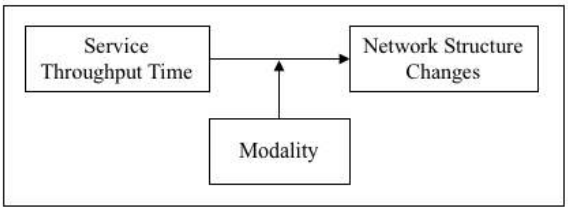


Figure 1: Conceptual model

1.4 Research Questions

The dataset contains snapshots from three years: 2016, 2017, and 2018. Service throughput time is used as a link weight on all three datasets in order to observe differences between them. Attention is paid to the changes in the network structure and to the attributes belonging to nodes. If a change is observed, a study is made on whether the influence of a change in service throughput time can be causing this. The possible influence of the transport modality is studied afterwards. Here the focus is on whether there is a difference in the impact of service throughput time across modalities. This results in the following three research questions:

1. *What are the differences in the network structures and node attributes for the three consecutive years of data?*
2. *How is the service throughput time influencing possible changes in the networks?*
3. *How does the influence of the service throughput time differ per mode of transport?*

1.5 Outline of the Thesis

This thesis continues with an exploration of theory on network structures in transportation. Next, the European network structure for container transport and the way hinterland transportation is performed in Europe are described. The influence of travel time, travel frequency, and the transport modality are explored in the last part of the literature review. The research methodology describes the structure of the dataset and the operationalisation of the variables from the conceptual model. A section with results is structured in the order of the three research questions: the network changes are described; the influence of service throughput time is measured; and the influence of service throughput time is split up per modality. The thesis finishes with the conclusions, a discussion of the results, and recommendations for further research.

2 Exploration of Theory

This chapter aims to give insight into the available literature on both network analysis and the hinterland network present in Europe. The available literature on the two independent variables, service throughput time and the differences per transport modality, are explored as well.

2.1 Background on Network Analysis in Transportation

Until the 1990s, graph theory was the primary way of studying transport networks (Ducruet & Lugo, 2013). This mathematical approach considers networks existing of nodes (i.e., cities) and links (also called edges, i.e., a waterway between two nodes). Edges can go one way, directed networks, or both ways, undirected networks. Nodes and links can also have weights placed to them such as capacity or time; this is called a weighted graph. The research focus was mainly on roads, railways, and canals, because these means of transport are running over fixed routes. This in contrary to maritime and air transport structures, who are almost not dependent on spatial factors (although maritime transport is somewhat constrained by coastlines). In order to establish a network for maritime or air transport, data of scheduled services or movements of ships or planes is required. This data was not available or extensive enough for a long time, and thus analysis on these two types of transport was lacking. In the late 1990s, more complex network measures were developed, introducing the concepts of small-world and scale-free networks. These two types of networks are often seen in transportation networks and give an intuitive abstraction of the real world.

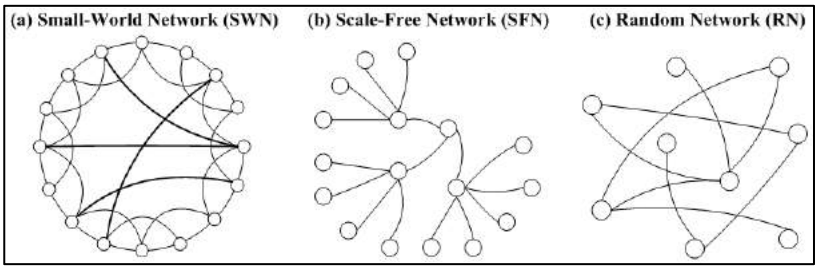


Figure 2: Complex networks. From Huang, Sun, & Lin (2005).

Networks in which nodes are highly clustered with a short width, and therefore have a high density of edges, are defined by Watts and Strogatz (1998) as small-world networks (figure 2a). There is a high chance two randomly selected nodes are direct neighbours of each other or require only few other nodes to reach many other nodes in the network. In a scale-free network (figure 2b), many nodes have few connections, and few nodes have many connections (Barabasi & Albert, 1999). This results in a disassortative network in which nodes with few connections are more likely to connect with highly connected nodes.

Although complex network measures are nowadays often used, not all transportation networks are suitable for this type of analysis because they do not hold the properties of a small-world or scale-free network. Therefore, researchers keep using the set of measures the original graph theory holds. These measures tell more about the local properties of a node, and less about the complete network. The definition used in this paragraph are taken from Demšar, Špatenková, & Virrantaus (2008). Centrality measures are values describing how important a node is in the network. The most common used types of centrality are closeness, betweenness, and degree. The closeness centrality indicates what the shortest distance of a specific node is to all other nodes in the network. Betweenness centrality is number of shortest paths connecting every pair of nodes passing through a certain node. The degree of a node is the total number of neighbours. The importance of a specific node it its direct network environment can be measured by the clustering coefficient. It can be calculated by summing all the edges between a certain node and all its neighbour nodes divided by all possible edges in the same neighbourhood of the specific node. These and other measures will be discussed in more detail in the methodology chapter.

2.2 Europe's Hinterland

In order to understand how or why a network is developing itself, knowledge about the network studied is required. The container transportation market by rail is for instance different in the US than Europe because railroads are owned by governments in the latter.

2.2.1 The European container network

Europe's container transportation network is a unique mix of ports of different sizes and types. The large economic hinterland, plus the relative proximity of competitors (Wang & Cullinane, 2006) results in a certain competition and hierarchy among the ports. European container port competition is, among others, studied by Marcadon (1999), Notteboom (2007), Veldman & Bückmann (2003), and Veldman et al. (2005). In one of the only longitudinal studies performed on the European container network, Notteboom (1997) concluded the containerisation of transport is not leading to further concentration of ports. Instead, he expected traffic flows to decentralise under the influence of competition between large consortia, the development of hinterland links, and policies from (port) authorities and governments. In his follow-up research in 2010, it became evident this was indeed happening. Large market players are being more orientated to the complete hinterland network and try to exploit their economies of scale on their own or by cooperating with other players. However, this was not the case for barge operators, they rather stay independent and not cooperate with other barge services to exploit the economies of scale (Van der Horst & De Langen, 2008).

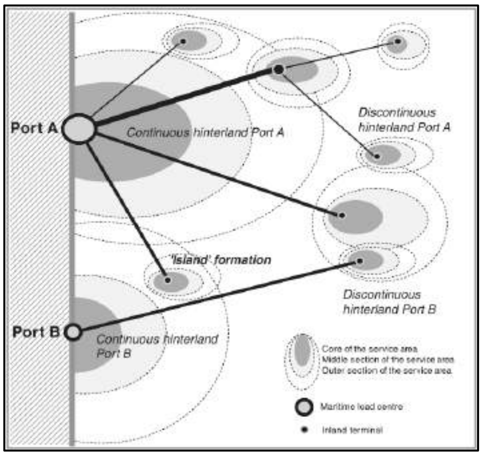


Figure 3: Overlapping hinterland markets through corridor-based container flows. From Notteboom & Rodrigue (2005).

So only economies of scale influencing cost factors are not sufficient to explain the flows of containerised traffic through Europe. Service quality effects, connectivity to hinterland markets, and synchromodal bundling effects result in forces pushing container flows to certain gateway ports and not to necessarily to the closest gateway (Notteboom, 2010). This type of port development in Europe, called port regionalisation (centralisation in network terms), can be drawn schematically as shown in figure 3. The activities normally taking place in the port are decentralised, but the import or export of goods is running through the deep-sea ports with a high hinterland connectivity. The formation of corridors is creating discontinuous hinterlands (as in figure 3). Typically, the high-volume corridors offer a better balance in lead time, price, and distance compared to the established modes of inland transport (Notteboom, 2008). The size of the resulting hinterland is depending on the frequency of the services over the corridor and the competitiveness of the operators. A drawback of this development is hinterland areas are depending on the competitive offering of services from operators. An unpredictable and highly competitive marketplace for intermodal transport is thus not leading to stable hinterland areas.

2.2.2 The influence of a modality

There seems to be a logical relationship between physical distance and the travel frequency: when the first is increasing, the latter is decreasing. This is also the case for intermodal transport in Europe where differences are observed if the analysis is split up for barge and rail services (figure 4) (De Langen et al., 2017). The travel frequency per week for transport by barge drops from an average of four per connection with a distance of less than 200 km, to two per connection with a distance between 400-600 km. No barge connections exist for distances over 600 km. Travel frequency per week by rail is also decreasing with distance, but only from an average of six per connection with a distance of less than 200 km, to four per connection with distances over 1200 km.

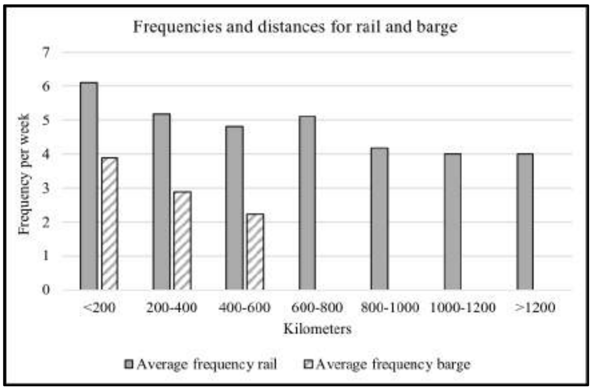


Figure 4: Frequencies and distances for rail and barge services. From: De Langen et al. (2017).

A prerequisite for effective waterway transport by barge is, besides the demand, the availability of infrastructure. This is of course also the case in rail transport, but a railway is often easier to construct than a canal. The length of waterways used for barge transport in Europe is about 52,000 km (Konings, 2009). Half of this is network can be found in France (14,900 km), Germany (7,500 km), The Netherlands (5,000 km), and Belgium (1,570 km). This explains why this modality is of such popularity in these regions.

De Langen et al. (2017) agree on several concluding points after their analysis. First, barge and rail services are complementing each other (i.e. the number of pairs with the same origin and destination is limited). Barge services are mainly used for short distances and rail services for longer distances. They are however both competing with road haulage for short (barge) and long (rail) distances (Meers et al., 2017). Second, if there is more competition between service providers at a certain port, this may lower the distance of the shortest rail service. Third, the transportation distances are often smaller than they assumed to be economically viable, especially between ports and over mountain crossings. Fourth, inland-to-inland services by barge rarely exist, the focus is more on deep-sea port to hinterland connections. The large train operators primarily specialise in port to hinterland connections, but inland-to-inland services are offered as well (although no service provider is specialised in these connections by rail).

2.2.3 *Changes in network structure*

In order to find evidence for changes in the hinterland, measures and statistical methods to suit this case are required. Examples are derived from studies looking at networks on a more global level because research on network structure changes in hinterlands is lacking (see table 1 for the complete overview). In a study on the maritime network structure of the Atlantic, the maritime degree (i.e., the number of direct connections a node has) is plotted against the cumulative number of nodes (Ducruet et al., 2010). By applying the power-law rule on the slope of the line (slope > 1), the Atlantic maritime network proves to show scale-free properties. The slope was however changing and getting closer to 1 between the two snapshots of 1996 and 2006; indicating the network was becoming less centralised. In the same study the density of the network was measured by comparing the observed connectivity with the optimal connectivity. This interconnectivity measure can also be applied on one node, indicating

Author	Area	Network	Longitude	Measure
Fleming et al. (1994)	US	Air, maritime	1 year	Centrality, intermediacy
Guimera et al. (2005)	World	Air	1 year	Centrality, hierarchy, communities, shortest path
Choi et al. (2006)	World	Internet, air	1 year	Centrality
Sales-Pardo et al. (2007)	World	Air	1 year	Hierarchy, communities, clustering
Ducruet et al. (2009)	Europe	Road, rail, river, short-sea	1 year	Gini-index, diversity
Ducruet et al. (2010)	Atlantic	Maritime	1996 & 2008	Centrality, hierarchy, clustering, gamma index
Lam et al. (2011)	World	Maritime	1995-2006	Centrality
Parshani et al. (2011)	World	Air, maritime	1 year	Centrality, clustering
Laxe et al. (2012)	World	Maritime	2008-2010	Centrality, vulnerability
Ducruet et al. (2012)	World	Maritime	1 year	Degree, communities
Fraser et al. (2016)	South-Africa	Maritime	1996, 2006, 2011	Centrality, eccentricity
Wang et al. (2016)	World	Maritime	1 year	Centrality
Ciliberto et al. (2017)	US	Air	1990-2015	Centrality
Calatayud et al. (2017)	Americas	Maritime	1 year	Centrality, gamma index, clustering, diameter, shortest path
Ducruet (2017)	World	Maritime	1977-2008	Centrality, gamma index clustering, assortativity, rich-club
Liu et al. (2018)	World	Maritime	1 year	Centrality, hierarchy

Table 1: Overview of network measures used in other publications.

whether it is well connected to its direct neighbours. A third method Ducruet et al. (2010) are using for gathering evidence of the changes in the network, is by applying the Gini-index (a measure of equality/inequality) on the distribution of traffic on nodes and edges. Through an observation of a decrease in the Gini-index, they conclude there is decrease in hierarchy and therefore greater complexity in the network.

The maritime degree of a node says something about the importance of a node, though not anything about its role in the network. Guimera, Mossa, Turtschi & Amaral (2005) use betweenness centrality, the number of shortest paths connecting any two cities through a specific node, for the air transportation network. Nodes with a small degree and large centrality are considered outliers since other complex networks do not behave like this according to Guimera et al. These outliers are creating a network looking like two communities only connected through a specific node, something which was first observed in air transportation. The combination of these two measures can thus be used to identify communities in a network and the importance of nodes in those situations. It is often referred to as the connectivity of a node (Xu et al., 2015). An increase of centrality on its own could mean the node is acting as an important place in the integration process of distant hubs (Seoane, Laxe, & Montes, 2013).

In recent research the assortativity coefficient and the rich-club coefficient are used to measure how large nodes connect with each other over time in the global shipping network (Ducruet, 2017). This gives insight on the centrality of the network through the connectivity of both large and small nodes. The assortativity coefficient is based on the Pearson correlation between the degree centrality of two connected nodes and indicates to what extent nodes with comparable connectivity connect with each other (Ducruet & Lugo, 2013). The rich-club coefficient is the ratio between the gamma index of the entire network (the reason why it sometimes referred to as the rich-club index), and the gamma index of the nodes connected with at least a certain number of nodes (this threshold is explained later on). The gamma index is a measure between zero and one and considers the relation between the observed links and the number of possible links. The rich-club index can be used to tell how well nodes with a certain degree are connected with each other. A decreasing gamma index means the network is becoming more simple and central, so larger nodes connect more with smaller nodes. A decreasing rich-club index indicates the most connected ports become less connected with each other.

2.3 Using Service Throughput Time as Link Weight

The economies of scale generated by the large European ports enable them to operate high frequency intermodal transport to many hinterland destinations (Van Klink & Van den Berg, 1998). The case of a new rail shuttle service between Rotterdam and Northern Italy demonstrates a higher travel frequency promotes a larger flow of container traffic between these regions and lowers the number of containers handled by Italian ports (Van Klink & Van den Berg, 1998). How this influences the network structure is not known, but since higher container volumes in hubs enable them to strengthen their competitive position (Rodrigue, Debrie, Fremont, & Gouvernal, 2010), it is probable the network structure surrounding the Northern Italy rail destinations is affected. There is also a direct link between travel frequency and the inland service area of a port: increasing the travel frequency by rail or barge shuttle services has a positive impact on the service area of a specific hinterland area (Notteboom, 2008). With a larger flow of containers to a hinterland nub, is it economically viable to offer extra services from a hub onwards, thus changing the connectivity of the entire area.

Indicators of connectivity and accessibility prove to be useful in an assessment of the European deep-sea ports and their intermodal transportation options to the hinterland (De Langen & Sharypova, 2013). Connectivity is a network measure indicating the possibility to reach all nodes from all other nodes, and accessibility is a node measure indicating the possibility to reach all or certain nodes form a specific node. While the research is limited to 26 ports, it does indicate an increase in travel frequency of rail and barge strengthens the connectivity of specific nodes and is altering the overall connectivity of the network. No connection to measures indicating the network structure are made however. The same goes for the influence of travel time on the choice of modality. Research is widely available (see Reis (2014) for an overview) and known to influence the connectivity of a node (e.g., the shuttle service Rotterdam-Northern Italy), but no connection to the network structure is made. However, for both shippers and operators, it is an important variable being considered when choosing between road or rail and barge transport to the hinterland (Reis, 2014).

The time dimension has proven to be of importance to implement in intermodal transport models (Crainic & Kim, 2007). First because the time dimension is a requirement for the effective scheduling, routing, and coordination of assets. Second, the wide range of time-factors (dwell-times, frequencies, transport times, number of stops) can result in significant different outcomes of expected service quality when the network is arranged in a different way (Ypsilantis, 2016). As service quality is often associated with, among others, frequency of service (Li & Tayur, 2005) in the decision for a certain carrier (Crevier, Cordeau, & Savard, 2012), it is of importance to be considered. Operators should therefore make strategic use of the time dimension for increasing their competitiveness.

3 Research Methodology

In this chapter the dataset for this research is presented and methods for cleaning and making the dataset useful for analysis are described. In the second part, the dependent and independent variables from the conceptual model are operationalised.

3.1 Structure of the Dataset

Ecorys is a leading international research and consultancy company aiming primarily at handling important societal challenges. Established in 1929, it has an extensive history of basing their advice on sound research. Clients are, among others, the European Commission, the World Bank, airports, hospitals, universities, national and local governments, and a long list of private companies such as KPN Telecom and Royal HaskoningDHV. The main areas which Ecorys operates in are: Economic Growth, Social Affairs & Health, Natural Resources, Regions & Cities, Public Sector Reform, Security & Justice, and Transport & Infrastructure.

A Transport & Infrastructure project Ecorys has initiated in 2013 is the development of a tool for scheduling intermodal container transport in Europe: Intermodal Links. The database contains around 150 intermodal operators, carrying out 25,000 weekly departures by rail, barge, Roll-on Roll-off ferries, and short-sea shipping between 1,000 terminals. Companies are, after payment, able to plan their transport through this tool, or even integrate it within their enterprise system. The goal is to make intermodal transport more efficient in terms of time and costs. Another group of Intermodal Links-users are port authorities, local governments, and terminal operators. Since the database with service connections is updated continuously, this group of users can track if, and how, the interest of service operators, and thus the shippers as well, is changing. The port authority, local government, or terminal operator can adapt its strategy to become more competitive in this changing environment.

The dataset provided by Ecorys contains snapshots of three consecutive years, 2016, 2017, and 2018. A summary of the development over the three years in given in table 2 (excluding short-sea shipping and ferries). An entry is the connection between two terminals, listed as country-city-terminal, by a certain service provider, with a specific mode of transport, including the departure days in a week and the travel time in days. The specific modes of transport are railway and inland barge shipping. The top five connected cities for 2016 are presented in table 3 (excluding short-sea shipping and ferries).

Descriptive	**2016**	**2017**	**2018**
Rail connections	2,126	2,171	2,111
Barge connections	567	734	808
Number of countries	24	24	25
Number of cities	324	334	330
Number of terminals	443	485	481
Number of operators	102	112	116
Unique directional links between cities	1,648	1,656	1,673
Average departures per week rail	4.8	4.8	4.7
Average departures per week barge	3.2	3.1	2.8
Average travel time rail [days]	1.8	1.9	1.9
Average travel time barge [days]	2.4	2.6	2.5
Average distance rail [km]	536	543	651
Average distance barge [km]	198	368	201

Table 2: Database descriptives over time (excluding short-sea shipping and ferries).

Country and city of origin	Weekly outgoing connections	Weekly incoming connections	Connected unique countries *	Connected unique cities **
Netherlands – Rotterdam	1,011	994	9	103
Germany – Hamburg	842	781	9	64
Belgium – Antwerp	577	540	12	90
Italy – Milan	447	471	8	42
Germany – Bremerhaven	365	376	5	40

*Table 3: The top five connected terminals in 2016 (excluding short-sea shipping and ferries). *Max:24, **Max: 324.*

Several modifications are made to the dataset in order to make it suitable for this research. The short-sea shipping links are removed from the database for three reasons. First, the focus of this research is on intermodal hinterland transport and short-sea shipping is not part of this in most definitions (De Langen et al., 2017). Second, the short-sea shipping connections include feeder services, but these are not stable on the service schedules provided. Third, inter-continental container vessels with multiple calls in Europe are not included in the database. Out of all the entries, approximately 500 entries contain the input “Terminal not specified” (for 2016 and 2017, this issue was fixed in 2018). Since the research will focus on connection between cities, and not terminals, this does not pose a problem. However, when the data is aggregated on a city level, some connections within one city arise. These three entries are removed from the dataset in 2016.

Next, the services running on different days of the week with the same origin and destination, and provider are merged. Sometimes services are registered as two separate services running on different days a week but are the same from a customer perspective. An example is a service with a travel time of four days leaving on Monday and Friday: the shipment leaving on Monday will arrive on Thursday, but the service on Friday will arrive on Tuesday since most services do not operate on Sundays. The travel time is therefore recorded as five instead of four days while it is actually the same service. This is solved by averaging the travel time of the two separate services. The last modification is the merging of services with the same origin and destination, operator, modality, frequency per week, and days of the week in service. This aggregates the calls a service provider makes to multiple terminals within one city on the same service. It is for instance rational for a barge service from Rotterdam to Duisburg to load its ship with containers from multiple terminals located in Rotterdam. These are registered as separate services but there is still only one ship sailing between the two cities for the particular service.

Modifications	2016 [remaining]	2017 [remaining]	2018 [remaining]
Raw dataset	12,968 [100%]	14,400 [100%]	15,876 [100%]
1 Short-sea shipping	- 6,385 [50.8%]	- 7,558 [47.5%]	- 8,878 [44.1%]
2 Within city connections	- 3 [50.7%]	0 [47.5%]	- 16 [44.1%]
3 Service aggregation 1	- 2,118 [34.4%]	- 2,226 [32.1%]	- 1,989 [31.5%]
4 Service aggregation 2	- 1,769 [20.8%]	- 1,711 [20.2%]	- 2,074 [18.4%]
Final dataset	2,693 [20.8%]	2,905 [20.2%]	2,919 [18.4%]

Table 4: Cleaning the dataset.

The dataset is continuously updated by employees from Ecorys to guarantee the data quality demanded by their customers: e.g. the port authorities of Zeebrugge, Amsterdam and Rotterdam, Zeeland Seaports, and Kombi-Terminal Ludwigshafen. By experience Ecorys knows in which period of the year service providers change their schedules. The new schedules are looked-up on the websites of the service providers, or the service providers supply the schedules themselves. It is not known what percentage of the total volume transported is done through these scheduled links (i.e., the volume over unscheduled links is not known as well), but Ecorys does estimate its dataset contains over 90 percent of all scheduled services in Western Europe. In 2018, mainly East and South-East European services have been added. It should be remarked the dataset contains the supply of services operating at least one times a week, not the actual demand, capacity, or prices. The services only span container transport, not bulk, break-bulk, or services dedicated to specific companies (e.g., a car manufacturer shipping a full trainload of cars from a factory to a port each week).

3.2 *Research Design*

This research is a study on the effects of service throughput time on changes in network structures. The transport modalities rail and barge are used as moderating variables between service throughput time and the changes in network structures. Observations are made on a longitudinal dataset of three years. For each snapshot, global network measures and node attributes will be calculated. By observing both the changes in the network measures and node attributes, and the change in service throughput time, a relationship between the two is established. The last part of the data analysis is focussed on whether the influence of service throughput time on changes in network structure is different if only the rail or barge network is considered.

For visualising the network and extracting the measures from it, three statistical programs are used. Tulip 5.1.0 (Auber & Mary, 2007) and Gephi 0.9.2 (Bastian, Heymann, & Jacomy, 2009) are used to visualise the network, but also have the ability to extract measures. The NetworkX 2.1 package (Hagberg, Schult, & Swart, 2008) for Python 3.6.4 (using Visual Studio Code 1.23.1 for the interface) and the tnet 3.0.14 (Opsahl, 2009) and brainGraph 2.2.0 (Watson, 2018) packages for R 3.5.0 (using Rstudio 1.1.453 for the interface) are used for statistical analysis on the networks. Microsoft Excel is in the first instance used for cleaning the data and making it suitable for import in one of the three programs above; but is also used to perform additional computations on output from these programs. IBM SPSS is used for establishing the correlations, regressions, and levels of significance between service throughput time and the network measures.

3.2.1 Operationalising the variables

The dependent variable, network structure changes, is measured on both a global, network, level and a local, node, level. Table 5 provides an overview of measures for both of the levels. All measures are taking into account the network is directed. Nodal measures can be aggregated to a network level by taking the sum of the nodal measures and dividing it by the number of nodes (N) in the network.

The network consists of a graph $G = (V, E)$ where there is a set of nodes V with index v, and a set of edges E with index e. An edge e_{ij} connects node v_i with node v_j. M is the number of edges in set E, N is the number of nodes in set V. $d(v_i, v_j)$ denotes the shortest path between two nodes. ℓ is the sum of the length of $d(v_i, v_j)$ and α is the symbol for a link weight

Network measure	Definition	Formula
Hierarchy	Exponent of the slope of the power-law line (h) drawn by plotting node frequency over degree centrality	$y = \alpha x^h$
Assortativity coefficient	Pearson correlation between the degree (k) of each two neighbouring nodes	$\tau = \frac{M^{-1}\sum_i j_i k_i [M^{-1}\sum_i \frac{1}{2}(j_i + k_i)]^2}{M^{-1}\sum_i \frac{1}{2}(j_i^2 + k_i^2) - [M^{-1}\sum_i \frac{1}{2}(j_i + k_i)]^2}$
Weighted assortativity coefficient	Pearson correlation between the weighted degree (k^w) of each two neighbouring nodes	$\tau^w = \frac{M^{-1}\sum_i j_i k_i^w [M^{-1}\sum_i \frac{1}{2}(j_i + k_i^w)]^2}{M^{-1}\sum_i \frac{1}{2}(j_i^2 + (k_i^w)^2) - [M^{-1}\sum_i \frac{1}{2}(j_i + k_i^w)]^2}$
Average shortest path length	Average number of stops between two nodes	$l_G = \frac{1}{N(N-1)\sum_{i,j} d(v_i, v_j)}$
Weighted average shortest path length	Average number of stops between two nodes with link weights (α)	$l_G^w = \frac{1}{N(N-1)\, min \sum_{i,j} d(v_i, v_j)\alpha_{ij}}$
Gamma index	Observed links over the possible number of links	$\gamma = \frac{2e}{[N(N-1)]}$
Rich-club coefficient	Extent to which nodes above a certain degree threshold (k) are connected	$\phi(k) = \frac{2Ek}{Nk(Nk-1)}$
Weighted rich-club coefficient	Extent to which nodes above a certain weighted degree threshold (k^w) are connected	$\phi(k^w) = \frac{2Ek^w}{Nk^w(Nk^w-1)}$
Gini-index	Concentration of a variable over nodes. σX and σY denote the cumulative proportions	$Gini = \left\lvert 1 - \sum_i (\sigma Y_{i-1})(\sigma X_{i-1} - \sigma X_i) \right\rvert$
Network diameter	The longest shortest path (ℓ) in the network	$max(\ell)$
(Weighted) Modularity	Splits the network so it will form high connectivity communities	Algorithm from Blondel, Guillamme, Lambiotte, & Lefebvre (2008)
Nodal measure		
Degree centrality	Number of neighbouring nodes	$k_i = C_D(i) = \sum_j^N e_{ij}$
Weighted degree centrality	The sum of the average link weights (α) to all neighbouring nodes	$k_i^w = C_D^w(i) = \sum_j^N \alpha_{ij}$
(Weighted) Betweenness centrality	Number of times a node is crossed by (weighted) shortest paths	$C_B(i)\ or\ C_B^w(i) = \frac{d_{jk}(i)}{d_{jk}}$
Closeness centrality	Sum of the length (ℓ) of the shortest paths to all other nodes	$C_c(i) = \frac{N}{\sum_{k \neq i} \ell(i,k)}$
Clustering coefficient	Share of actual links between nodes within the neighbourhood (L_i) divided by the maximum possible number of links	$C_l(i) = \frac{L_i}{\frac{k_i(k_i-1)}{2}}$
Average nearest neighbour degree	Average degree centrality of neighbouring nodes	$K_{nn}(i) = \frac{1}{k_i}\sum_j^N k_j$
Average weighted nearest neighbour degree	Average degree centrality of neighbouring nodes with link weights (α)	$K_{nn}^w(i) = \frac{1}{kw_i}\sum_j^N k_j \alpha_{ij}$

Table 5: Measures used for complex networks on a global and local level. From Ducruet & Lugo (2013), Guimera et al. (2005), Van den Heuvel, De Langen, Van Donselaar, & Fransoo (2013), De Langen & Sharypova (2013), Ciliberto, Cook, & Williams (2017), and Calatayud, Mangan, & Palacin (2017).

The independent variable is the service throughput time: a link weight based on the frequency of incoming and outgoing services from direct links with any modality, and the travel time of each of those services. It is calculated by adding the average waiting time if a customer would need the service to the travel time of the service (equation 1). The variable frequency is measured as number of times per week, therefore it is the denominator. The function is based on the queuing formula commonly known as Little's Law (Little, 1961).

$$Service\ Throughput\ Time = Travel\ Time + \frac{1}{2} * \frac{7}{Frequency}$$

Equation 1: Service throughput time of one service between two links.

The average service throughput time of all links between two nodes is calculated by summing the service throughput time of all the services and divide it by the number of services between two nodes (equation 2). Examples of equation 1 and equation 2 are demonstrated in table 6.

$$Average\ Service\ Throughput\ Time = \frac{\sum Service\ Throughput\ Time}{n}$$

Equation 2: Average service throughput time of all services between two nodes.

Services	**Travel time in days**	**Frequency per week**	**Average waiting time**	**Service throughput time**
A → B	1	7	0.5	1.5
A → B	2	14	0.25	2.25
			Average A → B	**1.7**
C → D	1	12	0.29	1.29
C → D	2	3	1.17	3.17
			Average C → D	**2.23**

Table 6: Examples of service throughput time as link weight for link A→B (1.7) and C→D (2.23).

Whether the impact of the service throughput time is different for barge and rail transport is analysed by splitting the network for the two modalities. The same measures from table 5 will be used to describe the networks. Service throughput times will be calculated separately for these two networks by only taking the specific modality into account. By observing the development of the modalities separate from each other, the impact of service throughput time can be interpreted and correlated.

4 Results

4.1 Changes in Network Structure

Before the possible effects of service throughput time are examined, it is important to observe in what way the network structure is changing. The database descriptives (table 2) show the number of rail connections is stable over the course of 3 years, just like the average number of departures and the average travel time of this modality. The only remarkable change for rail transport is the 21 percent increase from 2018 compared to 2016 in average distance. Part of this increase can be attributed to the inclusion of a rail connection between Tilburg (The Netherlands) and Chengdu (China) of over 8,500 km (and vice versa). However, if this connection is removed, the average distance is still 639 km. Developments in barge transport seem to be more significant: over 200 new services are added and the average number of departures per week is decreasing by 0.4 from 2016 to 2018. Finally, the number of unique directed links show if services were added on already established links, or if new connections between cities were made. These numbers are slightly increasing over time; meaning new connections between cities are made each year. The gamma index (table 7), which is the fraction between the number of actual connections versus the total number of possible connections, confirms this. It is stable because only a small amount of unique connections is added to a network which has many possible routes between its 330 cities.

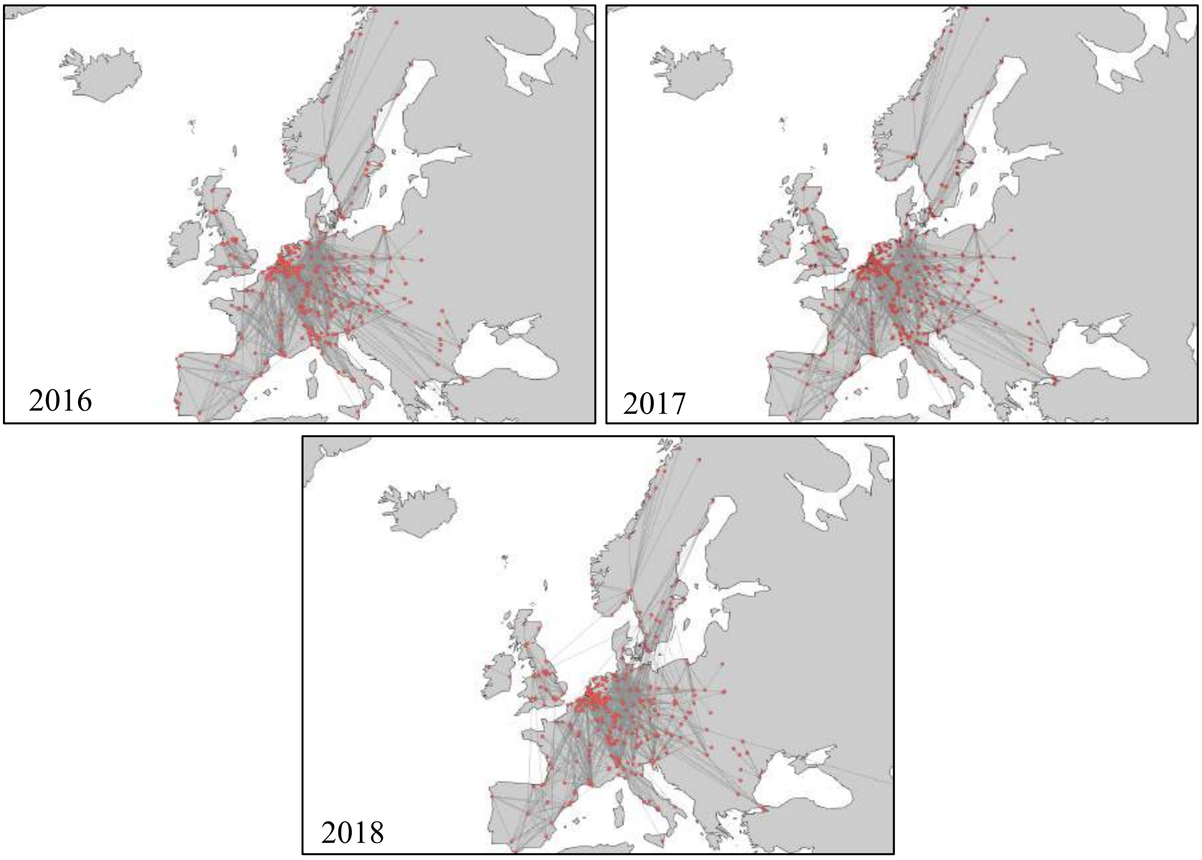


Figure 5 (.1, .2, .3): The complete network of barge and rail connections, no link weights. Created with Tulip.

Figure 5 visualises the complete network with all barge and rail connections. The absence of clear visual changes supports the observations from the database descriptives. In order to be able to see new connections arise on already existing routes, the networks are visualised with size and colour scales in appendices 2-4. In all the visualisations in the appendices which have a circular shape, the force-directed Fruchterman-Reingold algorithm is used (Fruchterman & Reingold, 1991). In the appendices 2-4 it is clear from the thickening and colouring of edges new services were mainly established on already existing routes in 2017 and 2018. The absence of changes in the outskirts of the figures in the appendices indicates new services are primarily added on routes from and to cities with a high number of connections.

4.1.1 Hierarchy

The exponent of the slope of the power-law line (hierarchy) is increasing in each consecutive year (figure 6 and table 7). The power law is a functional relationship between two values: node frequency and degree centrality in this case. One value will vary as the power of another regardless of the initial amount of the value. If the exponent is larger (>1), a relative change in degree centrality will result in a larger decrease of node frequency. Therefore, if the exponent is high, there will be a large number of low degree nodes and a low number of high degree nodes (a steep downward line). This indicates a centralised network in which high degree nodes connect to many low degree nodes. When the exponent is above 1, a network can be described as scale-free (Barabasi & Albert, 1999). As the measures from the datasets are below 1, the network for container transport by rail and barge is not scale-free. However, as the exponent is increasing, the network shows signs of slowly becoming so. The reason the network is not scale-free is primarily due to the clique located on the bottom left of the trend line; if the eight nodes with degree centrality values 1, 3 and 5 are removed in 2017, the hierarchy is 1.021. For the structure of the hinterland transport the increase in hierarchy indicates more importance is placed on nodes with a large number of connections.

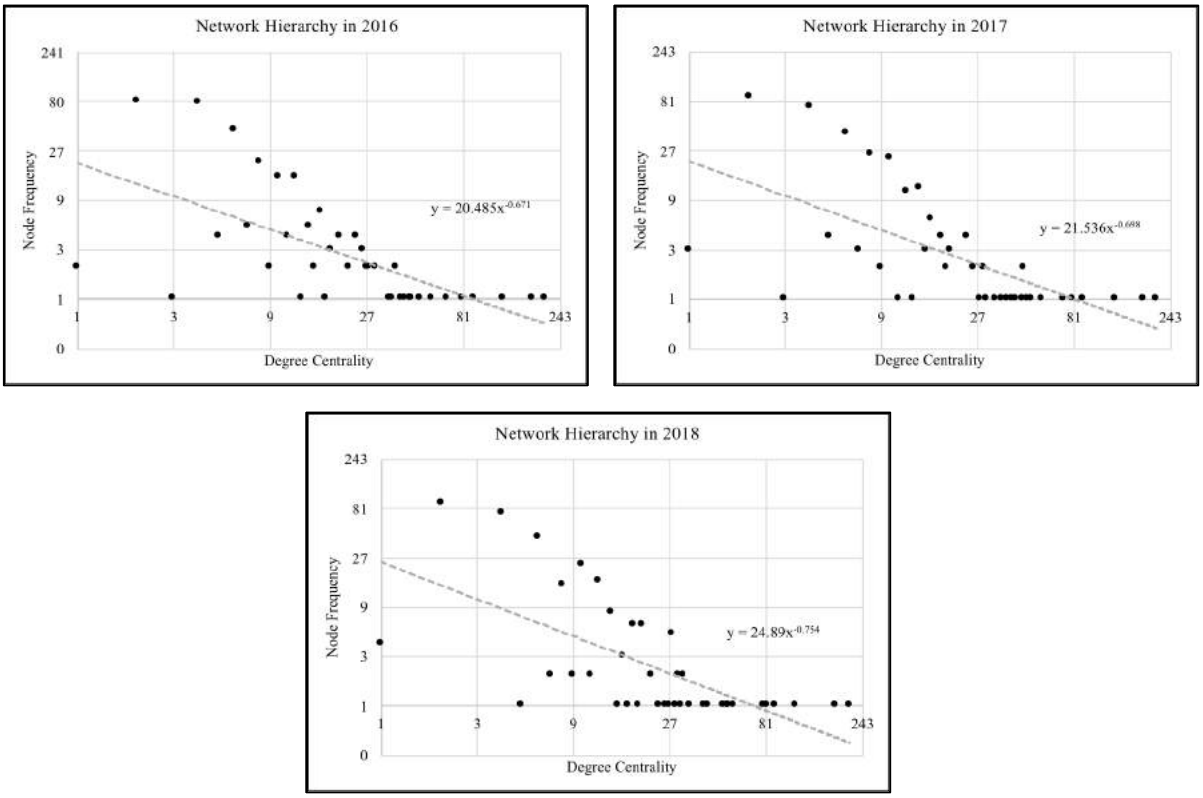


Figure 6 (.1, .2, .3): The exponent of the slope of the power-law line indicates whether the network is becoming more (1) or less (0) centralised. Computed in Gephi and visualised in Microsoft Excel.

4.1.2 Assortativity

The assortativity coefficient shows a weak to moderate negative linear relationship between the degree centrality of two neighbouring nodes. Positive values of the coefficient mean there is a correlation between nodes of a similar degree, while negative values indicate nodes of a different degree are connected each other. So, if two connected nodes are randomly selected and one of them has a high degree, there is a weak to moderate chance the other node has a lower degree (and vice versa). The decrease of the assortativity coefficient indicates the chance of a high degree node connecting with a low degree node is getting larger. In terms of hinterland connectivity this corresponds with the hierarchy measures implying higher degree nodes are increasing in importance.

4.1.3 Rich-club

The actual change in connectivity of low and high degree nodes can be separated with the rich-club coefficient. The rich-club measures the extent to which nodes above a certain threshold (degree) are linked with each other. Table 7 lists both the rich-club coefficient for nodes with a degree over 5 and nodes with a degree over 20. A high rich-club coefficient means many links exist between nodes over a certain degree. The increase in rich-club for nodes with a degree over 5 indicates the connectivity between nodes is getting larger; this aligns with the observations from the previous two paragraphs. However, it is not clear whether these new links are made between low and high degree nodes or between high degree nodes. In order to see where the changes occur, the rich-club coefficient with a degree threshold of 20 is used. The rich-club [≥20] increases significantly from 2016 to 2017, just like the rich-club [≥5]; so, the increasing connectivity between these two years cannot solely be attributed to nodes with a degree between 5 and 20. However, the rich-club [≥20] is decreasing in the following year, while the rich-club [≥5] is still increasing. The change in connectivity can thus be attributed to nodes with degrees between 5 and 20. If the measures in 2016 are compared with 2018, it can

Network and node level measures	**2016**	**2017**	**2018**
Hierarchy	0.671	0.698	0.754
Assortativity coefficient	-0.272	-0.333	-0.353
Average shortest path length	3.760	3.830	3.653
Gamma index	0.016	0.015	0.015
Rich-club coefficient [degree ≥5]	0.069	0.093	0.097
Rich-club coefficient [degree ≥20]	0.421	0.476	0.409
Gini-index [degree centrality]	0.573	0.576	0.582
Gini-index [edge distribution]	0.670	0.718	0.706
Average degree centrality	5.055	5.006	5.018
Average betweenness centrality	848	881	805
Average closeness centrality	0.290	0.288	0.307
Average clustering coefficient	0.361	0.331	0.315
Average nearest neighbour degree	59.051	84.591	76.760
Network diameter	11	12	11
Number of communities	13	13	14

Table 7: Network and node level measures. Extracted from Tulip, Gephi, and computed in Microsoft Excel, brainGraph for R, and NetworkX for Python.

be concluded nodes with a degree above 5 but under 20 are increasing in connectivity. Whether this is with high degree nodes or with nodes of a similar degree cannot be determined from the rich-club measures, but the decreasing assortativity coefficient yields this is primarily with high degree nodes. An interesting observation for the structure of the hinterland is the decreasing number of connections between nodes with a degree over 20. This can imply the exchange of containers between large hubs is declining and hinterlands are more served over specific corridors rather than a flexible route which can cross several large hubs.

4.1.4 Gini-index

The Gini-index is a construct used to measure (in)equality by calculating the difference between the frequency of observations with a line of perfect equality. It is computed by constructing a Lorenz curve ranking the frequency of observations (degree or number of edges is this case) from low to high and calculating the deviation from the line of perfect equality. The share of the area between the line of perfect equality and the Lorenz curve compared to the total area below the line of perfect equality is the Gini-index. If the area is small, so closer to zero, there is more equality because the Lorenz curve will be closer to the line of perfect equality. The increasing Gini-index [degree centrality] (table 7) indicates the distribution of unique edges is less dispersed; so, nodes already having a large number of connections are only getting more of them. The rise in the Gini-index [edge distribution] (table 7) confirms this: the total number of edges is distributed over less nodes.

This supplements the assortativity measures used to conclude low degree nodes are increasing their connectivity with high degree nodes based on the rich-club index. Thus, hubs in the European container transport network are getting stronger because operators choose to increase the number of services departing from or arriving at these hubs. The hubs are not necessarily getting stronger because the number of destinations which can be reached is increasing. Appendix 1 support this: the table shows the development of degree values of the 30 largest degree nodes over the course of three years and indicates the number of destinations reachable from these 30 nodes is actually on average decreasing over three years (14 out of the 30 see an increase in degree). It is possible these connections are primarily between the nodes listed in the table because the rich-club [≥20] already showed the connectivity between these nodes is declining.

4.1.5 Centralities

The difference between average degree centrality and average nearest neighbour degree give insights in how nodes with a different degree are linked to each other. The average degree centrality is the average number of unique connections per node for all nodes in the network. The average nearest neighbour degree indicates the average degree centrality of all the direct neighbours of a certain node aggregated to a network measure for all nodes. If the measures are equal, nodes with a similar number of unique links to other nodes are strongly connected to one another. If the average degree centrality is lower than the average nearest neighbour degree, then low degree nodes are more often linked to high degree nodes. In the observed networks the latter is the case: the average degree is smaller than the average nearest neighbour degree (table 7); so, low degree nodes are primarily connected to high degree nodes.

This is supported by the high average betweenness centrality (table 7): the number of shortest paths running through a node averaged over all nodes in the network. High betweenness centrality values can also occur if a node is acting as a bridge linking two communities with one another; any shortest path between these two communities is running over the bridge. This is however not the reason for the high values in the observed networks as several measures have already showed low degree nodes are primarily connected to high degree nodes. The high average betweenness centrality value is caused by the shortest paths in the network running through a selection of nodes, the hubs. The hubs link different parts of the network with one another; and link cities strongly connected to one hub but not to each other. The hubs have to be crossed in order to create a shortest path between two cities not located in the same part of the network or between cities in the same area but not connected to each other.

Both the average degree centrality in combination with the average nearest neighbour degree, and the average betweenness centrality indicate a centralised network because of the importance of hubs and the low connectivity between low degree cities. Over the three years in the dataset, the betweenness centrality is first increasing and then decreasing. An increase suggests certain hubs have gained importance and interconnectivity between low degree cities is decreasing even further. The subsequent decrease could be the results of the offering of more direct services: if these by-pass hubs, then the number of shortest paths going through the hubs will decrease. This would also justify the increase in average distance of train services. The increase in the average closeness centrality in 2018 compared to 2017 is in line with the trend of the average betweenness centrality in those years. Closeness centrality is the number of nodes in the network divided by the average length of all shortest paths from one node. For Rotterdam is it for instance close to 1 because it has many direct connections to other nodes. Hence, as the average closeness centrality of the network is increasing, the average shortest path length is decreasing, and the network is becoming more central. Just like the rich-club index, the increase of direct services could signal the increasing importance of corridors crossing less cities on their way.

4.1.6 Communities

It is possible networks are both scale-free and small-world. This is achieved by adding a few random edges in a structured scale-free network (Watts & Strogatz, 1998). These new edges are characterised as weak links because the high degree nodes in the scale-free network are already connecting most nodes. As is concluded, the network of container transport is currently not scale-free; although it shows signs of becoming so and calculations have shown a lack of connections to only 8 out of the 330 nodes cause the network not to be classified as scale-free. Since small-world networks show regional specialisation and efficient transfer of information within its communities (Watts & Strogatz, 1998), it is interesting to research whether this type of network classification can be applied to the datasets.

Small-world networks tend to consist of communities which are interconnected by hubs. The communities in the network often have high average clustering coefficients (> 0.7), small diameters (< 5), and low average shortest path lengths (< 3) (Ducruet & Notteboom, 2012). The clustering coefficient is based on the number of triplets in the network. A triplet is a set of three nodes which is connected by two (open triplet) or three (closed triplet) edges. The number of closed triplets divided by the sum of open and closed triplets, averaged over all nodes in the community or network, is the average clustering coefficient. If the coefficient is higher, the network is more connected because there is a larger number of closed triplets; however, it is not yet clear whether this is because just a few high degree nodes increase the average. The longest shortest path in the network is defined as the diameter. If the diameter would be large, there is a lower chance the average shortest path length is low because some nodes are positioned far away from each other. To define small-world networks, the average shortest path length has to be low in addition to a high clustering coefficient. This shows there is indeed a good connectivity between nodes since the average shortest path would be longer if only some high degree nodes would have caused a high clustering coefficient.

Appendices 5-7 display the communities identified in the unweighted networks for 2016-2018. Based on maximizing the modularity with an algorithm (Blondel et al. 2008) adding and removing nodes to or from a certain module (i.e., establishing strong connections within a community, and weak ones with other modules), communities are detected and visualised using Gephi. The communities show several small-world characteristics as it comes to the average shortest path length and the network diameter. However, the clustering coefficient is not over

0.7 for any of the communities in the three years of data. This means the interconnectivity between all nodes within a community is low, but the interconnectivity through central nodes in the communities is high as average shortest paths are small and the diameter is so as well.

Although the network is not small-world, it is still interesting to make some observations about the detected communities. The number of communities is for instance stable (table 7), but the composition of the communities is changing (appendices 5-7). For example, Norway and Sweden are in 2017 and 2018 considered as two separate modules, meaning there is a lack of interconnectivity although they are neighbouring countries positioned on outskirts of Europe; which was not the case in 2016. The turnaround can be explained by comparing the measures between the three years. The combination Norway/Sweden displays a relatively low average clustering coefficient in 2016. When taken apart in 2017, Norway shows a significantly higher clustering than Sweden (then belonging to Germany/Sweden). For Sweden this means the interconnectivity to other modules just needs a little increase in order to be separated from Norway. Cities in smaller countries such as Hungary, Czech, Austria, Bulgaria, Romania, and Swiss tend to be switching communities based on how good they are connected to the most nearby hub. This can be useful strategical information for rail or barge operators when they want to increase access to certain hinterland areas.

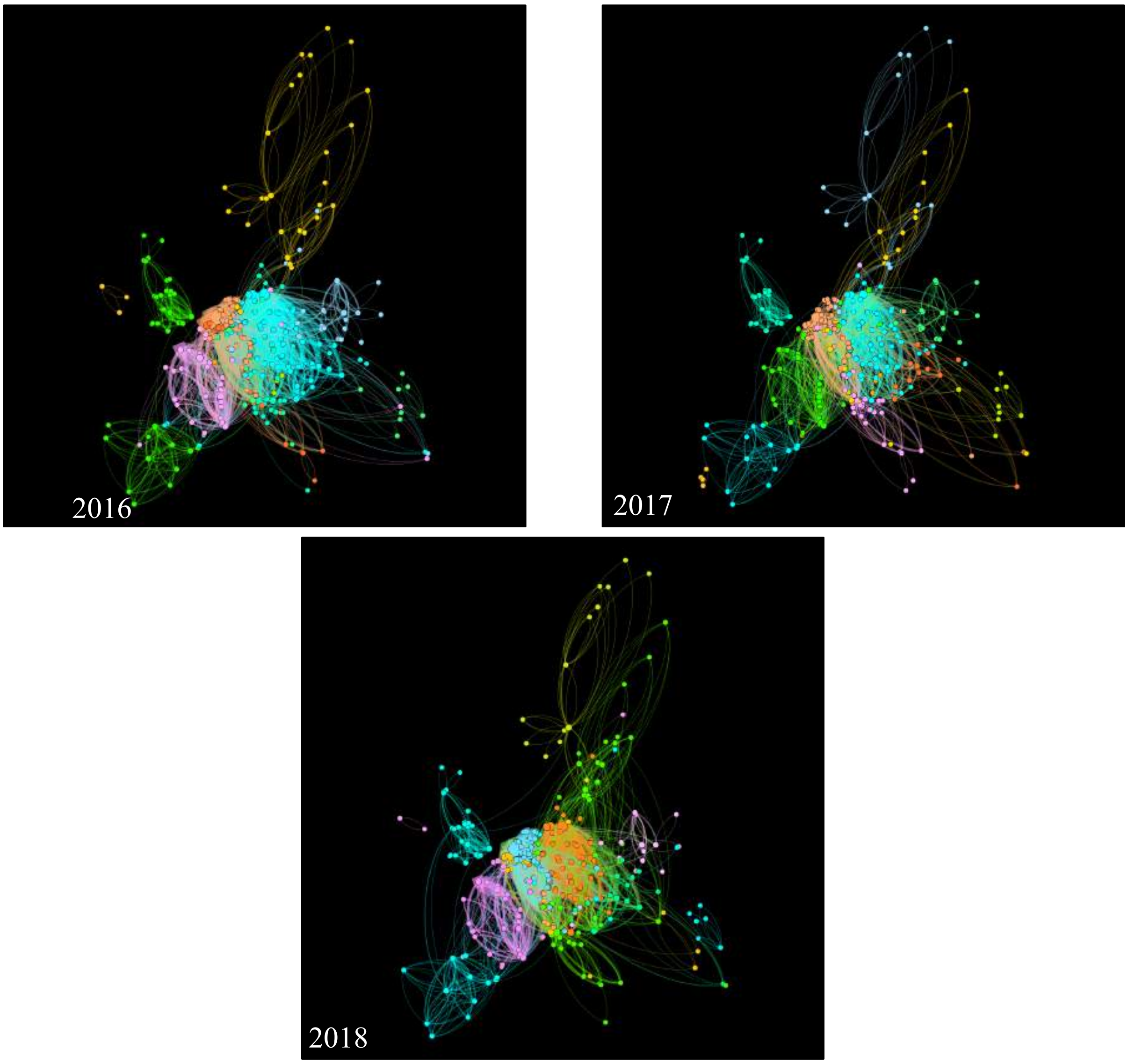


Figure 7: The community structures in 2016-2018 with a geographical layout. Visualised in Gephi.

However, there are also stable communities. Examples are the UK and Ireland for the obvious reasons they are islands and the short-sea shipping links have been removed; the barge network in France sailing on the Rhône, Somme and Seine rivers; and remote areas such as the Trans Alp Roll-on Roll-off train. Not completely remarkable is the observation the communities consisting primarily of a single country show some of the highest clustering coefficients and thus look more like small-world networks. Most striking is how geographically dispersed some of the communities are which consist of two or more countries and have an equal number of nodes in each country: Germany/Spain in 2018 for example. When looking at the actual links in this specific community, a high frequency train link offered by the operator Kombiverkehr from Ludwigshafen (Germany) to Barcelona, Granollers, Madrid, and Tarragona is responsible for establishing the modularity. So, train links with a sufficient frequency have the ability to form corridors and make connections between distant countries.

4.2 Service Throughput Time as Link Weight

Service throughput time is now applied as a link weight in order to analyse its possible influence on changes in the network structures described in chapter 4.1. Figure 8 shows the structure of the frequency per travel time is not changing significantly. What is odd about figures 8.1 and 8.2, is the increase in frequency if the transport time is larger than seven days. This trend is reversed in 2018: an increase in travel time leads to even lower frequencies (the highest travel time of 17 days is a rail service from The Netherlands to China). In 2016, some of the services covered a large distance, from Belgium to Turkey and Greece for example, explaining their long travel times. However, these services do not have a high frequency. The high frequency and long travel time services are primarily barge routes from Antwerp sailing over the Rhine river, and rail connections from Genk (Belgium) to Bulgaria and Romania. The latter are offered six times a week and have a travel time varying between the eight and ten days. These rail connections disappear on the right side of figure 8.3 because the operators have decided to reduce the number of cities with direct connections from Genk. An explanation for the long travel times in barge transport is two-fold: barges are slow, and there are a lot of ports to call along the way, resulting in dwell-time. The 2017 data shows a similar trend.

Figure 9 compares the average distance in kilometre versus the travel time in days. As already indicated in figure 4 by De Langen et al. (2017), the frequency of service is decreasing when the distance is increasing. The 2017 and 2018 datasets show a relative stable trend in average distance when the travel time is increasing, but 2016 shows the opposite: the average distance is increasing as the travel time is getting longer. The changes in 2017 and 2018 compared to 2016 can be attributed to the increase in slow barge services from Rotterdam and Antwerp sailing over the Rhine, Maas, and Moselle rivers to Germany, France, and Switzerland. The disappearing of the longer rail connections described in the previous paragraph are also contributing.

The average service throughput time is slowing down from 3.1 in 2016, to 3.2 in 2017, and finally 3.3 in 2018. In the visualisations of the weighted networks in appendices 5-7, the slower average service throughput time is also visible: there are less brown edges, and the green edges, indicating slow service throughput times, are becoming thicker. The slow connections are primarily seen around large hubs such as Antwerp, Rotterdam, and Hamburg. This is rational as these cities have the cargo volumes to fill trains or barges leaving to more distant areas which take more days to reach and have a less departures. A second reason, specifically for Antwerp and Rotterdam, for the large amount of thick green edges, is the development of a relatively slow barge network sailing over the Rhine, Maas, and Moselle rivers.

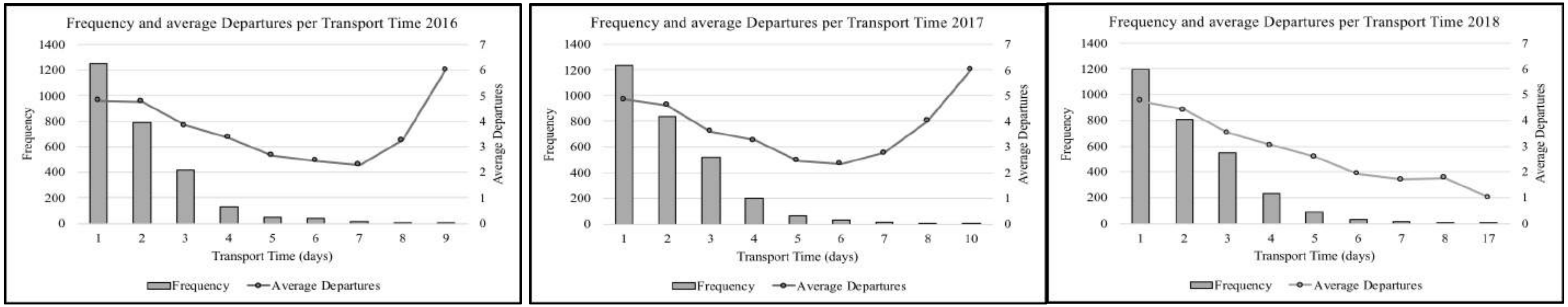


Figure 8 (.1, .2, .3): If the transport time is increasing, the number of services offered and the average number of departures of a service are (generally) decreasing.

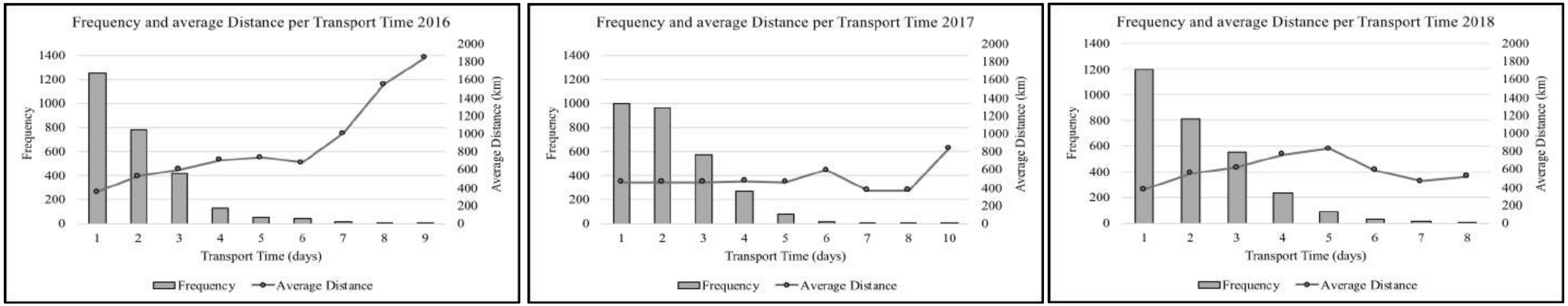


Figure 9 (.1, .2, .3): If the transport time is increasing, the number of services offered is decreasing and the average distance is increasing. The link from The Netherlands to China is left out in figure 9.3 because it distorted the graph.

4.2.1 *Weighted centralities*

For a network with link weights, many of the same measures can be computed as one without (table 8). The average weighted degree centrality is the sum of the average service throughput times over all unique edges a node has. This is different than the degree centrality in an unweighted network because it does not take the number of links into account. However, it is the preferred measure for analysing weighted networks (Barrat, Barthelemy, Pastor-Satorras, & Vespignani, 2004). An increase of this measure over the course of three years means the average service throughput times of unique edges are slower, or the number of unique edges per node is decreasing. With average service throughput times of 3.1 in 2016, 3.2 in 2017, and 3.3 in 2018, the number of unique edges (degree centrality in an unweighted network) would be estimated at 4.979 (15.343 / 3.1) per node in 2016, 4.883 in 2017, and 4.956 in 2018. This is stable, just like in the unweighted network, and also close to the values in table 2. The increase of the average weighted degree centrality is thus caused by slower average service throughput times and not by a decrease in edges. As a consequence of the increasing average degree, the average weighted nearest neighbour degree is increasing as well.

The betweenness centrality for weighted graphs is unlike the unweighted variant not based on shortest paths crossing the least number of nodes, but on shortest paths with the lowest total sum of weights. A higher betweenness centrality in a weighted network compared to an unweighted network means there are faster routes connecting any two random nodes than the one with the least number of edges. This is because the high degree nodes often have faster connections, so their betweenness centrality will be higher. As the observed weighted betweenness values are consequently higher than the unweighted values (848, 881, 805), more shortest paths routed through a selection of high degree nodes. The difference between the weighted and unweighted variant is however decreasing (110, 80, 67); thus, less shortest paths based on service throughput times are going through high degree nodes. This could be caused by cities with an average number of connections, which can be classified as inland hubs, having fast services between cities which have less connectivity to the network.

The trend in average weighted shortest path length from 2016 to 2017 can be explained by the increase of average travel time of both rail and barge connections (table 2). However, the decrease from 2017 to 2018 can only by explained by the decreasing average travel time of barge services (table 2); all the other variables of which the average service throughput time is constructed are moving in the direction of longer shorter paths. It is though unlikely the relative small number of barge connections compared to rail connections result in the significant drop (-1.690) in average shortest path length. Especially since rail services are on average faster, so there is little chance the decreasing average travel time of barge services is resulting in a faster path than a rail service. The developments causing the shorter paths are therefore probably on specific links or corridors which have a large influence in the network connectivity.

Weighted network measures	**2016**	**2017**	**2018**
Weighted assortativity coefficient	-0.262	-0.256	-0.236
Weighted rich-club coefficient [degree ≥15]*	0.953	0.955	0.982
Weighted rich-club coefficient [degree ≥60]*	0.479	0.412	0.454
Gini-index [weight distribution]	0.610	0.750	0.742
Average weighted degree centrality*	15.434	15.627	16.354
Average weighted shortest path length*	9.644	10.961	9.271
Average weighted betweenness centrality*	958	961	872
Average weighed nearest neighbour degree	23.859	25.960	28.304
Number of communities	11	13	15

*Table 8: Weighted network and node level measures. Extracted from Gephi and computed in R using the tnet package, and Python using NetworkX. *Service throughput times on parallel edges are averaged and recorded as one link.*

4.2.2 Other weighted network measures

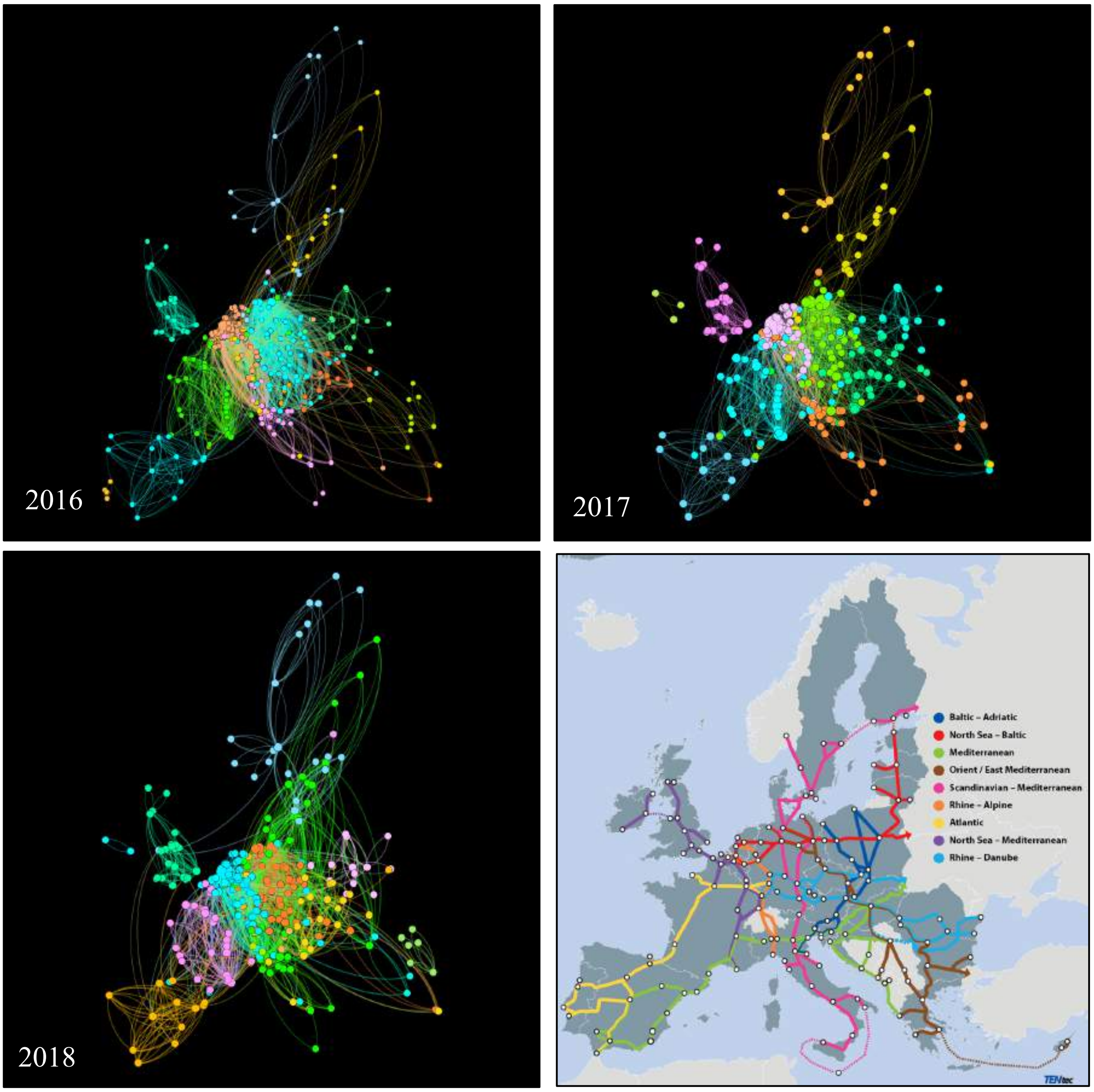
The weighted assortativity is based on the weighted degree centrality of nodes. This is slightly different from the Pearson correlation of unweighted graphs since it does take the number of edges a node has into account. The resulting values in table 8 are close to zero and therefore there is only a weak negative linear relationship. So, nodes with a high average weighted degree centrality do still have slightly more chance to be connected to low average weighted degree centrality nodes (and vice versa). The Pearson correlation for unweighted graphs is lower because low degree nodes with slow incoming and outgoing average service throughputs times are assigned a higher weighted degree centrality compared to the unweighted graph. Thus, more nodes in the network will hold higher weighted degree centrality values. As the values of the assortativity coefficient are stable over time, the average weighted degree centrality of nodes is not changing much; and this is indeed the case as shown in table 8.

Just like the assortativity coefficient, the rich-club is based on the degree centralities which do not take the number of links a node has into account. Different degree thresholds are used for the rich-club analysis compared to the unweighted graph since the weighted degree centralities have higher values. The thresholds are set by multiplying the average service throughput time over three years (3.2) with the degree thresholds from the unweighted graph. As a consequence of the stable average weighted degree centrality, the rich-club coefficient is so as well. The combination of the two rich-club values for each year show the same development as the unweighted graphs: the connectivity between high degree nodes is decreasing and the connectivity between low and high degree nodes is increasing. Since the Gini-index is not based on the average service throughput time of all parallel edges, but on the sum of all parallel edges, nodes with more services have higher degree values than nodes with an equal number of edges but a smaller number of services (assuming all services have the same service throughput time for a moment). The increasing Gini-index indicates more inequality, so a larger number of services belong to a smaller number of nodes. Calculations show 20 out the 330 nodes in 2018 not only hold 50 percent of all services, but these services are also responsible for 50 percent of the total amount of service throughput time in the network.

4.2.3 Communities

A quick glance on the tables containing the communities (appendices 11-13) reveal the extra communities are mainly small ones; especially in 2018. The communities in the unweighted network, which were often just consisting of two or three countries, do barely exist anymore. The new large communities rely on fast average service throughput times for their interconnectivity: otherwise the nodes located far apart would not be in the same community. Just like the unweighted network, the community creation in the weighted network does not make the network a small-world: the clustering coefficients are too low, and the diameter is too large. Worthy remarking is the absence of an increase in the average shortest path lengths although the communities have become more geographically scattered and the diameter has increased. It indicates services with a low throughput time have the ability to effectively form communities with nodes far apart from one another; but still with a limited number of nodes in between.

An example is the spread of nodes belonging primarily to The Netherlands all the way to Northern Italy (orange in figure 10.1). These services belong to a rail corridor which is part of Trans-European Transport Network (TEN-T) from Rotterdam (The Netherlands) to the Italian cities of Milan and Genoa. TEN-T has in total nine of these corridors (figure 11) and multiple of them are visible in the visualisations of the community structures: Atlantic, North Sea-Baltic, Scandinavian-Mediterranean, Baltic-Adriatic, Rhine-Alpine. The corridors are less visible when the communities are established without link weights; pointing out service throughput time is a useful link weight for detecting communities and specifically corridors.



10 (.1, .2, .3): The community structures with service throughput time as link weight for 2016-2018 with a geographical lay-out. The size of a node represents the number of connections and the thickness of an edge represents the number of services. Visualised in Gephi. Figure 11: The TEN-T corridors: Source: Eurostat (2017b).

Relating back to what is observed in the networks without link weights, adding average service throughput time as a link weight is resulting in the same developments regarding network structure; and it is even better suitable for finding communities. The Gini-index shows the number of link weighs is clustering around higher degree nodes, just like the Gini-indices in table 7 have shown low degree nodes are increasing their connectivity to high degree nodes. Weighted centrality measures confirm this trend in both the weighted and unweighted analyses. For establishing communities and finding corridors, service throughput time as a link weight shows the ability to resemble the real-world situation.

4.3 Correlation of Results

So far the network is described based on a list of 15 measures and service throughput time is used as a link weight. The network measures in the weighted network show the same trends as the unweighted network; this indicates the service throughput time is at least not conflicting the observations in network changes. However, in order to test whether the changes in service throughput times are related to the network changes, correlations are calculated between the average service throughput time and the list of 15 measures. The average departures per week and the average travel time are added as control variables as the service throughput time is based on these two variables. The average distance is added to check whether using it can be interesting for a follow up research.

Table 9 contains all the results and shows five of the 15 measures have a strong correlation (≥0.9) with the service throughput time. Another seven show a high correlation (≥0.7). The hierarchy has a strong positive correlation with the average service throughput time (0.972): so, if the service throughput time is higher, the high degree nodes increase in importance. This can be attributed to the fact high degree nodes have links reaching in the far hinterlands, which take more time to reach and have less departures per week. If the service throughput time is higher, it could mean more low degree nodes ,which are located farther away, are connecting to the high degree nodes. For the same reason there is high negative correlation (-0.970) with the assortativity coefficient: a slower service throughput time will result in more low degree nodes connecting to high degree nodes, thus decreasing the assortativity. A decreasing gamma index, although only minimal, is highly correlated (-0.884) with the service throughput time. The level of significance is not good, but the developments do make sense as more direct connections from large degree nodes to low degree nodes result in less connections between low degree nodes and former intermediary hubs. The high correlation (-0.991) between the average service throughput time and the clustering coefficient supports this: there are less closed triplets, so more nodes are only reached from one point.

As the rich-club coefficient and the Gini-index [degree centrality] are based on the degree centrality (with which there is indeed a strong correlation), it is more helpful to see if the service throughput time is correlating with the average degree centrality. The correlation between the two is -0.749, indicating a high chance a slower service throughput time is resulting in a higher degree centrality. However, as both the increase in service throughput time and degree centrality are minimal in table 7, the significance of this result is not good. The logic a slower service throughput time is resulting in more services to low degree nodes, and thus increasing the average degree, is nevertheless still there.

Not completely remarkable is the two variables from which the service throughput time is constructed are also showing high and strong correlations with the network measures. The average travel time also has strong and high correlations with 12 out of the 15 measures, just like the average service throughput time. The average number of departures per week has eight strong correlations, three more than the average service throughput time, but five less in total. The strong correlations between the underlying variables of the service throughput time give an indication both of variables are of importance in the aggregated measure. Average distance shows some high and strong correlations with measures related to the shortest paths: as the average distance increases, the shortest paths decrease in number of nodes to cross. This can point towards the use of more direct links.

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

		Average Service Throughput	Hierarchy	Assortativity coefficient	Average shortest path length	Gamma index	Rich-club coefficient [≥5]	Rich-club coefficient [≥20]	Gini-index [degree centrality]	Gini-index [edge distribution]	Average degree centrality	Average betweenness centrality	Average closeness centrality	Average clustering coefficient	Average nearest neighbour degree	Network diameter	Number of communities	Average distance	Average travel time	Average departures per week
Average Service Throughput	**Pearson Cor.**	1																			
	Sig. (2-tailed)																				
Hierarchy	**Pearson Cor.**	**0.972**	1																		
	Sig. (2-tailed)	0.150																			
Assortativity coefficient	**Pearson Cor.**	**-0.970**	**-0.885**	1																	
	Sig. (2-tailed)	0.158	0.308																		
Average shortest path length	**Pearson Cor.**	-0.570	-0.746	0.352	1																
	Sig. (2-tailed)	0.614	0.463	0.771																	
Gamma index	**Pearson Cor.**	**-0.884**	-0.750	**0.972**	0.120	1															
	Sig. (2-tailed)	0.310	0.460	0.152	0.924																
Rich-club [≥5]	**Pearson Cor.**	**0.938**	**0.831**	**-0.994**	-0.250	**-0.991**	1														
	Sig. (2-tailed)	0.226	0.376	0.068	0.839	0.084															
Rich-club [≥20]	**Pearson Cor.**	-0.131	-0.360	-0.115	**0.889**	-0.347	0.221	1													
	Sig. (2-tailed)	0.916	0.766	0.926	0.302	0.774	0.858														
Gini-index [degree centrality]	**Pearson Cor.**	**0.974**	**1.000****	**-0.890**	-0.741	-0.756	**0.836**	-0.351	1												
	Sig. (2-tailed)	0.145	0.006	0.302	0.469	0.454	0.370	0.772													
Gini-index [edge distribution]	**Pearson Cor.**	0.746	0.569	**-0.886**	0.122	**-0.971**	**0.931**	0.563	0.577	1											
	Sig. (2-tailed)	0.464	0.614	0.307	0.922	0.154	0.239	0.620	0.609												
Average degree centrality	**Pearson Cor.**	-0.749	-0.574	**0.889**	-0.117	**0.972**	**-0.932**	-0.558	-0.581	**-1.000****	1										
	Sig. (2-tailed)	0.461	0.611	0.303	0.925	0.151	0.235	0.623	0.605	0.003											
Average betweenness centrality	**Pearson Cor.**	-0.533	-0.716	0.310	**0.999***	0.076	-0.207	**0.909**	-0.710	0.166	-0.161	1									
	Sig. (2-tailed)	0.642	0.492	0.799	0.028	0.952	0.867	0.274	0.497	0.894	0.897										
Average closeness centrality	**Pearson Cor.**	0.792	**0.913**	-0.619	**-0.953**	-0.415	0.531	-0.709	**0.909**	0.184	-0.189	**-0.939**	1								
	Sig. (2-tailed)	0.418	0.268	0.575	0.196	0.728	0.643	0.498	0.273	0.882	0.879	0.224									
Average clustering coefficient	**Pearson Cor.**	**-0.991**	**-0.931**	**0.994**	0.453	**0.939**	**-0.977**	-0.005	**-0.934**	**-0.830**	**0.833**	0.413	-0.701	1							
	Sig. (2-tailed)	0.087	0.237	0.070	0.701	0.223	0.138	0.997	0.232	0.377	0.374	0.729	0.505								
Average nearest neighbour degree	**Pearson Cor.**	0.703	0.518	**-0.856**	0.183	**-0.954**	**0.906**	0.612	0.525	**0.998***	**-0.998***	0.226	0.123	-0.794	1						
	Sig. (2-tailed)	0.503	0.654	0.346	0.883	0.193	0.278	0.581	0.648	0.039	0.043	0.855	0.921	0.416							
Network diameter	**Pearson Cor.**	0.037	-0.198	-0.281	**0.800**	-0.500	0.381	**0.986**	-0.189	0.693	-0.689	**0.826**	-0.581	-0.173	0.736	1					
	Sig. (2-tailed)	0.976	0.873	0.819	0.410	0.667	0.751	0.107	0.879	0.512	0.516	0.382	0.606	0.889	0.473						
Number of communities	**Pearson Cor.**	**0.847**	**0.948**	-0.691	**-0.920**	-0.500	0.610	-0.638	**0.945**	0.277	-0.283	**-0.901**	**0.995**	-0.766	0.218	-0.500	1				
	Sig. (2-tailed)	0.357	0.207	0.514	0.257	0.667	0.582	0.559	0.212	0.821	0.818	0.285	0.061	0.444	0.860	0.667					
Average distance	**Pearson Cor.**	0.709	**0.854**	-0.514	**-0.984**	-0.296	0.420	-0.793	**0.849**	0.058	-0.064	**-0.975**	**0.992**	-0.605	-0.003	-0.679	**0.975**	1			
	Sig. (2-tailed)	0.499	0.349	0.656	0.115	0.809	0.724	0.417	0.354	0.963	0.960	0.143	0.081	0.586	0.998	0.525	0.142				
Average travel time	**Pearson Cor.**	**0.992**	**0.936**	**-0.992**	-0.464	**-0.935**	**0.974**	-0.007	**0.939**	**0.823**	**-0.826**	-0.424	0.710	**-1.000****	0.786	0.161	0.774	0.615	1		
	Sig. (2-tailed)	0.079	0.229	0.078	0.693	0.231	0.146	0.995	0.224	0.385	0.382	0.721	0.497	0.008	0.424	0.897	0.436	0.578			
Average departures per week	**Pearson Cor.**	**-0.991**	**-0.995**	**0.927**	0.677	**0.812**	**-0.882**	0.266	**-0.996**	-0.648	0.652	0.644	**-0.868**	**0.963**	-0.60	0.100	**-0.912**	-0.798	**-0.966**	1	
	Sig. (2-tailed)	0.087	0.063	0.245	0.527	0.397	0.313	0.829	0.057	0.551	0.548	0.555	0.331	0.174	0.590	0.936	0.270	0.412	0.166		

Table 9: Correlations between the average service throughput time and all network measures for 2016-2018 (N=3). Correlations above 0.8 are in bold for more clarity. Computed in IBM SPSS.

4.4 Differences Between Barge and Train

It is possible a change in the service throughput time for barge and rail services is of different influence on the networks because of the fundamental differences between barge and rail transport: barges are slower, train are more used for longer transport, and it is easier to construct a railroad than a canal; so, barges are more restricted to corridors. Therefore, in the final part of the analysis, the networks are split up for each modality.

4.4.1 The rail network

When the networks are separated, and the average service throughput times are calculated for each network (table 10), both rail and barge services are increasing in service throughput time and are thus responsible for the increase in average service throughput time on an aggregated level. Because of the lower number of barge services compared to rail services, the larger increase of service throughput time for barges is not proportionally reflected in the aggregated network measure. Splitting the network for the two modalities is therefore useful.

Service throughput time measures	**2016**	**2017**	**2018**
Average service throughput time	3.1	3.2	3.3
Average service throughput time rail	2.9	2.9	3.0
Average departure per week	*4.8*	*4.8*	*4.7*
Average travel time	*1.8*	*1.9*	*1.9*
Average service throughput time barge	3.8	4.1	4.3
Average departure per week	*3.2*	*3.1*	*2.8*
Average travel time	*2.4*	*2.6*	*2.5*

Table 10: Service throughput times for all datasets.

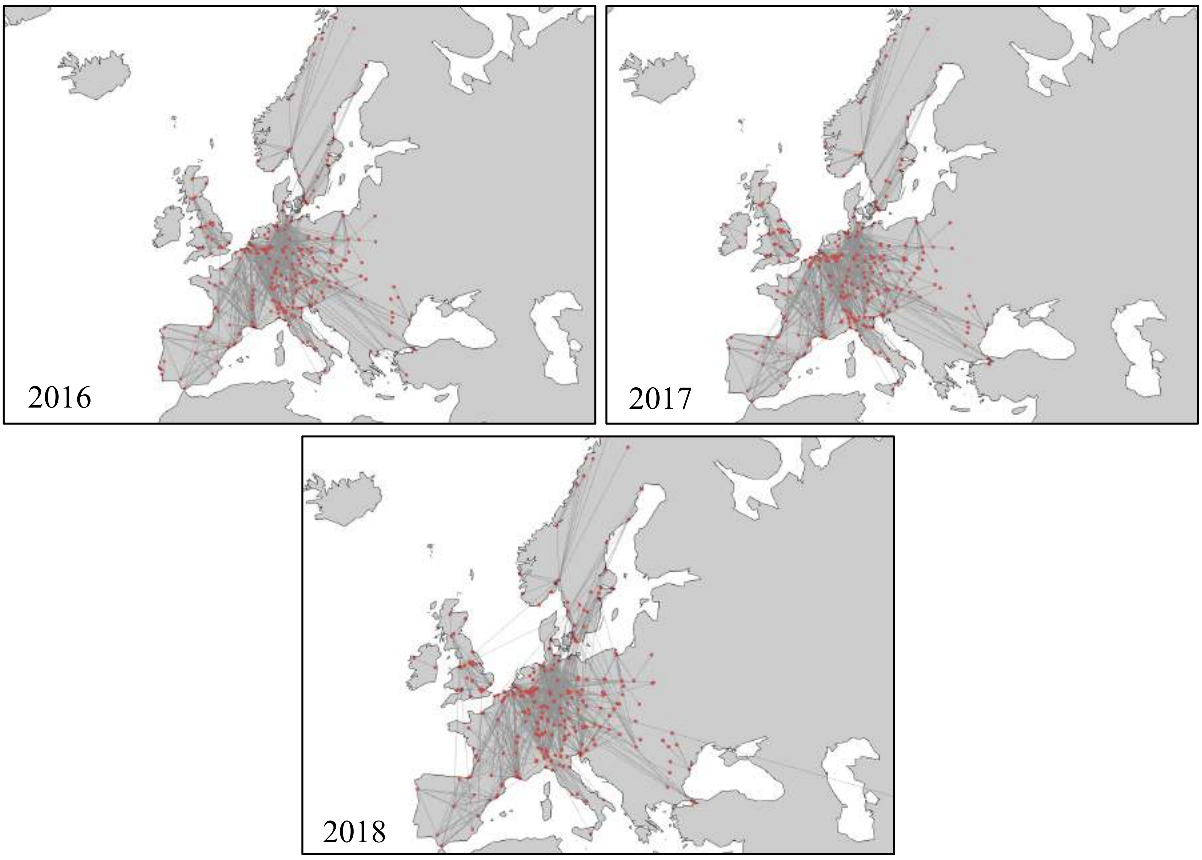
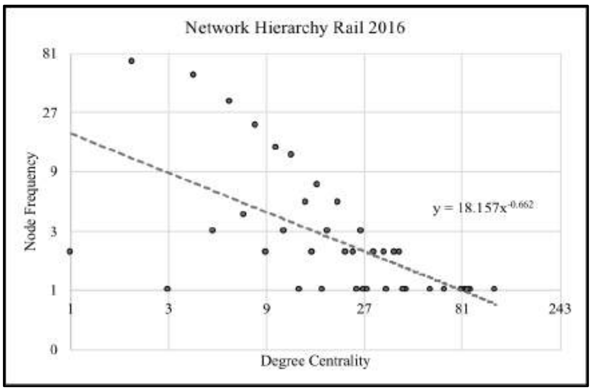
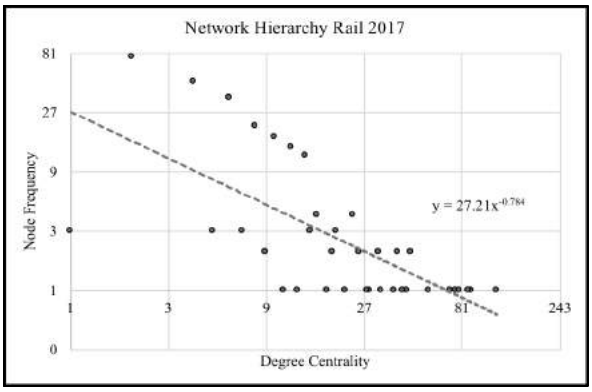


Figure 12 (.1, .2, .3): The complete network of rail connections. Created with Tulip.

From the visuals in figure 12 there is (again) little change visible; this matches with the stable number of rail services displayed in table 2. Visualisation in appendix 14 show Rotterdam is for a change not the most important European node: Hamburg is dominating the rail market and other German nodes such as Bremerhaven and Duisburg show good connectivity as well. The rail service market is primarily focussed on The Netherlands, Belgium, Northern-Italy, and Germany and the number of services is increasing between these regions as indicated by the change of thickness and colour of the edges. There is however also a change visible from the large degree nodes to the lower degree nodes in their proximity: these edges are becoming thicker are more orange/red, which means an increase in services. Both the edges between large degree nodes, and between large degree nodes and lower degree nodes are thus increasing while the number of services (table 2) is actually decreasing between 2016 and 2018. It could indicate services are more offered over certain corridors because this would increase the number of connections to a particular set of high degree nodes (hubs) and decrease the number of connections to small degree nodes.

As the number of services is slightly lower in 2018 compared to 2016, some of the measures in table 11 are returning to their former values after the increase of services in 2017. However, some measures end up higher or lower than before. The main changes for rail services table 2 presents are the increase in average distance, the increase in travel time, and the decrease in the number of weekly departures. Just like the complete networks, the rail networks are not showing signs of being scale-free (hierarchy <1, figure 13); the increasing hierarchy is however indicating the network is centralising. This supports the change of colours and the thickening of edges around high degree nodes visualised in appendix 14. The stable gamma index means





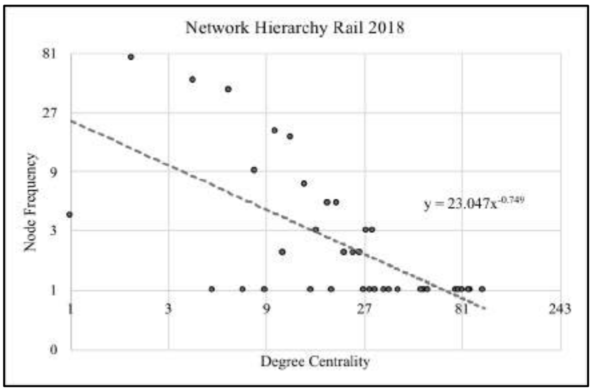


Figure 13 (.1, .2, .3): The exponent of the slope of the power-law line indicates whether the network is becoming more (1) or less (0) centralised. Computed in Gephi and visualised in Microsoft Excel.

there are not many new connections arising between nodes which are not already connected; in line with the stable and even decreasing number of services from table 2. A decreasing weak to moderate negative linear relationship of the assortativity coefficient also indicates there is a larger chance high degree nodes are being connected to lower degree nodes; signalling centralisation.

The use of intermediary nodes and centralisation in the rail network is also supported by the other measures in table 11. As the interconnectivity of higher degree nodes is increasing, the rich-club for both [>5] and [>20] is so as well. The decrease of rich-club [>20] to its former value in 2018 (compared to 2016) can be attributed to an increase of connections to intermediary hubs with a degree lower than 20. This would shift some of the edges away from the rich-club [≥20] but still measure an increase in the rich-club [≥5]. In a real-world hinterland structure this would look like a corridor because cities with many unique connections are foremost making connections with intermediary cities, and these intermediary cities have the services for the distribution of containers to most of the small cities. The slight increase in Gini-index is supporting this reasoning since it points towards centralisation while the number of services has actually decreased; so more of the remaining services remain on the edges leading from or towards high degree nodes.

Network and node level measures rail	**2016**	**2017**	**2018**
Hierarchy	0.662	0.784	0.749
Assortativity coefficient	-0.166	-0.257	-0.253
Average shortest path length	4.044	4.101	3.887
Gamma index	0.019	0.018	0.019
Rich-club coefficient [degree ≥5]	0.075	0.095	0.097
Rich-club coefficient [degree ≥20]	0.418	0.485	0.419
Gini-index [degree centrality]	0.550	0.552	0.557
Average degree centrality	5.090	5.000	5.004
Average betweenness centrality	762	782	692
Average closeness centrality	0.274	0.274	0.294
Average clustering coefficient	0.277	0.243	0.241
Average nearest neighbour degree	27.370	29.800	27.904
Network diameter	11	12	11
Number of communities	14	15	12

Table 11: Network and node level measures for rail transport. Extracted from Tulip, Gephi, and computed in Microsoft Excel, brainGraph for R, and NetworkX for Python.

Therefore, the average shortest path length is also decreasing: the increasing number of hubs shorten connections between small nodes normally displaying high average shortest path values. Consequently, the betweenness centrality will go down as less shortest paths will go through the high degree nodes and more through intermediary nodes acting as hubs. When the betweenness of high degree nodes is decreasing proportionally faster than the increase of the betweenness of the intermediary hubs, the average betweenness centrality on a network level will go down. Finally, an increase in the closeness centrality is expected, and observed, since the average shortest path length is decreasing.

Appendix 15 visualises the rail network with the links weights. As the networks are drawn with a force-directed algorithm, the distance of nodes on the graph is representing the number of services between them. Antwerp is located close to some large French cities, and it is also clear Rotterdam and Duisburg have strong connections with each other. In 2017 and 2018, the size of smaller nodes located just around the centre of the graph is growing, and the graph is showing more brown edges. The combination of these two observations signals the number of services to middle degree nodes are increasing. The indication faster service often belong to high degree nodes is supported by the increasing weighted assortativity index. Although it is again based on the sum of the average throughputs from a node, and not on the number of links, it signals nodes with a low average service throughput time are connecting more to high average service throughput time nodes.

Weighted network measures rail	**2016**	**2017**	**2018**
Weighted assortativity coefficient	-0.174	-0.183	-0.228
Weighted rich-club coefficient [degree ≥15]*	0.980	0.984	0.981
Weighted rich-club coefficient [degree ≥60]*	0.520	0.463	0.474
Gini-index [weight distribution]	0.582	0.661	0.658
Average weighted degree centrality*	14.896	14.990	15.380
Average weighted shortest path length*	10.185	11.380	9.410
Average weighted betweenness centrality*	853	857	755
Average weighed nearest neighbour degree	10.542	10.753	10.186
Number of communities	14	15	14

*Table 12: Weighted network and node level measures for rail transport. Extracted from Gephi and computed in R using the tnet package, and Python using NetworkX. *Service throughput times on parallel edges are averaged and recorded as one link.*

A slight increase in weighted degree centrality could be the cause of the improved connectivity of middle degree nodes. Unlike the betweenness centrality in the unweighted graph, the addition of extra edges to the middle degree nodes does not influence the degree values of the high degree nodes; they will, due to their high number of services, probably still have a service running on a unique link although the number of services might have decreased. If the average weighted degree centrality is divided by the average service throughput time for rail services, the measures are similar to the unweighted graph; 5.137 (14.896/2.9) for 2016, 5.169 for 2017, and 5.127 for 2018. Placing link weights on the graph results in weighted betweenness values higher than the unweighted betweenness values; indicating the fastest paths do go through a particular set of nodes, creating a high average. The differences between the weighted and unweighted measure are however decreasing (91, 75, 63), signalling the set of nodes holding fast services is increasing. This confirms a trend of increased importance of intermediary nodes. The decreasing average shortest path length supports this as well: nodes located far away from one another can reach each other faster because intermediary nodes with fast services are increasing the interconnectivity of the network. Finally, the rich-club coefficients are not showing much variation because the average weighted degree is relatively stable, but it does show the same trend as in the unweighted graph: nodes with a degree under 20, intermediary nodes, are increasing in connectivity. The Gini-index moving towards centralisation supports this again.

4.4.2 The barge network

The final network under review is the one from the inland shipping services by barge. The number of services included in the Intermodal Links database has increased significantly over the course of three years: 175 extra in 2017 and 74 extra in 2018. The average distance is first increasing from 198 km in 2016 to 368 km in 2017 and then decreasing to 201 km in 2018. Although a lot of services have been added, the network is not visually changing as figure 14 shows. This is reasonably logical as the barges are restricted to rivers and canals. The gamma index in table 13 also indicated this: it is stable, so few new unique connections arise while extra services are added. Rotterdam and Antwerp are by far the most important spots where barges sail from or to and this is also not changing with the addition of extra services (appendix 14). Other cities with a relative large number of connections are Terneuzen (The Netherlands), Zeebrugge (Belgium), Hamburg (Germany), Strasbourg, and Le Havre (both France).

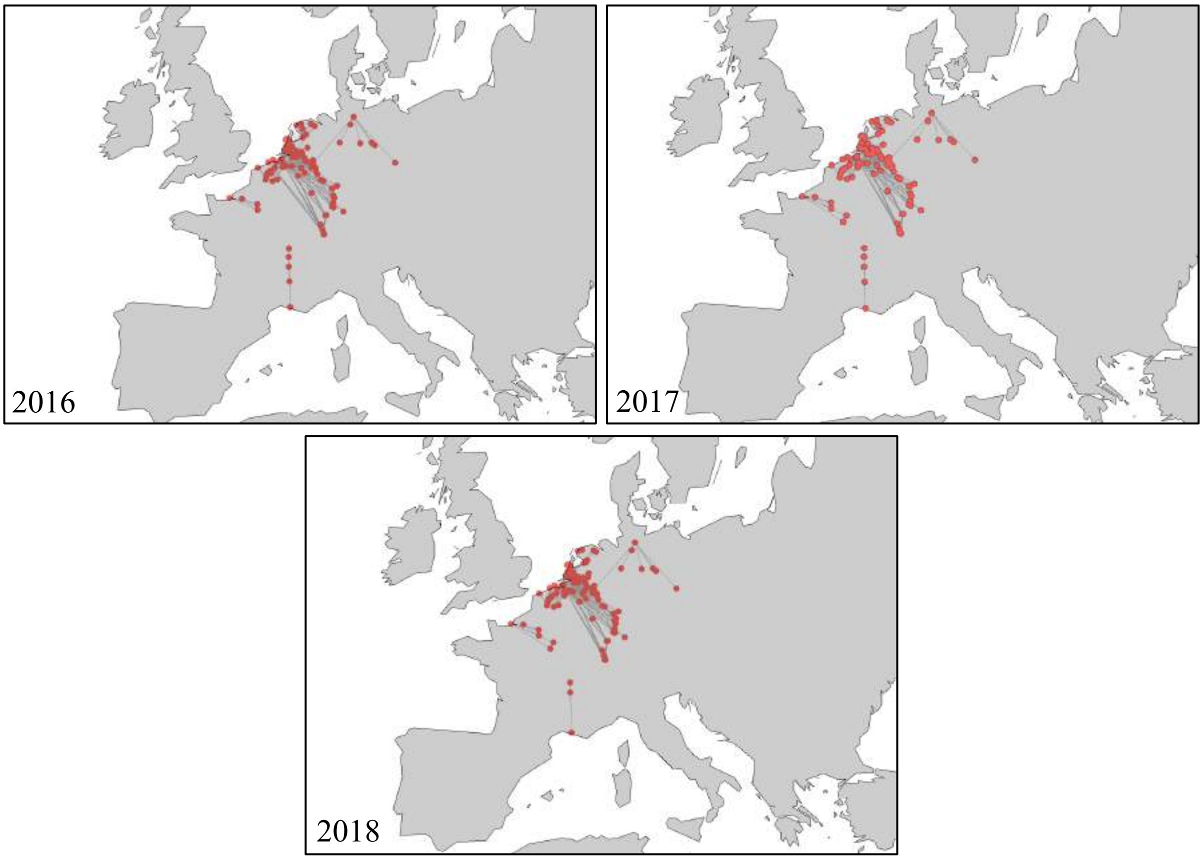


Figure 14 (.1, .2, .3): The complete network of barge connections. Created with Tulip.

With the addition of extra barge services, the hierarchy is increasing towards a more centralised, yet not scale-free, network (figure 15). In order to become a scale-free network, the nodes currently having a relative large degree (located between eight and 27) compared to all the small degree nodes will have to increase their number of connections, or nodes with low degrees have to increase their connectivity to other nodes. Both are not likely to happen due to the nature of the barge network; it is restricted to the rivers and canals. Compared to the complete network and the separated rail network, the assortativity coefficient displays a strong negative linear relationship. Hence, if a high degree node is selected, the chance is significant a randomly picked node to which it is connected is of a small degree (and vice versa). The huge difference between the average nearest neighbour degree and the average degree centrality shows the similar trend: most nodes connect to a large degree node.

Since extra services are primarily added on already existing links, the average shortest path length is not changing significantly. As the closeness centrality is based on the shortest paths and the number of nodes in the network, both of which are stable, it is stable as well. For the rich-club analysis in the barge network the degree thresholds are lower than the rail or complete network since the average degree centrality and the maximum degree centrality are lower compared to them. The rich-club coefficients show the large increase in connections in 2017 is creating some new unique connections between both along high and low degree nodes. However, as the Gini index is increasing, resulting in more inequality of unique edge distribution, the share of extra connections belongs primarily to high degree nodes.

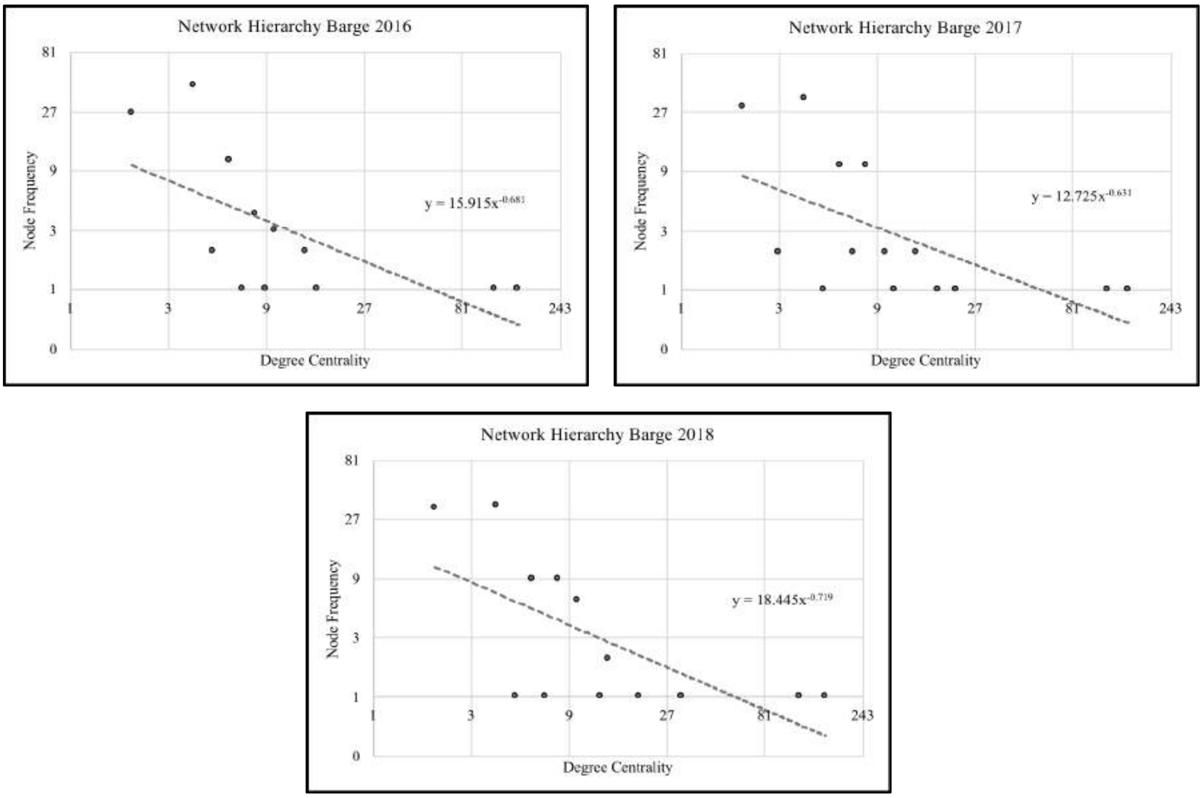


Figure 15 (.1, .2, .3): The exponent of the slope of the power-law line indicates whether the network is becoming more (1) or less (0) centralised. Computed in Gephi and visualised in Microsoft Excel.

Network and node level measures barge	2016	2017	2018
Hierarchy	0.681	0.631	0.719
Assortativity coefficient	-0.646	-0.637	-0.647
Average shortest path length	1.990	1.976	1.979
Gamma index	0.037	0.038	0.039
Rich-club coefficient [degree ≥3]	0.099	0.218	0.228
Rich-club coefficient [degree ≥8]	0.448	0.559	0.579
Gini-index [degree centrality]	0.536	0.554	0.569
Average degree centrality	3.588	3.717	3.851
Average betweenness centrality	64	62	68
Average closeness centrality	0.534	0.524	0.534
Average clustering coefficient	0.590	0.578	0.542
Average nearest neighbour degree	72.216	131.901	122.077
Network diameter	4	3	3
Number of communities	10	9	9

Table 13: Network and node level measures for barge transport. Extracted from Tulip, Gephi, and computed in Microsoft Excel, brainGraph for R, and NetworkX for Python.

As the average service throughput time is increasing for barge services, and the number of unique connections is fairly stable, the weighted network measures will show a similar increase. This steady rise of service throughput time can directly be translated to the increases in the weighted degree centrality, weighted shortest path length, and weighted nearest neighbour degree as they are all based on the sum of the service throughput time for a single node. For some other measures the latter is also the case, but they require some extra explanation. The weighted betweenness is for instance increasing because of the centralisation of the network (supported by the increasing Gini-index) around high degree nodes and not necessarily because the average service throughput time is increasing. As the weighted betweenness is higher than the unweighted measure, more shortest paths go through a particular set of high degree nodes rather than following the route with the least nodes in between.

Weighted network measures barge	**2016**	**2017**	**2018**
Weighted assortativity coefficient	-0.674	-0.632	-0.634
Weighted rich-club coefficient [degree ≥12]*	0.779	0.933	0.821
Weighted rich-club coefficient [degree ≥32]*	0.423	0.525	0.406
Gini-index [weight distribution]	0.600	0.726	0.745
Average weighted degree centrality*	12.892	13.460	14.248
Average weighted shortest path length*	5.795	6.378	6.906
Average weighted betweenness centrality*	70	71	80
Average weighed nearest neighbour degree	30.448	36.260	42.981
Number of communities	9	8	9

*Table 14: Weighted network and node level measures for barge transport. Extracted from Gephi and computed in R using the tnet package, and Python with NetworkX. *Service throughput times on parallel edges are averaged and recorded as one link.*

The weighted assortativity coefficient is relatively stable since the average weighted degree is proportionally increasing for all the nodes in the network due to an increasing average service throughput time. The strong negative linear relationship between a high degree node connecting to a low degree node is therefore sustained. For the rich-club coefficient new degree thresholds are set by multiplying the average service throughput time for barges by the thresholds from the unweighted rich-club measure. The increase of both the rich-club measures in 2017 is caused by the offering of new services between primarily high degree nodes; this is supported by the simultaneous increase in Gini-index. The subsequent decrease of rich-club coefficients in 2018 could be caused by a faster service throughput time between a particular set of nodes; this would result in a lower average weighted degree. So, some nodes from the [≥32] category end up in the [≥12] category, and some nodes get a weighted degree [<12]. This line of thought is supported by the weighted assortativity index and the Gini-index indicated the weights are not shifting between high and low degree nodes.

4.5 *Correlation of Results*

4.5.1 The rail network

The average service throughput time correlations for the rail network are in general less strong than for the complete network (table 15). However, the measures which do have a strong (≥0.9) or high (≥0.7) correlation are primarily the same. The assortativity coefficients show a high negative correlation (-0.821), which can mean increasing service throughput times do indeed connect higher degree nodes to lower degree nodes. This is supported by the high correlations in degree (0.843), closeness (0.888), clustering (0.868), and a strong correlation in the Gini-index (0.981). All of them point towards a trend of centralisation and an increasing connectivity between high and low degree nodes.

Similar to the complete network, the average travel time and average number of departures show strong correlations with the network measures; in this case even better ones than the average service throughput time. Both of them correlate high or strong with nine out the 14 measures. This is the same number as for the average service throughput time, but the significances are better for the average number of weekly departures. The average distance shows similar strong correlations to the shortest path lengths compared to the correlations for the complete network.

4.5.2 The barge network

The last sub-network up for analysis is the barge network: by far the smallest network in terms of number of connections, number of nodes, and geographical spread. In combination with significant changes in average service throughput time and its underlying variables, it is therefore expected to show the best correlations of all three networks. Out of the 14 measures, ten show a high to strong correlation with the average service throughput time (table 16). Remarkable is the strong positive correlation (0.958) with the rich-club [>20], as this measure shows low correlations for the complete and rail network. A cause of this could be the relative stability of the set of nodes belonging to this degree threshold whereas this set seemed to fluctuate for the rail network as the number of intermediary nodes is increasing there. It could also explain the strong correlations with others measure based on the degree centrality; 0.934 for the rich-club [≥5] and 0.999 for the Gini-index. The correlation with the degree centrality is 0.995, so higher degree values for more connected nodes could be the result of an increased average service throughput time.

High and strong negative correlations are detected between the average number of departures per week and two of the centrality measures (-0.944 for degree, -0.872 for betweenness). As the frequency of departures steadily decreases, the network is becoming more central. This is supported by the strong positive correlation (0.998) between the average number of weekly departures and the clustering coefficient. The latter is decreasing, thus there are less closed triplets in the network, indicating most nodes only have services from or to one other (high degree) node and not to other (low degree) nodes. This is some way embedded in the barge network as the rivers are connecting specific cities and do not offer the possibility to easily create services to other cities out of their river system. Average distance shows low correlations compared to the other two networks. This raises the question whether it is only primarily useful for explaining changes in the rail network.

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Rail network

		Average Service Throughput	Hierarchy	Assortativity coefficient	Average shortest path length	Gamma index	Rich-club coefficient [≥5]	Rich-club coefficient [≥20]	Gini-index [degree centrality]	Average degree centrality	Average betweenness centrality	Average closeness centrality	Average clustering coefficient	Average nearest neighbour degree	Network diameter	Number of communities	Average distance	Average travel time	Average departures per week
Average Service Throughput	**Pearson Cor.**	1																	
	Sig. (2-tailed)																		
Hierarchy	**Pearson Cor.**	0.659	1																
	Sig. (2-tailed)	0.542																	
Assortativity coefficient	**Pearson Cor.**	**-0.821**	**-0.971**	1															
	Sig. (2-tailed)	0.387	0.155																
Average shortest path length	**Pearson Cor.**	-0.740	0.019	0.223	1														
	Sig. (2-tailed)	0.470	0.988	0.857															
Gamma index	**Pearson Cor.**	0.046	-0.721	0.533	-0.706	1													
	Sig. (2-tailed)	0.971	0.487	0.642	0.501														
Rich-club [≥5]	**Pearson Cor.**	**0.884**	**0.934**	**-0.993**	-0.339	-0.427	1												
	Sig. (2-tailed)	0.310	0.232	0.077	0.780	0.719													
Rich-club [≥20]	**Pearson Cor.**	-0.033	0.730	-0.544	0.697	**-1.000****	0.439	1											
	Sig. (2-tailed)	0.979	0.479	0.634	0.509	0.008	0.711												
Gini-index [degree centrality]	**Pearson Cor.**	**0.981**	0.499	-0.693	**-0.857**	0.240	0.775	-0.228	1										
	Sig. (2-tailed)	0.125	0.667	0.513	0.344	0.846	0.435	0.854											
Average degree centrality	**Pearson Cor.**	**0.843**	**0.960**	**-0.999***	-0.262	-0.499	**0.997**	0.510	0.721	1									
	Sig. (2-tailed)	0.362	0.180	0.026	0.831	0.667	0.052	0.659	0.487										
Average betweenness centrality	**Pearson Cor.**	-0.771	-0.028	0.268	**0.999***	-0.672	-0.383	0.662	**-0.880**	-0.307	1								
	Sig. (2-tailed)	0.440	0.982	0.827	0.030	0.531	0.750	0.539	0.315	0.802									
Average closeness centrality	**Pearson Cor.**	**0.888**	0.239	-0.466	**-0.966**	0.500	0.569	-0.489	**0.961**	0.501	**-0.977**	1							
	Sig. (2-tailed)	0.304	0.846	0.691	0.166	0.667	0.614	0.675	0.179	0.666	0.136								
Average clustering coefficient	**Pearson Cor.**	**-0.868**	**-0.945**	**0.996**	0.308	0.457	**-.999***	-0.468	-0.754	**-0.999***	0.352	-0.542	1						
	Sig. (2-tailed)	0.331	0.211	0.056	0.801	0.698	0.021	0.690	0.456	0.031	0.771	0.635							
Average nearest neighbour degree	**Pearson Cor.**	0.164	**0.850**	-0.698	0.542	**-0.978**	0.607	**0.981**	-0.032	0.669	0.502	-0.308	-0.633	1					
	Sig. (2-tailed)	0.895	0.353	0.508	0.635	0.134	0.585	0.126	0.980	0.533	0.665	0.801	0.564						
Network diameter	**Pearson Cor.**	-0.046	0.721	-0.533	0.706	**-1.000****	0.427	**1.000****	-0.240	0.499	0.672	-0.500	-0.457	**0.978**	1				
	Sig. (2-tailed)	0.971	0.487	0.642	0.501	0	0.719	0.008	0.846	0.667	0.531	0.667	0.698	0.134					
Number of communities	**Pearson Cor.**	-0.689	0.092	0.151	**0.997***	-0.756	-0.269	0.747	**-0.817**	-0.190	**0.993**	**-0.945**	0.237	0.602	0.756	1			
	Sig. (2-tailed)	0.516	0.941	0.904	0.047	0.454	0.827	0.463	0.391	0.878	0.077	0.212	0.847	0.588	0.454				
Average distance	**Pearson Cor.**	**0.901**	0.268	-0.492	**-0.958**	0.474	0.594	-0.462	**0.969**	0.527	**-0.971**	**1.000***	-0.567	-0.279	-0.474	**-0.935**	1		
	Sig. (2-tailed)	0.285	0.827	0.672	0.185	0.686	0.595	0.694	0.160	0.647	0.155	0.019	0.616	0.820	0.686	0.231			
Average travel time	**Pearson Cor.**	**0.992**	0.556	-0.740	**-0.821**	0.175	**0.816**	-0.162	**0.998***	0.766	**-0.847**	**0.940**	-0.796	0.035	-0.175	-0.777	**0.950**	1	
	Sig. (2-tailed)	0.083	0.625	0.470	0.387	0.888	0.393	0.896	0.042	0.445	0.357	0.221	0.414	0.978	0.888	0.434	0.202		
Average departures per week	**Pearson Cor.**	**-1.000***	-0.673	**0.832**	0.727	-0.026	**-0.893**	0.013	**-0.977**	**-0.853**	0.758	**-0.879**	**0.877**	-0.183	0.026	0.674	**-0.893**	**-0.989**	1
	Sig. (2-tailed)	0.012	0.530	0.375	0.482	0.983	0.298	0.991	0.138	0.349	0.452	0.316	0.319	0.883	0.983	0.529	0.297	0.095	

Table 15: Correlations between the average service throughput time and all network measures in the rail network for 2016-2018 (N=3). Correlation above 0.8 are in bold for more clarity. Computed in IBM SPSS.

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Barge network

		Average Service Throughput	Hierarchy	Assortativity coefficient	Average shortest path length	Gamma index	Rich-club coefficient [≥5]	Rich-club coefficient [≥20]	Gini-index [degree centrality]	Average degree centrality	Average betweenness centrality	Average closeness centrality	Average clustering coefficient	Average nearest neighbour degree	Network diameter	Number of communities	Average distance	Average travel time	Average departures per week
Average Service Throughput	**Pearson Cor.**	1																	
	Sig. (2-tailed)																		
Hierarchy	**Pearson Cor.**	0.349	1																
	Sig. (2-tailed)	0.773																	
Assortativity coefficient	**Pearson Cor.**	-0.002	**-0.938**	1															
	Sig. (2-tailed)	0.999	0.225																
Average shortest path length	**Pearson Cor.**	**-0.802**	0.280	-0.595	1														
	Sig. (2-tailed)	0.407	0.820	0.594															
Gamma index	**Pearson Cor.**	**0.996**	0.430	-0.091	-0.746	1													
	Sig. (2-tailed)	0.057	0.717	0.942	0.464														
Rich-club [≥3]	**Pearson Cor.**	**0.934**	-0.009	0.355	**-0.963**	**0.899**	1												
	Sig. (2-tailed)	0.232	0.994	0.769	0.175	0.289													
Rich-club [≥8]	**Pearson Cor.**	**0.958**	0.064	0.286	-0.940	**0.928**	**0.997***	1											
	Sig. (2-tailed)	0.186	0.960	0.815	0.221	0.243	0.046												
Gini-index [degree centrality]	**Pearson Cor.**	**0.999***	0.383	-0.038	-0.780	**0.999***	**0.921**	**0.946**	1										
	Sig. (2-tailed)	0.023	0.750	0.976	0.430	0.033	0.256	0.209											
Average degree centrality	**Pearson Cor.**	**0.995**	0.440	-0.102	-0.739	**1.000****	**0.894**	**0.924**	**0.998***	1									
	Sig. (2-tailed)	0.064	0.710	0.935	0.471	0.007	0.296	0.250	0.040										
Average betweenness centrality	**Pearson Cor.**	0.585	**0.964**	**-0.812**	0.015	0.655	0.257	0.326	0.614	0.663	1								
	Sig. (2-tailed)	0.602	0.171	0.396	0.991	0.546	0.835	0.788	0.579	0.539									
Average closeness centrality	**Pearson Cor.**	-0.089	**0.903**	**-0.996**	0.666	0	-0.438	-0.372	-0.052	0.011	0.756	1							
	Sig. (2-tailed)	0.943	0.283	0.058	0.536	1	0.711	0.757	0.967	0.993	0.454								
Average clustering coefficient	**Pearson Cor.**	**-0.932**	-0.664	0.363	0.532	**-0.961**	-0.742	-0.788	**-0.945**	**-0.964**	**-0.839**	-0.277	1						
	Sig. (2-tailed)	0.236	0.538	0.763	0.643	0.179	0.468	0.422	0.212	0.172	0.367	0.821							
Average nearest neighbour degree	**Pearson Cor.**	**-0.913**	0.064	-0.406	**0.976**	**-0.873**	**-0.998***	**-0.992**	**-0.897**	**-0.868**	-0.203	0.488	0.704	1					
	Sig. (2-tailed)	0.268	0.959	0.734	0.140	0.324	0.035	0.081	0.291	0.331	0.870	0.676	0.503						
Network diameter	**Pearson Cor.**	**-0.907**	0.078	-0.419	**0.979**	**-0.866**	**-0.998***	**-0.990**	**-0.891**	**-0.860**	-0.189	0.500	0.693	**1.000****	1				
	Sig. (2-tailed)	0.277	0.950	0.725	0.130	0.333	0.044	0.091	0.300	0.340	0.879	0.667	0.512	0.009					
Number of communities	**Pearson Cor.**	**-0.907**	0.078	-0.419	**0.979**	**-0.866**	**-0.998***	**-0.990**	**-0.891**	**-0.860**	-0.189	0.500	0.693	**1.000****	**1.000****	1			
	Sig. (2-tailed)	0.277	0.950	0.725	0.130	0.333	0.044	0.091	0.300	0.340	0.879	0.667	0.512	0.009	0				
Average distance	**Pearson Cor.**	0.599	**0.960**	**-0.802**	-0.002	0.667	0.273	0.342	0.627	0.675	**1.000***	0.745	**-0.848**	-0.219	-0.206	-0.206	1		
	Sig. (2-tailed)	0.591	0.182	0.407	0.999	0.535	0.824	0.778	0.568	0.528	0.011	0.465	0.356	0.859	0.868	0.868			
Average travel time	**Pearson Cor.**	**0.819**	-0.252	0.572	**-1.000***	0.765	**0.970**	**0.950**	0.797	0.758	0.013	-0.644	-0.556	**-0.982**	**-0.984**	**-0.984**	0.030	1	
	Sig. (2-tailed)	0.389	0.838	0.612	0.018	0.446	0.157	0.203	0.412	0.453	0.991	0.554	0.625	0.122	0.112	0.112	0.981		
Average departures per week	**Pearson Cor.**	**-0.907**	-0.711	0.423	0.476	**-0.941**	-0.697	-0.747	**-0.922**	**-0.944**	**-0.872**	-0.339	**0.998***	0.656	0.645	0.645	**-0.880**	-0.501	1
	Sig. (2-tailed)	0.277	0.496	0.722	0.684	0.220	0.509	0.463	0.254	0.213	0.325	0.780	0.042	0.545	0.554	0.554	0.314	0.666	

Table 16: Correlations between the average service throughput time and all network measures in the barge network for 2016-2018 (N=3). Correlations above 0.8 are in bold for more clarity. Computed in IBM SPSS.

		Complete	Rail	Barge
Hierarchy	Pearson Cor.	**0.972**	0.659	0.349
	Sig. (2-tailed)	0.150	0.542	0.773
Assortativity coefficient	Pearson Cor.	**-0.970**	**-0.821**	-0.002
	Sig. (2-tailed)	0.158	0.387	0.999
Average shortest path length	Pearson Cor.	-0.570	-0.740	**-0.802**
	Sig. (2-tailed)	0.614	0.470	0.407
Gamma index	Pearson Cor.	**-0.884**	0.046	**0.996**
	Sig. (2-tailed)	0.310	0.971	0.057
Rich-club [≥5]	Pearson Cor.	**0.938**	**0.884**	**0.934**
	Sig. (2-tailed)	0.226	0.310	0.232
Rich-club [≥20]	Pearson Cor.	-0.131	-0.033	**0.958**
	Sig. (2-tailed)	0.916	0.979	0.186
Gini-index [degree centrality]	Pearson Cor.	**0.974**	**0.981**	**0.999***
	Sig. (2-tailed)	0.145	0.125	0.023
Gini-index [edge distribution]	Pearson Cor.	0.746	-	-
	Sig. (2-tailed)	0.464	-	-
Average degree centrality	Pearson Cor.	-0.749	**0.843**	**0.995**
	Sig. (2-tailed)	0.461	0.362	0.064
Average betweenness centrality	Pearson Cor.	-0.533	-0.771	0.585
	Sig. (2-tailed)	0.642	0.440	0.602
Average closeness centrality	Pearson Cor.	0.792	**0.888**	-0.089
	Sig. (2-tailed)	0.418	0.304	0.943
Average clustering coefficient	Pearson Cor.	**-0.991**	**-0.868**	**-0.932**
	Sig. (2-tailed)	0.087	0.331	0.236
Average nearest neighbour degree	Pearson Cor.	0.703	0.164	**-0.913**
	Sig. (2-tailed)	0.503	0.895	0.268
Network diameter	Pearson Cor.	0.037	-0.046	**-0.907**
	Sig. (2-tailed)	0.976	0.971	0.277
Number of communities	Pearson Cor.	**0.847**	-0.689	**-0.907**
	Sig. (2-tailed)	0.357	0.516	0.277
Average distance	Pearson Cor.	0.709	**0.901**	0.599
	Sig. (2-tailed)	0.499	0.285	0.591
Average travel time	Pearson Cor.	**0.992**	**0.992**	**0.819**
	Sig. (2-tailed)	0.079	0.083	0.389
Average departures per week	Pearson Cor.	**-0.991**	**-1.000***	**-0.907**
	Sig. (2-tailed)	0.087	0.012	0.277

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 17: Average service throughput time correlations for all networks (N=3). Composed from tables 9, 15, and 16

4.6 Regression analysis

A linear regression analysis is done in IBM SPSS in order to determine whether the established correlations are contributable to changes in the independent variable: the service throughput time. Before doing the linear regression analysis, the curve estimation regression analysis function in IBM SPSS was used to get the best fit with the model. Curve estimations were made for a linear, power, logarithmic, and exponential model. Most variables have the best fit with the linear model. Some did not, but since the differences between the linear and the other models were so minimal, the choice was made to use identical model for all variables.

All squared R values in table 18 which contain an asterisk (*) also show high or strong correlations in table 17. Six out of the 15 variables for the complete network have both a high or strong correlation and a relatively good fit (R2 ≥ 0.75, Sig. ≤ 0.31) with the linear regression model. For the rail network this is four out of 14, and for the barge network it is eight out of 14. However, the significance levels are for most of these values are not near the 95 percent confidence level. Only the Gini-index for the barge network is significant above the 95 percent level. Another three measures reach the 90 percent confidence level and another five the 80 percent level.

	Complete network		**Rail network**		**Barge network**	
Network measure	**R2**	**Sig. [2-t]**	**R2**	**Sig. [2-t]**	**R2**	**Sig. [2-t]**
Hierarchy	**0.945***	**0.150**	0.434	0.542	0.122	0.773
Assortativity coefficient	**0.940***	**0.158**	0.673	0.387	0.000	0.999
Average shortest path length	0.325	0.614	0.547	0.470	0.644	0.407
Gamma index	**0.781***	**0.310**	0.002	0.971	**0.992***	**0.057**
Rich-club [5]/[5]/[3]	**0.880***	**0.226**	**0.781***	**0.310**	**0.873***	**0.232**
Rich-club [20]/[20]/[8]	0.170	0.916	0.001	0.979	**0.917***	**0.186**
Gini-index [degree distribution]	**0.949***	**0.145**	**0.962***	**0.125**	**0.999***	**0.023**
Gini-index [edge distribution]	0.556	0.464	-	-	-	-
Average degree centrality	0.561	0.461	0.710	0.362	**0.990***	**0.064**
Average betweenness centrality	0.284	0.642	0.594	0.440	0.342	0.520
Average closeness centrality	0.628	0.418	**0.789***	**0.304**	**0.869**	**0.236**
Average clustering coefficient	**0.981***	**0.087**	**0.753***	**0.331**	0.008	0.943
Average nearest neighbour degree	0.495	0.503	0.027	0.895	**0.833***	**0.268**
Network diameter	0.001	0.976	0.002	0.971	**0.823***	**0.277**
Communities	0.717	0.357	0.474	0.516	**0.823***	**0.277**
Average distance	0.502	0.499	**0.812***	**0.285**	0.358	0.591
Average travel time	**0.985***	**0.079**	**0.983***	**0.083**	0.670	0.389
Average departures per week	**0.981***	**0.087**	**0.999***	**0.012**	**0.822***	**0.277**

Table 18: : Linear regression analysis for all network. R2 values above 0.75 are in bold for clarity. Computed in IBM SPSS.
** Also has a strong correlation.*

5 Conclusion

The goal of this master thesis is to explain if network structure changes of container transport in the hinterland of Europe can be attributed to changes in service throughput time, and whether the changes are different if the modality used is considered. For this purpose, complex network measures are computed on both weighted and unweighted networks, and correlations between the change in average service throughput time and the computed complex network measures are extracted. Finally, a regression analysis is performed to detect whether the changes are caused by the independent variable, service throughput time.

Over the course of three years the complete hinterland network with both rail and barge connections shows a trend towards an increasing importance around high degree nodes. These high degree nodes act as hubs in a hub-and-spoke structured network for connecting communities consisting of smaller degree nodes. Analysis has showed the number of unique connections between hubs has decreased; this does however not mean the number of services has decreased between these hubs. Since the increase in the total number of services included in the database is larger than the increase in unique connections, new services are primarily added on already existing routes. The decrease in the number of unique connections between highly connected hubs, and the increasing connectivity between hinterland destinations and hubs, shows hubs are serving specific hinterlands: their communities. As these communities are geographically dispersed over Europe, certain corridors, linkable to the TEN-T, are detected. The increasing use of corridors to specific hinterland communities for container transport in Europe would explain why the hubs need less connectivity with one another and connect more with distant, lower in connectivity, cities in the hinterland.

The use of service throughput time as link weight has showed faster shortest paths exist than the one crossing the least number of nodes; this was indicated by the weighted betweenness being higher than the unweighted betweenness. So, there are alternative connections between cities both connected to the same hub, but these connections are slower than a route crossing the hub. This signals the importance of the hubs in the network. The communities connected by the hubs have become more geographically dispersed, but the average shortest path length has not increased. Correlations between the network measures and the service throughput time are strong and cohere with the line of thought for why measures are moving in a certain direction. Six out of the 15 variables in the regression analysis also have a high squared R value. None of the correlations or squared R values are however significant on a 95 percent confidence level. So, while the relations between the service throughput time and the network measures are seemingly logical, it cannot be statistically established whether the service throughput time is a cause for the network structure changes.

Rail transport in the European hinterland has a relativity stable average service throughput time, but changes in the network structure do however take place. These are not significantly reflected in the average service throughput times as the number of connections is large. There seems to be a development in the use of intermediary cities with a significant number of connections; yet not as much as the large hubs have. This can be cities which are positioned on the corridors and act as entry points for hinterland destinations receiving and sending containers through a corridor. If specific routes are examined, the changes in average service throughput time make the formation of corridors in the network more visible. The correlations confirm this for developments on a network level. However, squared R values and a lack of significance on a 95 percent confidence level are, just like the in the complete network, not substantiating enough to prove service throughput time changes are causing the network structure changes for rail transport.

For the barge network the corridor structure was already in place and is dominated by primarily Antwerp and Rotterdam. The extra services included in the database mainly go from these two hubs to cities along the Rhine river in Germany. These relative long routes for the new services explain the significant changes in average service throughput time. Changes in average service throughput time are thus, compared to the rail network, not showing the formation of corridors; these are already in place and do not have the ability to change much since the rivers and canals are effectively already natural corridors on their own. Correlations show service throughput time is correlating with network structure changes around Antwerp, Rotterdam, and some other small cities. The offering of extra services here, not necessarily fast or frequent, on already existing routes increases the centrality. Regression analysis is though only resulting in one squared R value with a confidence level over 95 percent: the rich-club index [≥8]. So, while developments in the barge network are seemingly logical (and correlating) with the average service throughput time, no statistical relationship can be established.

It has proved to be useful to split the network for the modalities used in the European hinterland since the average service throughput time is showing different network structure developments for barge and rail. Eventually, service throughput time is however not able to statistically prove it is of influence in changing the network structure. The main reason for the lack of evidence is probably the limited timespan of the data, as the literature has showed a relationship between time and network structure changes is highly likely.

5.1 Discussion

The case of container transport by barge and rail in the European hinterland was until this thesis not researched with complex network measures on a longitudinal dataset. The complex network measures used are well established (see table 1 for the overview) and result in a high validity of observations and subsequent interpretations. The dataset gives a proper reflection of the real-world hinterland transportation network as the schedules are obtained from the operators and are continuously updated. Ecorys estimates over 90 percent of all scheduled services for container transport by barge and rail are included for Western Europe. These two modalities are the most important scheduled ways of container transport in Europe; transport by truck has a higher market share (Eurostat, 2015) but is more ad-hoc. The percentage of transport done through scheduled versus unscheduled services is not known. However, estimations based on the capacity of trains and barges, service frequency, and the recorded throughput of containers in a certain port, show almost all transport (if the services would actually take place on a fully loaded train or barge) is done through scheduled services (De Langen et al., 2017).

Developments observed in the network over the course of three years seem to for a large extent represent the trends of containerised transport in Europe (Notteboom, 1997 & 2010). For rail transport there is no further concentration of ports but rather a decentralisation of distribution activities to hubs located on corridors. This results in a network structure of hubs serving discontinuous hinterlands (Notteboom & Rodrigue, 2005). Whether this was under the influence of competition between large consortia was not in scope of this research. Policies from (port) authorities and governments were in scope neither, but the growing visibility of the TEN-T corridors suggest EU policy is being executed. The increasing number of containers in hinterland hubs enable operators in the hub to exploit economies of scale in handling the containers and loading full train and barge loads (Notteboom, 2010). The economies of scale therefore increase the competitive position of a hub (Rodrigue et al., 2010) and can subsequently cause network structure changes in their direct neighbourhood (De Langen & Sharypova, 2013). This is in some extent also visible in this research, yet not specifically measured, as there are communities attached to hubs located on the corridors (figure 10).

In the network for barge transportation the trend in containerised transport is slightly the same, although the corridors were already in place. The development of the strong links from hubs such as Antwerp and Rotterdam could be the result of policies stimulating 'greener' types of transport than trucking. Especially in the congested and polluted areas from Rotterdam and Antwerp to the Ruhr area in Germany (and vice versa). Growth in unique connections in the barge network could theoretically happen between nodes currently displaying low degree values. It is however probably not economically viable to establish these links as there currently is a lack of volume to be transported. The use of synchromodal planning tools could create opportunities to reach a cost-effective number of containers to be transported, which would in turn create the possibility to establish new services.

Current literature has addressed the barge and rail network in Europe separately and has in some cases also looked at its development over time. There is however an absence of literature on applying complex network measured to the European hinterland. This thesis is contributing to the existing literature by using throughput time as a link weight in calculating complex network measures for the European network for container transport by barge and rail. This is a novel way of applying a link weight in combination with complex network measures to the European network. Throughput time is correlating with a set of complex network measures, it can however not be established whether the developments are actually caused by changes in throughput time. Nevertheless, it signals throughput time could be of importance and it indicates further research on this is required.

The influence of time of the network structure is of practical relevance for ports designing new (trans)port policies or initiating large infrastructure projects. When the enormous investments are made in for example railways, the economic effects have to be well-defined in order to justify the investment. Besides the guarantee the investment or policy will have its desired effects, spill-over effects of an investment or a policy in the context of the larger transportation network could also help or hamper making this justification. Ports may also want to be attractive for service providers shipping large volumes on corridors or seeking access to certain hinterlands. The relevance of service throughput time in this process is the port authority can estimate what the frequency and transport time of services should be in order to establish a certain position in the network. A service provider could also benefit from the information how a transport network changes if a faster or more frequent service is offered. It could for instance mean a switch of modality, an increase of demand, or an opportunity to enter new regions. For rail operators the practical relevance is more specific on what throughput time is needed to establish a corridor, and for barge operators it is mainly how they can be more attractive than using a truck on a short distance range to cities with a limited number of connections. For the latter, the service throughput time can be included in the trade-off between offering a fast service and being cost-effective in transporting a small number of containers to a city with a limited number of connections.

As indicated by Van Langen et al. (2017), who used the same 2016 dataset, an important next step in research is to observe how the network changes over time. With two years more of stable data this research has made a start to this, yet a lack of longitudinal data has limited this research in increasing the correlations and significances between the average service throughput time and the network measures. A second limitation of this study was the limited amount of literature exploring complex network structures in hinterlands over time. However, methods used on static, social, and neural networks, and transportation network addressing global structures have proved to be useful. Research could however be improved if more specific measures were developed taking characteristics of rail and barge transport into account. A third and final limitation was the geographical scope of this research: only the hinterland network in Europe was considered.

Future research is therefore desired in testing whether the influence of service throughput time is also of influence in other hinterland transportation networks than the European one. Europe has for instance a large population in a compact geographical area, while the US has larger distances between its economical centres. Service throughput times could be of less influence there as the travel times are long anyways. It could influence differently on barge operators in an environment where 'greener' modes of transport are less supported or where the natural connectivity between river systems is greater or more limited. Confirming the influence of service throughput time in different environments would strengthen its validity as a link weight. The confirmation is of importance because it is relevant to many actors what a change in the network will result in. Extensions to this research based on the same dataset could be in the direction of testing its resilience if nodes are removed, adding dwell-times in hubs to the weighted shortest path lengths in order to increase the real-world representativeness, assigning properties to transport mode (shuttle service, hub-hub services, pre- and end-haulage), and exploring the competition between operators and its consequences on the offering of services and the network structure.

6 Reference List

Auber, D., & Mary, P. (2007). Tulip: data visualization software. LaBRI, University of Bordeaux, France.

Barabasi, A., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*, 509–512.

Barrat, A., Barthelemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences 101*, *11*, 3747–3752.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media.

Blondel, V., Guillamme, J., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), 10008.

Calatayud, A., Mangan, J., & Palacin, R. (2017). Connectivity to international markets: A multi-layered network approach. *Journal of Transport Geography*, *61*(February), 61–71.

Choi, J., Barnett, J., & Chon, B. (2006). Comparing world city networks: a network analysis of Internet backbone and air transport intercity linkages. *Global Networks*, *6*(1), 81–99.

Ciliberto, F., Cook, E., & Williams, J. (2017). Network Structure and Consolidation in the U.S. Airline Industry, 1990-2015. *CEPR Discussion Paper*, *83885*.

Crainic, T., & Kim, K. (2007). Intermodal transportation. *Handbooks in Operations Research and Management Science*, *14*, 467–537.

Crevier, B., Cordeau, J., & Savard, G. (2012). Integrated operations planning and revenue management for rail freight transportation. *Transportation Research Part B: Methodological*, *46*(1), 100–119.

De Langen, P., Lases Figueroa, D., Van Donselaar, K., & Bozuwa, J. (2017). Intermodal connectivity in Europe, an empirical exploration. *Research in Transportation Business and Management*, *23*, 3–11.

De Langen, P., Nijdam, M., & Van der Horst, M. (2007). New indicators to measure port performance. *Journal of Maritime Research*, *4*(1), 23–36.

De Langen, P., & Sharypova, K. (2013). Intermodal connectivity as a port performance indicator. *Research in Transportation Business and Management*, *8*, 97–102.

Demšar, U., Špatenková, O., & Virrantaus, K. (2008). Identifying critical locations in a spatial network with graph theory. *Transactions in GIS*, *12*(1), 61–82.

Ducruet, C. (2017). Multilayer dynamics of complex spatial networks: The case of global maritime flows (1977–2008). *Journal of Transport Geography*, *60*, 47–58.

Ducruet, C., & Lugo, I. (2013). Structure and Dynamics of Transportation Networks: Models, Methods and Applications. *The SAGE Handbook of Transport Studies*, 347–365.

Ducruet, C., & Notteboom, T. (2012). The worldwide maritime network of container shipping: Spatial structure and regional dynamics. *Global Networks*, *12*(3), 395–423.

Ducruet, C., Rozenblat, C., & Zaidi, F. (2010). Ports in multi-level maritime networks: Evidence from the Atlantic (1996-2006). *Journal of Transport Geography*, *18*(4), 508–518.

Ducruet, C., & Van der Horst, M. (2009). Transport integration at European ports. *European Journal of Transport and Infrastructure Research*, *9*(2), 121–142.

Ducruet, C., & Zaidi, F. (2012). Maritime constellations: a complex network approach to shipping and ports. *Maritime Policy and Management*, *39*(2), 151–168.

Eurostat. (2015). Modal split of freight transport. Retrieved March 15, 2018, from http://ec.europa.eu/eurostat/data/database

Eurostat. (2017a). Country level - gross weight of goods handled in all ports. Retrieved February 22, 2018, from http://ec.europa.eu/eurostat/data/database

Eurostat. (2017b). TEN-T Core Network Corridors (Freight and Passenger). Retrieved June 9, 2018, from http://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:TEN-T_Core_Network_Corridors_(Freight_and_Passenger)_RYB17.png&oldid=338850#file

Fleming, D., & Hayuth, Y. (1994). Spatial characteristics of transportation hubs: centrality and intermediacy. *Journal of Transport Geography*, *2*(1), 3–18.

Fraser, D., Notteboom, T., & Ducruet, C. (2016). Peripherality in the global container shipping network: the case of the Southern African container port system. *GeoJournal*, *81*(1), 139–151.

Fruchterman, T., & Reingold, E. (1991). Graph Drawing by Force-Directed Placement. *Software: Practice and Experience*, *21*(11).

Guerrero, D. (2014). Deep-sea hinterlands: Some empirical evidence of the spatial impact of containerization. *Journal of Transport Geography*, *35*, 84–94.

Guimera, R., Mossa, S., Turtschi, A., & Amaral, L. (2005). The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proceedings of the National Academy of Sciences*, *102*(22), 7794–7799.

Hagberg, A., Schult, D., & Swart, P. (2008). Exploring network structure, dynamics, and function using NetworkX. *Proceedings of the 7th Python in Science Conference (SciPy2008)*, 11–15.

Huang, C., Sun, C., & Lin, H. (2005). Influence of Local Information on Social Simulations in Small-World Network Models. *Journal of Artificial Societies and Social Simulation*, *8*(4).

Konings, R. (2009). *Intermodal barge transport: network design, nodes and competitiveness*. *TU Delft*.

Lam, J., & Yap, W. (2011). Dynamics of liner shipping network and port connectivity in supply chain systems: Analysis on East Asia. *Journal of Transport Geography*, *19*(6), 1272–1281.

Laxe, F., Seoane, M., & Montes, C. (2012). Maritime degree, centrality and vulnerability: port hierarchies and emerging areas in containerized transport (2008-2010). *Journal of Transport Geography*, *24*, 33–44.

Li, L., & Tayur, S. (2005). Medium-term pricing and operations planning in intermodal transportation. *Transportation Scienc*, *39*(1), 73–86.

Little, J. (1961). A Proof for the Queuing Formula: L = λW. *Operations Research*, *9*(3), 383–387.

Liu, C., Wang, J., & Zhang, H. (2018). Spatial heterogeneity of ports in the global maritime network detected by weighted ego network analysis. *Maritime Policy and Management*, *45*(1), 89–104.

Marcadon, J. (1999). Containerisation in the ports of Northern and Western Europe. *GeoJournal*, *48*(1), 15–20.

Martin, J., & Thomas, B. (2001). The container terminal community. *Maritime Policy and Management*, *28*(3), 279–292.

Meers, D., Macharis, C., Vermeiren, T., & Van Lier, T. (2017). Modal choice preferences in short-distance hinterland container transport. *Research in Transportation Business and Management*, *23*, 46–53.

Ng, A., Ducruet, C., Jacobs, W., Monios, J., Notteboom, T., Rodrigue, J., … Wilmsmeier, G. (2014). Port geography at the crossroads with human geography: Between flows and spaces. *Journal of Transport Geography*, *41*, 84–96.

Ng, A., & Lui, J. (2014). *Port-focal logistics and global supply chains*. Springer.

Notteboom, T. (1997). Concentration and load centre development in the European container port system. *Journal of Transport Geography*, *5*(2), 99–115.

Notteboom, T. (2007). Spatial dynamics in the container load centers of the Le Havre-Hamburg range. *Zeitschrift Fur Wirtschaftsgeographie*, *51*(2), 108–123.
Notteboom, T. (2008). The relationship between seasports and the intermodal hinterland in light of global supply chains. *International Transport Forum*, *10*.
Notteboom, T. (2010). Concentration and the formation of multi-port gateway regions in the European container port system: An update. *Journal of Transport Geography*, *18*(4), 567–583.
Notteboom, T., & Rodrigue, J. (2005). Port regionalization: Towards a new phase in port development. *Maritime Policy and Management*, *32*(3), 297–313.
Opsahl, T. (2009). Structure and Evolution of Weighted Networks. London, UK: University of London (Queen Mary College).
Pallis, A., Vitsounis, T., & De Langen, P. (2010). Port Economics, policy and management: Review of an emerging research field. *Transport Reviews*, *30*(1), 115–161.
Parshani, R., Rozenblat, C., Ietri, D., Ducruet, C., & Havlin, S. (2011). Inter-similarity between coupled networks. *EPL (Europhysics Letters)*, *92*(6), 68002.
Raimbault, N., Jacobs, W., & Van Dongen, F. (2016). Port Regionalisation from a Relational Perspective: The Rise of Venlo as Dutch International Logistics Hub. *Tijdschrift Voor Economische En Sociale Geografie*, *107*(1), 16–32.
Reis, V. (2014). Analysis of mode choice variables in short-distance intermodal freight transport using an agent-based model. *Transportation Research Part A: Policy and Practice*, *61*, 100–120.
Rodrigue, J., Debrie, J., Fremont, A., & Gouvernal, E. (2010). Functions and actors of inland ports: European and North American dynamics. *Journal of Transport Geography*, *18*(4), 519–529.
Sales-Pardo, M., Guimera, R., Moreira, A., & Amaral, L. (2007). Extracting the hierarchical organization. *Proceedings of the National Academy of Science*, *104*(39), 15224–15229.
Seoane, M., Laxe, F., & Montes, C. (2013). Foreland determination for containership and general cargo ports in Europe (2007-2011). *Journal of Transport Geography*, *30*, 56–67.
United Nations. (2015). *Review of Martime Transport*. New York, NY: United Nations Conference on Trade and Development.
Van den Heuvel, F., De Langen, P., Van Donselaar, K., & Fransoo, J. (2013). Spatial concentration and location dynamics in logistics: The case of a Dutch province. *Journal of Transport Geography*, *28*, 39–48.
Van der Horst, M., & De Langen, P. (2008). Coordination in hinterland transport chains: A major challenge for the seaport community. *Maritime Economics and Logistics*, *10*(1–2), 108–129.
Van Klink, H., & Van den Berg, G. (1998). Gateways and intermodalism. *Journal of Transport Geography*, *6*(1), 1–9.
Veldman, S., & Bückmann, E. (2003). A model on container Port competition: An application for the West European container Hub-Ports. *Maritime Economics and Logistics*, *5*(1), 3–22.
Veldman, S., Bückmann, E., & Saitua, R. (2005). River depth and container port market shares: The impact of deepening the Scheldt River on the West European container hub-port market shares. *Maritime Economics and Logistics*, *7*(4), 336–355.
Wang, T., & Cullinane, K. (2006). The efficiency of European container terminals and implications for supply chain management. *Maritime Economics and Logistics*, *8*(1), 82–99.
Wang, Y., & Cullinane, K. (2016). Determinants of port centrality in maritime container transportation. *Transportation Research Part E: Logistics and Transportation Review*, *95*, 326–340.

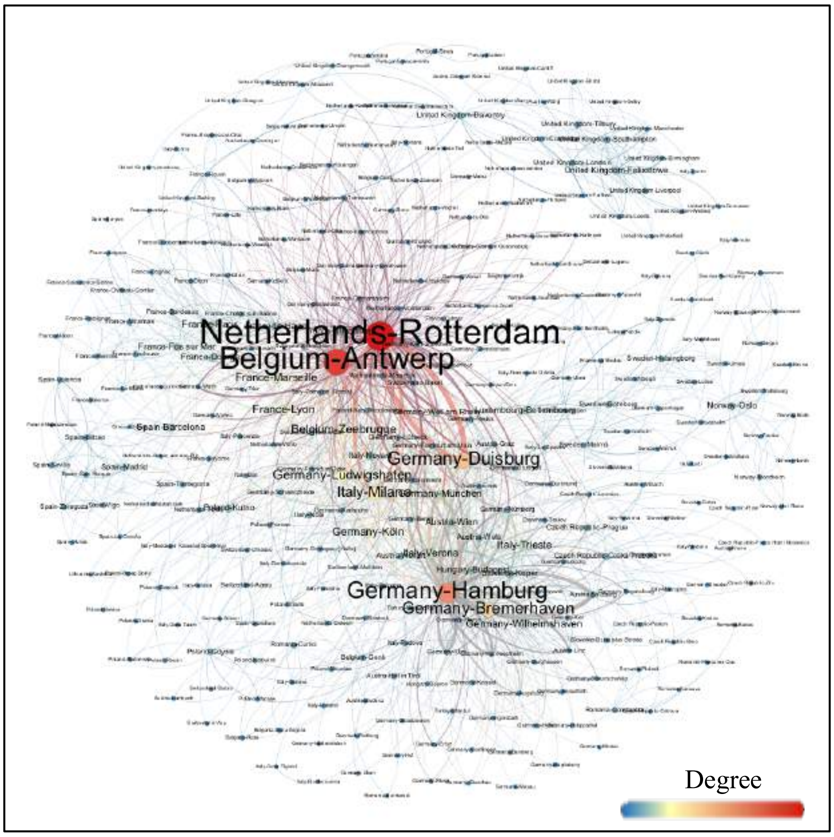
Watson, G. (2018). brainGraph: Graph Theory Analysis of Brain MRI Data.
Watts, D., & Strogatz, S. (1998). Collective dynamics of “small-world” networks. *Nature*, *393*(6684), 440–442.
Woo, S., Pettit, S., Kwak, D., & Beresford, A. (2011). Seaport research: A structured literature review on methodological issues since the 1980s. *Transportation Research Part A: Policy and Practice*, *45*(7), 667–685.
Xu, M., Li, Z., Shi, Y., Zhang, X., & Jiang, S. (2015). Evolution of regional inequality in the global shipping network. *Journal of Transport Geography*, *44*, 1–12.
Ypsilantis, P. (2016). The design, planning and execution of sustainable intermodal port-hinterland transport networks. *No. EPS-2016-395-LIS*.

7 Appendices

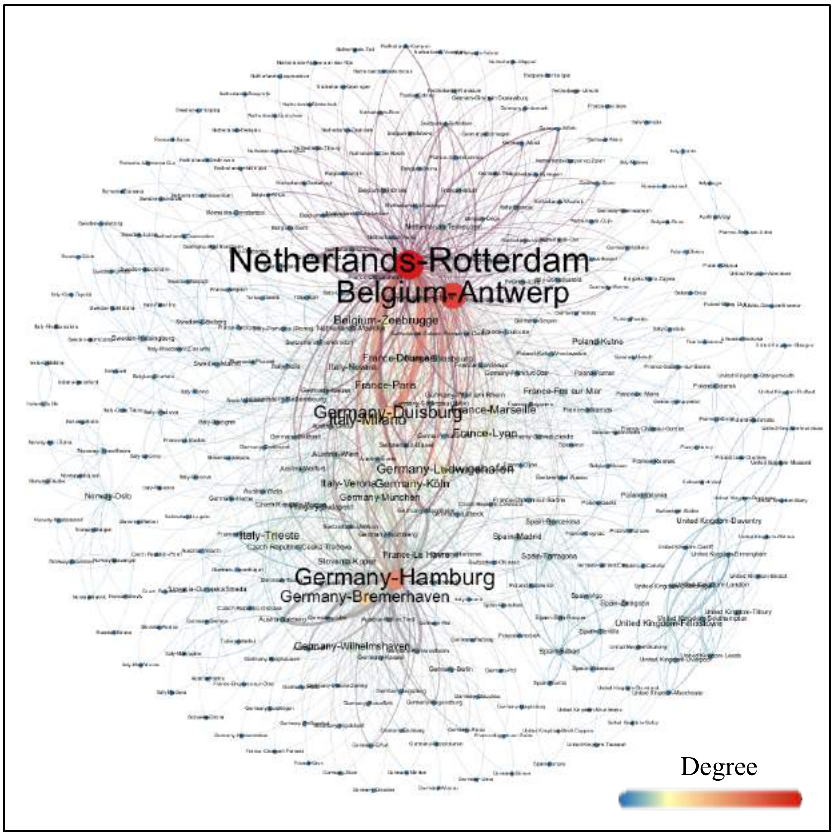
Appendix 1: Changes of degree in percentages for the 30 nodes with the highest degree in 2016. Only nodes already existing in 2016 are considered. Degree values extracted from Gephi.

Node	Degree			Delta [%]		
	2016	**2017**	**2018**	**'16-'17**	**'17-'18**	**'16-'18**
Netherlands - Rotterdam	205	207	210	1%	1%	2%
Belgium - Antwerp	179	179	181	0%	1%	1%
Germany - Hamburg	128	130	114	2%	-12%	-11%
Germany - Duisburg	92	90	90	-2%	0%	-2%
Germany - Bremerhaven	80	80	79	0%	-1%	-1%
Italy - Milan	67	72	82	7%	14%	22%
Germany - Ludwigshafen	57	56	53	-2%	-5%	-7%
France - Lyon	45	45	30	0%	-33%	-33%
Germany - Cologne	44	48	42	9%	-13%	-5%
France - Marseille	42	46	34	10%	-26%	-19%
Italy - Trieste	40	46	56	15%	22%	40%
Italy - Verona	38	40	40	5%	0%	5%
Germany - Wilhelmshaven	38	38	52	0%	37%	37%
France - Paris	36	42	32	17%	-24%	-11%
Germany - Munich	35	33	29	-6%	-12%	-17%
Austria - Vienna	30	24	28	-20%	17%	-7%
Spain - Barcelona	30	20	18	-33%	-10%	-40%
France - Fos sur Mer	28	29	28	4%	-3%	0%
Luxembourg - Bettembourg	28	15	12	-46%	-20%	-57%
France - Dourges	27	36	28	33%	-22%	4%
Hungary - Budapest	27	19	19	-30%	0%	-30%
United Kingdom - Felixstowe	26	28	31	8%	11%	19%
Poland - Kutno	26	26	18	0%	-31%	-31%
Norway - Oslo	26	24	30	-8%	25%	15%
France - Le Havre	24	30	32	25%	7%	33%
Italy - Novara	24	29	28	21%	-3%	17%
Slovenia - Koper	24	26	28	8%	8%	17%
France - Strasbourg	24	24	22	0%	-8%	-8%
Germany - Weil am Rhein	22	18	18	-18%	0%	-18%
			Average delta	**-0.02%**	**-2.87%**	**-2.91%**

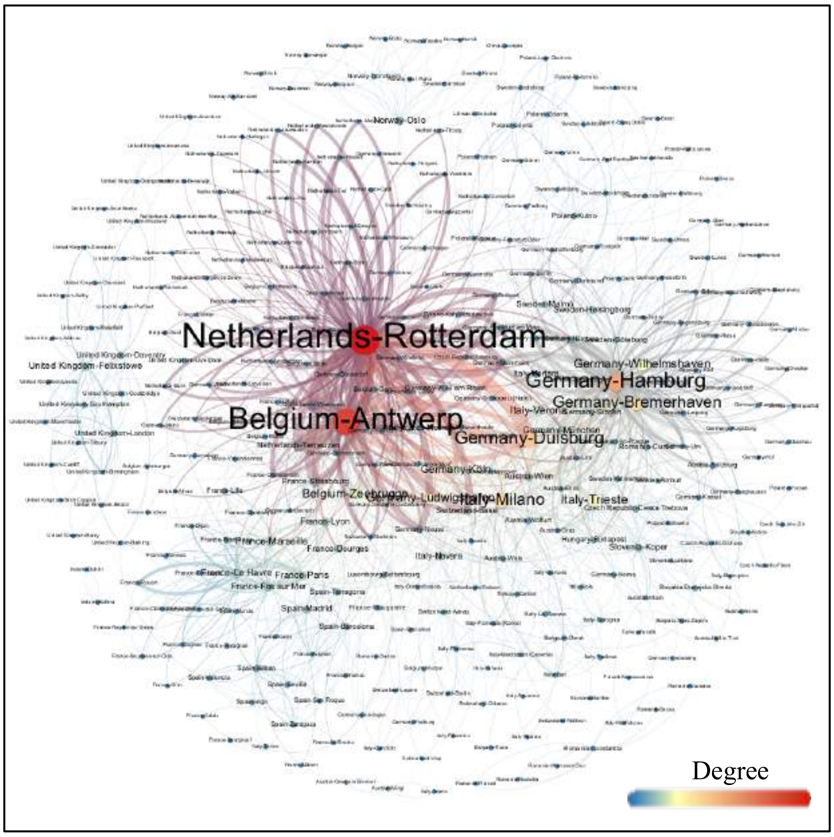
Appendix 2: The complete 2016 network with node weights and colouring. The thickness of the edges represents the number of services. The size of a node is based on the node degree. Visualised with Gephi.



Appendix 3: The complete 2017 network with node weights and colouring. The thickness of the edges represents the number of services. The size of a node is based on the node degree. Visualised with Gephi.



Appendix 4: The complete 2018 network node link weights and colouring. The thickness of the edges represents the number of services. The size of a node is based on the node degree. Visualised with Gephi.



Appendix 5: Average centrality and clustering measures for communities in 2016. Extracted from Gephi and Tulip and computed in Microsoft Excel.

Community 2016	**Nodes [#]**	**Average Degree**	**Diameter**	**Average Clustering**	**Average shortest path length**
Trans Alp	3	1.333	2	0	1.333
Portugal	4	1.500	2	0	1.500
Czech/Slovakia	11	2	3	0.188	2.164
Spain	13	5	4	0.459	1.688
Norway/Sweden	18	2.778	4	0.312	2.268
Germany/Sweden	21	3.333	3	0.440	2.148
Poland	21	2.429	4	0.064	2.436
Austria/Italy/Slovakia	21	3.571	4	0.402	2.217
UK	23	3.913	4	0.276	2.368
France	26	5.115	4	0.535	2.026
Germany	48	4.125	4	0.264	2.176
Romania/Italy/Germany	52	2.635	7	0.137	3.162
Netherlands	65	3.400	3	0.624	1.995

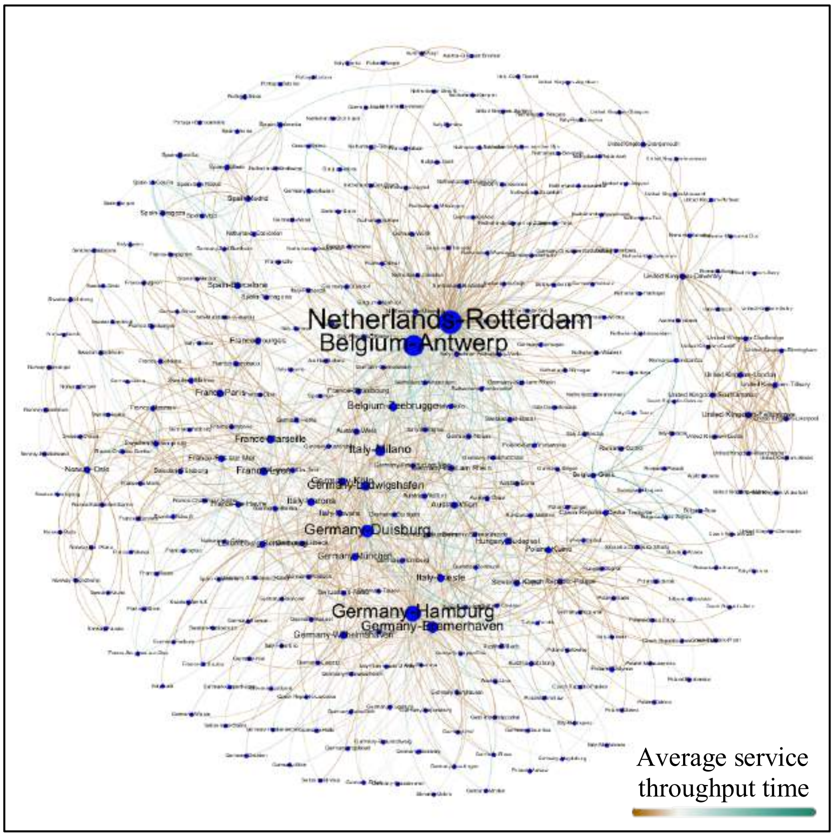
Appendix 6: Average centrality and clustering measures for communities in 2017. Extracted from Gephi and Tulip and computed in Microsoft Excel.

Community 2017	**Nodes [#]**	**Average Degree**	**Diameter**	**Average Clustering**	**Average shortest path length**
Lugo (ITA)/Arcis-Sur-Aube (FRA)	2	1	1	0	1
Trans Alp	3	1.333	2	0	1.333
Ireland	3	1.333	2	0	1.333
Bulgaria/Romania	10	2	3	0.207	2.111
Spain	12	5.833	3	0.569	1.515
Norway	13	2.769	3	0.344	1.872
Poland	13	2.462	3	0.000	2.128
Germany/Sweden	23	2.609	5	0.140	2.593
UK	27	3.222	5	0.124	2.467
BEL/DEU/ITA/ESP/CHE	36	2.972	7	0.183	2.904
France/Germany/Italy	42	4.571	4	0.438	2.429
AUT/CZE/DEU/ITA/SVK/POL	71	4.549	6	0.246	2.517
AUT/BEL/FRA/DEU/NLD/CHE	79	3.658	3	0.659	2.004

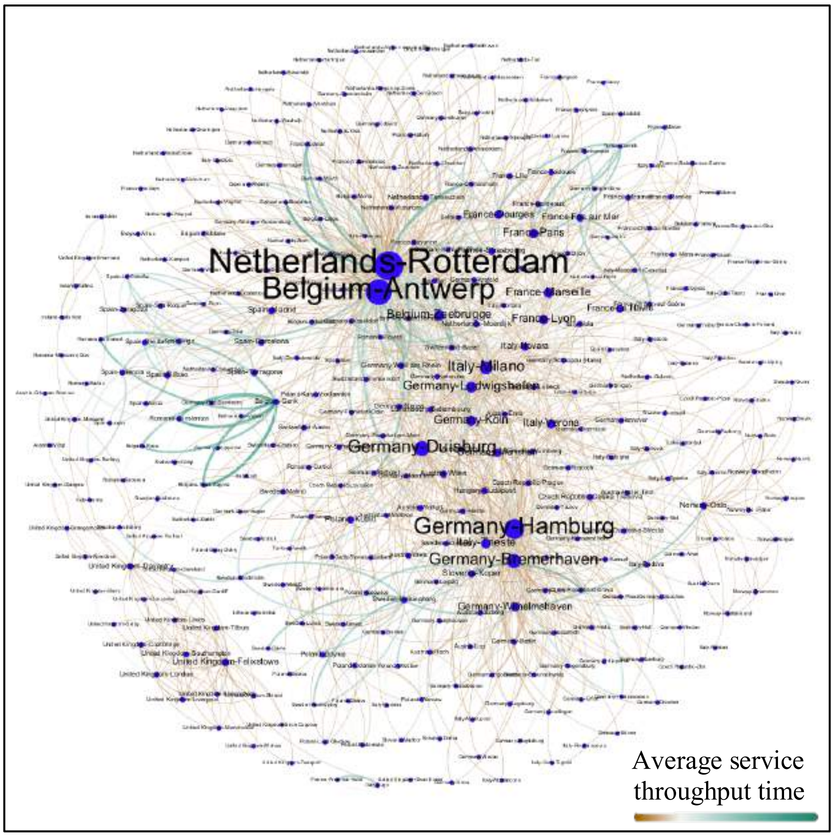
Appendix 7: Average centrality and clustering measures for communities in 2018. Extracted from Gephi and Tulip and computed in Microsoft Excel.

Community 2018	**Nodes [#]**	**Average Degree**	**Diameter**	**Average Clustering**	**Average shortest path length**
Ireland	2	1	1	0	1
Trans Alp	3	1.333	2	0	1.333
Romania	6	1.667	2	0	1.667
Germany/Switzerland	10	2	6	0	2.622
Poland	13	2.462	4	0	2.231
Norway/Sweden	14	2.429	3	0.319	1.934
Germany/Spain	19	4.421	5	0.366	2.538
Bulgaria/Germany/Italy	21	2.619	5	0.301	2.695
UK	28	2.964	5	0.114	2.403
France	31	4.226	6	0.281	2.464
AUT/DEU/ITA/SWE/CHE	49	4.449	5	0.256	2.529
AUT/CZE/DEU/SVK	58	4.328	5	0.303	2.424
BEL/FRA/DEU/NLD/CHE	76	3.645	5	0.546	2.044

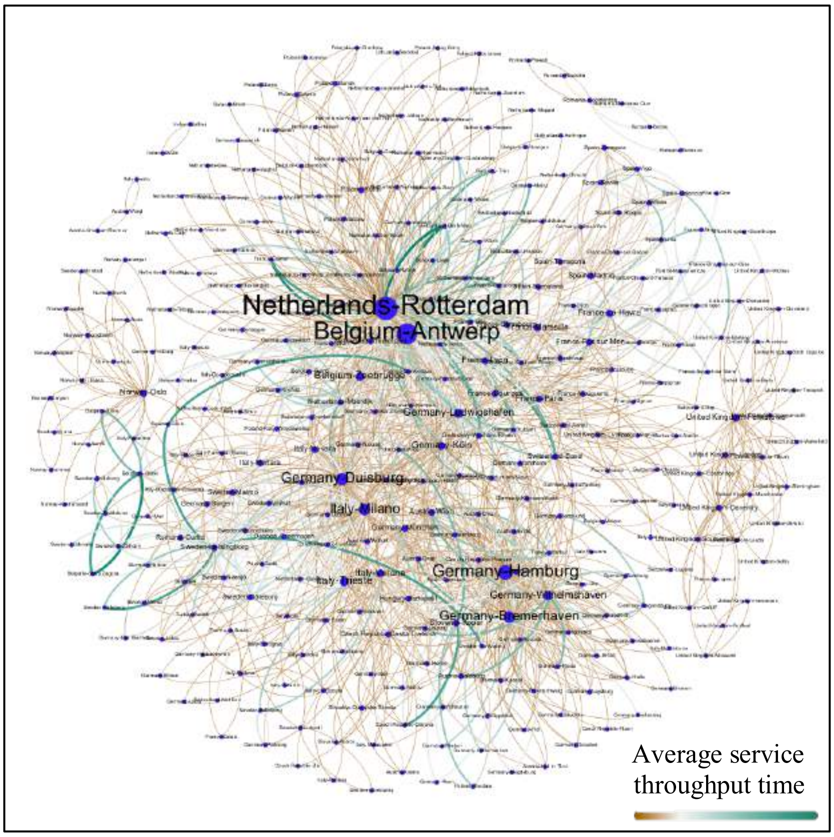
Appendix 8: The complete 2016 network with service throughput time as link weight. The colour and thickness of the edges represents the average service throughput time. The size of a node is based on the node degree. Visualised with Gephi.



Appendix 9: The complete 2017 network with service throughput time as link weight. The colour and thickness of the edges represents the average service throughput time. The size of a node is based on the node degree. Visualised with Gephi.



Appendix 10: The complete 2018 network except for the rail link with China (it distorted the colouring because of the long travel time) with service throughput time as link weight. The colour and thickness of the edges represents the average service throughput time. The size of a node is based on the node degree. Visualised with Gephi.



Appendix 11: Average centrality and clustering measures for communities with link weights in 2017. Extracted from Gephi and Tulip and computed in Microsoft Excel.

Community 2016	**Nodes [#]**	**Average Degree**	**Diameter**	**Average Clustering**	**Average shortest path length**
Trans Alp	3	1.333	2	0.000	1.333
Portugal	4	1.500	2	0.000	1.500
Romania/Bulgaria/Hungary	13	2.308	4	0.179	2.244
Germany/Italy/Spain	18	2.556	5	0.052	2.536
UK	23	3.913	4	0.276	2.368
Norway/Sweden	30	2.933	5	0.226	2.777
Germany/Poland	32	3.344	5	0.292	2.440
Italy/Netherlands/Switzerland	34	2.471	6	0.207	2.839
France/Spain	38	5.211	6	0.459	2.729
Austria/Czech/Germany	44	3.500	4	0.207	2.248
AUT/FRA/DEU/NLD/CHE/CZE	87	3.678	4	0.543	2.161

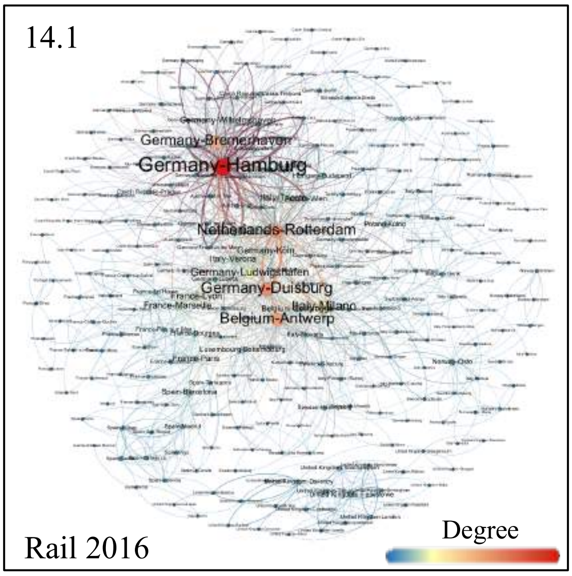
Appendix 12: Average centrality and clustering measures for communities with link weights in 2017. Extracted from Gephi and Tulip and computed in Microsoft Excel.

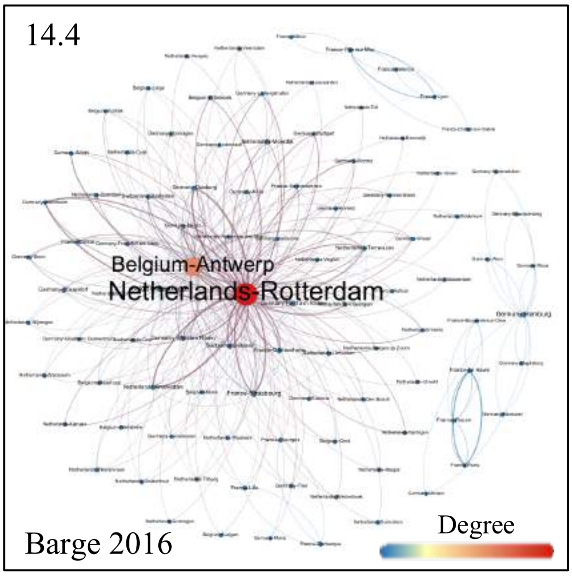
Community 2017	**Nodes [#]**	**Average Degree**	**Diameter**	**Average Clustering**	**Average shortest path length**
Luga (ITA)/Arcis-sur-Aube (FRA)	2	1	1	0	1
Ireland	3	1.333	2	0	1.333
Trans Alp	3	1.333	2	0	1.333
Spain	13	5.538	4	0.508	1.667
BEL/BGR/DEU/ITA/ROU	16	2.062	5	0.226	2.567
Germany/Poland/Spain	18	2.222	5	0.152	2.863
AUT/DEU/ITA/POL/SVK/SWE	26	2.962	5	0.094	2.594
UK	26	3.269	5	0.130	2.456
Norway/Sweden	28	2.714	6	0.276	3.101
BEL/DNK/DEU/ITA/CHE/NLD	30	3.233	5	0.107	2.680
France/Germany	33	4.697	4	0.510	2.212
AUT/CZE/DEU/ITA/SVK	56	3.839	7	0.259	2.932
BEL/FRA/DEU/NLD/CHE	80	3.612	4	0.626	2.051

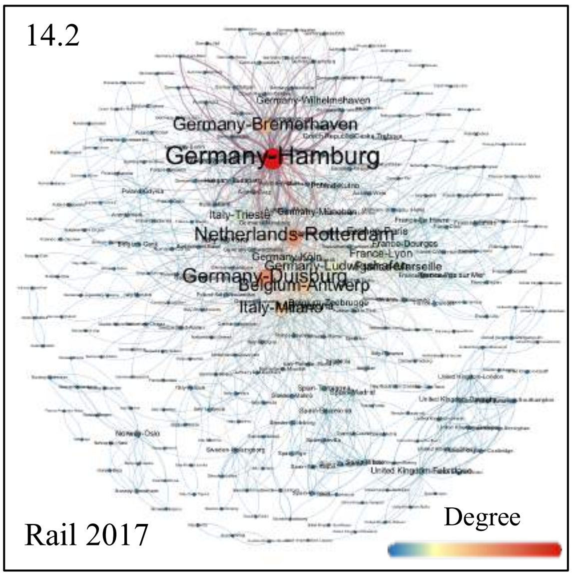
Appendix 13: Average centrality and clustering measures for communities with link weights in 2018. Extracted from Gephi and Tulip and computed in Microsoft Excel.

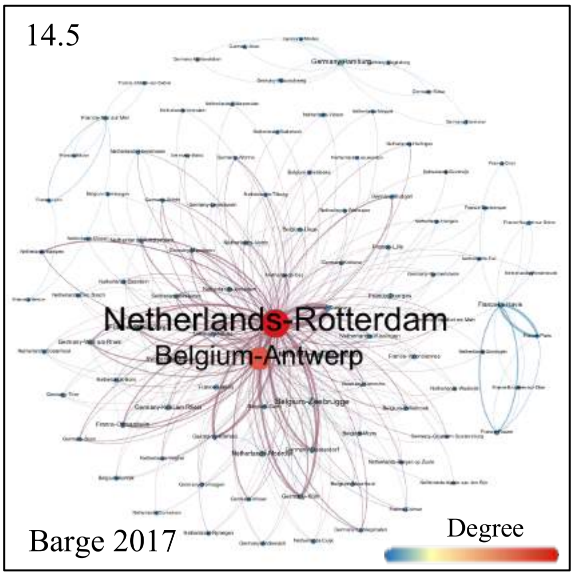
Community 2018	**Nodes [#]**	**Average Degree**	**Diameter**	**Average Clustering**	**Average shortest path length**
Netherlands/China	2	1	1	0	1
Ireland	2	1	1	0	1
Trans Alp	3	1.333	2	0	1.333
Romania	6	1.667	2	0	1.667
Poland	13	2.462	4	0	2.231
Bulgaria/Germany/Italy	15	2.267	6	0.137	2.733
Germany/Spain	16	4.812	4	0.412	1.933
BEL/FRA/DEU/ITA/NLD	18	2.611	5	0.256	2.706
AUT/CZE/DEU/POL/SVK/CHE	25	2.760	5	0.293	2.514
France	30	4.100	6	0.263	2.484
Norway	14	2.429	3	0.319	1.934
DNK/DEU/ITA/SWE	40	4.050	5	0.146	2.533
Austria/Germany	41	4.805	4	0.420	1.991
UK	26	3.077	4	0.132	2.362
BEL/FRA/DEU/NLD/CHE	79	3.367	4	0.515	2.155

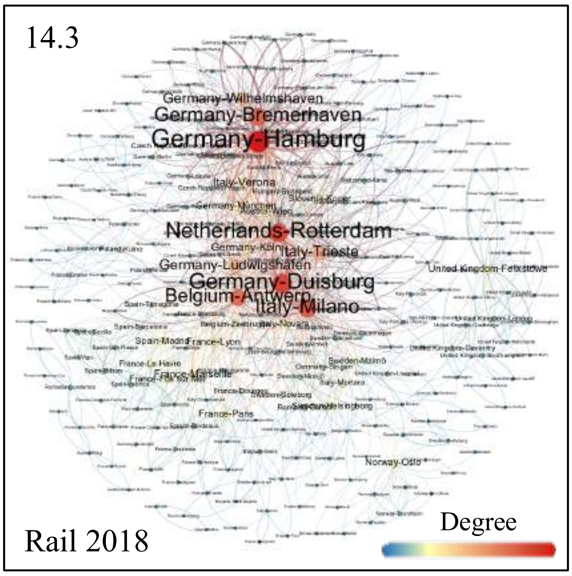
Appendix 14 (.1, .2, .3, .4, .5, .6): The rail and barge networks for 2016-2018 with node weights and colouring. The thickness of the edges represents the number of services. The size of a node is based on the node degree. Visualised with Gephi.

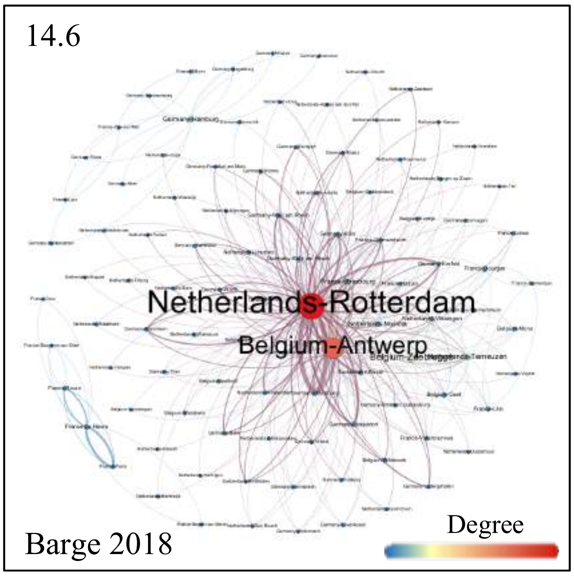












Appendix 15 (.1, .2, .3, .4, .5, .6): The 2016-2018 rail and barge networks (except the link with China in 13.4 because it distorted the colour scale with its long travel time) with service throughput time as link weight. The colour and thickness of edges represents the average service throughput time. The node-size is based on the degree. Visualised with Gephi.