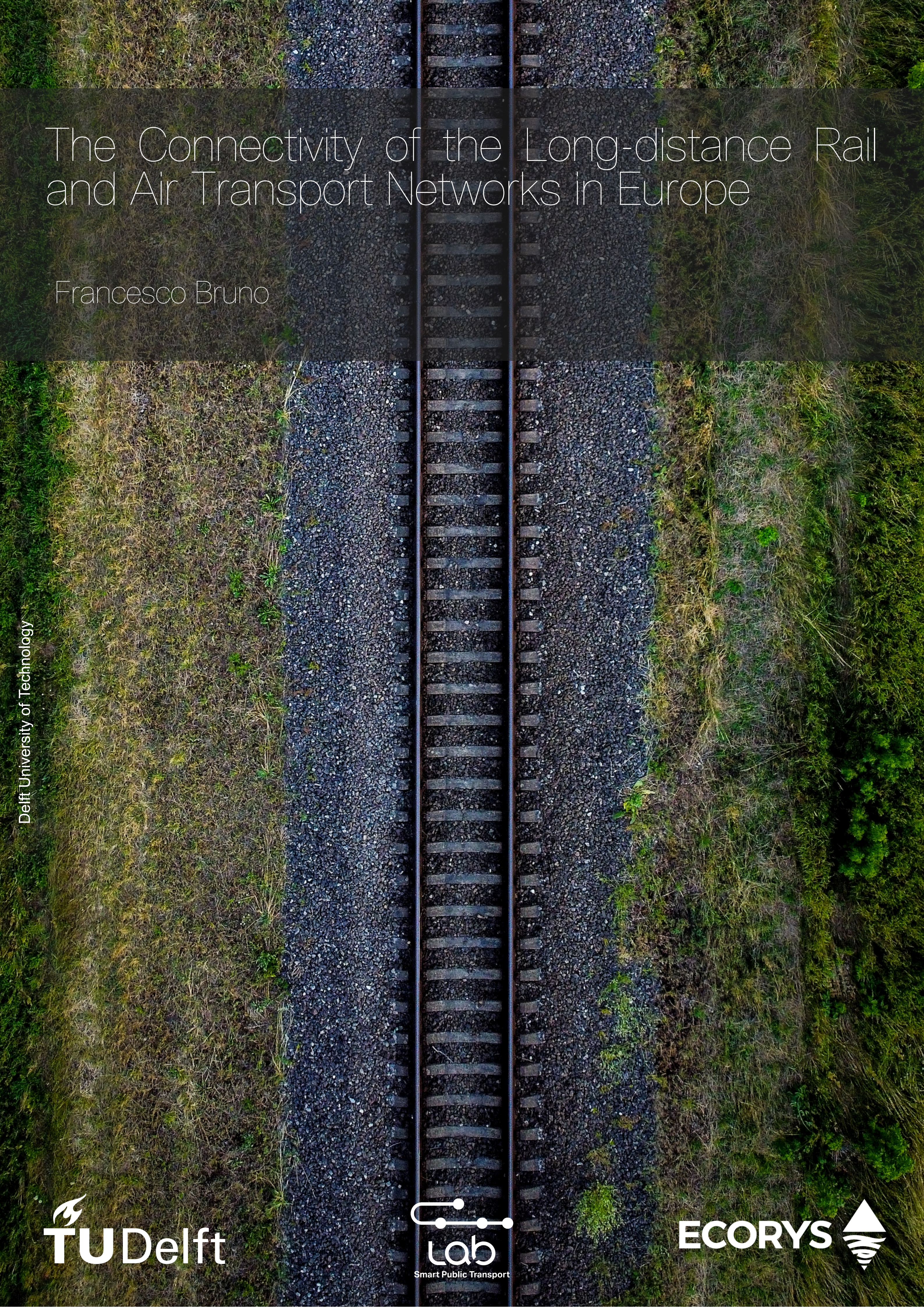


The Connectivity of the Long-distance Rail and Air Transport Networks in Europe

Francesco Bruno

Delft University of Technology

The Connectivity of the Long-distance Rail and Air Transport Networks in Europe

by

Francesco Bruno

in partial fulfilment of the requirements for the degree of

Master of Science in Transport, Infrastructure and Logistics

at the Delft University of Technology,

to be defended publicly on Monday December 19th, 2022

Student number:	5382122		
Project duration:	April 19, 2022 - December 19, 2022		
Thesis committee:	Prof. dr. G.P. van Wee	TU Delft, TPM	Chair
	Dr. O. Cats	TU Delft, CEG	Supervisor
	Dr. A. Bombelli	TU Delft, AE	Daily supervisor

Smart Public Transport

Preface

This thesis represents the culmination of my master's studies at TU Delft and is the result of two intense years, which have been particularly transformational for me. Looking back at the experiences that marked this journey I realise how differently I would approach them today. And this, I believe, represents the scale of my personal and academic growth. Working on this thesis made me confront my limits and face my fears. It tested my ideas, ideals, methods and habits. Tackling and solving practical problems, thinking through numbers and functions, being constantly open to learning, being patient and self-compassionate, asking for help and managing to strike a balance have all been challenging parts of this journey. However, I believe that what all these extraordinary challenges have brought to me in terms of personal and academic development is inestimable. Although this thesis required a considerable amount of autonomy and independent work, I would not have managed without the help and support of many people, whom I want to thank in the following paragraphs.

First of all, I would like to express my gratitude to the graduation committee for their unconditional support throughout the entire process. To Bert, thank you for your enthusiasm and for the insightful perspectives you have offered. Thank you, Oded, for not providing the easy way out, for making critical and thoughtful questions and for offering crucial food for thought, constantly stimulating my reasoning. Thanks for trusting me and my decisions, even when I did not and for pushing me to achieve things I did not think possible. To Alessandro, thank you for your attentive care and patience. I have always felt welcomed and listened to. Thanks for the openness you have put in the mindful feedback you provided and in our conversations.

Furthermore, I would like to thank Ecorys, the company whose project inspired this research and provided fertile terrain for this study to flourish. Thanks for welcoming me into the team and providing a stimulating environment to work in. My time at Ecorys has enriched me on many distinct levels, teaching me about human relations and providing specific knowledge in the transport field. I am deeply grateful for the opportunity I have received and would like to thank all Ecorys' colleagues for the moments we have shared in these months. A special thank you to Geert for believing in me. To Hugh for the fruitful conversations and exchanges of ideas. To Marco and Jasper for the significant patience and support. And to Michael and Niklas from Hacon for allowing me to access the data required to develop part of the analysis.

I also want to take this opportunity to thank my friends and family, who have relentlessly supported me throughout this journey. To my friends in the Netherlands, thank you for all the special moments and experiences that have made these two years unforgettable. A special thanks to Anekëtt for the constructive teamwork and intense intellectual exchanges. To Alex, for teaching me through your example. To Beatrice for being there when I needed the most. To Filippo, Jeronimo and Nick, for being the best friends and study partners I could have ever hoped for. To Marc, for guiding me through the hard times like an older brother would. To Pim for providing novel perspectives that have helped me to discover more about myself and understand better who I am. To my friends in Italy, thank you for *being there despite the distance that separates us. Special thanks to Piero, Davide, Andrea, Giacomo, Filippo and Lorenzo for being a safe haven to return to. To Chiara and Sheida, for pushing me to see the bright side when everything looked dark. To Sofia, for being my endless source of inspiration, proving through your brave determination that everything is attainable with the proper effort. To Martina, for the unconditional support and love you have constantly offered me. Finally, I want to thank my family, aunts, uncles and cousins, and especially my parents and grandparents, for their immeasurable support and education. Without you, I would not be who I am or where I am today.*

Lastly, thanks to TU Delft, to the professors and the fellow students I have met during these two years, for making this experience truly exceptional. As my future endeavours unfold, I will do my best to use the knowledge and skills gained throughout this journey to enhance the sustainability and ethics of the transport sector.

Francesco Bruno
Rotterdam, December 2022

Contents

List of Figures

List of Tables

1

Introduction

In the last decades, the figures of long-distance transport have exponentially grown all over the globe, and especially in Europe. As a consequence, the transportation industry has had to face its increasing limitations in terms of service capacity and environmental impact. The modal shift from air to rail, from short-haul flights to long-distance rail services has been proposed by many and widely discussed. However, the actual potential of this shift from a network perspective does not appear to have been assessed yet. In particular, comparing the connectivity of air and rail transport networks on a continental scale represent a novel and unexplored field of research, which is closely related to many others. This first chapter aims to present and introduce the broader context into which this research fits, highlighting the approach employed and the relevance of the topic. In particular, Section 1.1 introduces the background of the proposed study, exploring the current state of long-distance international travel practices in the European continent. Following this, the research problem is defined and scoped, with the problem definition being discussed in Section 1.2 and the scope of the study being provided in Section 1.8. Furthermore, Section 1.3 focuses on defining the research gap, both from a scientific and a practical perspective, whilst the empirical and the methodological objectives of this research are defined in Section 1.4. Then, Section 1.5 highlights the main research question and the sub-questions required to try to fill the aforementioned research gap and to reach the research aim. Finally, Sections 1.6, 1.7 and 1.9 conclude the chapter, pointing out the relevance of this study, its approach and the report structure respectively.

1.1. Research Context

1.1.1. The evolution of the transport market

A lot has changed since the times when a journey of a few hundreds kilometres was considered a rare and special event. In the Europe of 2022, it has become rather common to travel often, covering increasingly longer distances. As a reference, Eurostat (2019) highlights that in 2017 a total of more than 1 billion trips was made by European residents, with 267 million citizens, equivalent to 62% of the total population, going on at least one private trip over the course of the year. In particular, 73% of those trips are domestic whilst the trips within Europe account for another 23% (with almost 21% of being between EU Member States) whilst a mere 3% consists of intercontinental trips. And the figures of European tourism are not expected to slow down, with the WTO (2018) projecting arrivals in EU destinations from European source markets to grow by 1.9% a year on average through 2030. Furthermore, faster and more frequent transport connections have made it possible to commute not only within the vicinity of major urban areas but also across different regions and even nations. In 2018, 18.3 million citizens, corresponding to 8.3% of all the employed population, commuted between regions within their country of residence whilst 1.3 million (0.6% of all employed) regularly commuted across borders, living in one Member State and working in another one (Eurostat, 2019).

These figures are the outcome of the considerable infrastructural investments and the radical market changes that over the last century have reshaped the entire transport sector. In particular, the road and air modes of transport have been the two protagonists and key players in this revolution. The number of automobiles, from the second postwar period, has boomed, rapidly making it the most common transport mode across the continent, with passenger cars in the EU accounting for more than half of its population (Eurostat, 2022c). Similarly, the aviation sector, despite starting its ascent slightly later, has managed to attract enormous investments experiencing an impressive and constant growth ever since the 1980s. Between 2008 and 2017 air traffic grew by 30%, with over 1 billion passengers flying in 2017 across the EU, half of whom travelling within the continent (Eurostat, 2019). The victim of these trends was rail, whose market shares in passenger transport have kept plummeting from the 1970s until the 2000s when they have stabilised around 6% (Di Pietrantonio & Pelkmans, 2004). As a result, in 2018 almost 70% of the private trips across the European Union were made by car, 14% by air and only 9% by rail (Eurostat, 2019).

1.1.2. The environment and long-distance travel

The evolution of the transport market and travel behaviours over the last decades of the twentieth century has had a twofold impact. On the one hand, it has provided a more ”democratic” and widespread access to transport, making it increasingly easy to travel across the continent and opening to a world of new possibilities. On the other, the growing ease to travel, the enhanced accessibility of transport services, and the consequent boost in traffic and movement of passengers and freight have created a wide range of problems and negative externalities, which have been further exacerbated by the extensive dominion of fossil fuels modes. Over the last decades, public attention has increasingly focused on climate change shedding light over these issues. This, in turn has driven to the rise of widespread concerns regarding the environmental impact of transport, especially in first world countries. In fact, whilst greenhouse gas (GhG) emission levels are generally declining, the emissions caused by the transport sector are still on the rise. Eurostat (2021d) highlights that in 2019 the GhG emission in the European Union, despite recording a 24% drop compared to the overall levels of 1990, have increased by more than 30% in the transport sector, which now accounts for over 25% of the total GhG emissions of the EU (Eurostat, 2021a).

As much as the current scenario is the result of the political decisions of the past, which have oftentimes favoured air and road transportation in terms of infrastructural investments and tax regulations, the future is going to be shaped by the political decisions of today (Feuerstein et al., 2018). Thus, European politics, has taken these matters at heart, looking for possible solutions to change the tide. In doing so, particular attention has been given to the long-distance market, which currently represents the main source of transport related GhG emissions (Rich & Mabit, 2012). Long-distance is a rather vague concept and has been variously defined across the literature, generally taking into account either the distance between origin and destination or the length of the trip in term of overnight stay (Gerike & Schulz, 2018). However, as many researchers and the European Commission (EC) agree in defining trips longer than 100km as long-distance travel, this definition is adopted throughout this study (Frei et al., 2010; Limtanakool et al., 2006; Malichová et al., 2022; Petersen et al., 2009; Rich & Mabit, 2012). In recent decades, the figures in long-distance travel between the main urban areas of the European Union have soared dramatically, especially in terms of road and air transport (World Tourism Organization, 2018). And this trend is not projected to cease, as literature generally agrees in saying that an increase in the share of medium- and long-distance trips in Europe is to be expected (Limtanakool et al., 2006). In 2009, trips longer than 100 km, despite representing only 2.5% of all the European trips, accounted for 55% of the passenger-km travelled across the continent (Petersen et al., 2009). The high share of passenger-km of long-distance trips clearly proves the crucial role of this market in reducing the GhG emissions of the entire transport sector. However, currently a large majority of the long-distance market's passenger-km still relate to road and air transport, as the increases in demand did not manage to be reflected in the rail sector yet (Malichová et al., 2022). In particular, regarding the comparison between rail and air transport , the European Environment Agency (EEA) points out that over the last decades the passenger-km of the aviation sector have risen much more sharply and steadily in comparison to rail in the European market, as illustrated by Figure 1.1 (European Environment Agency, 2021).

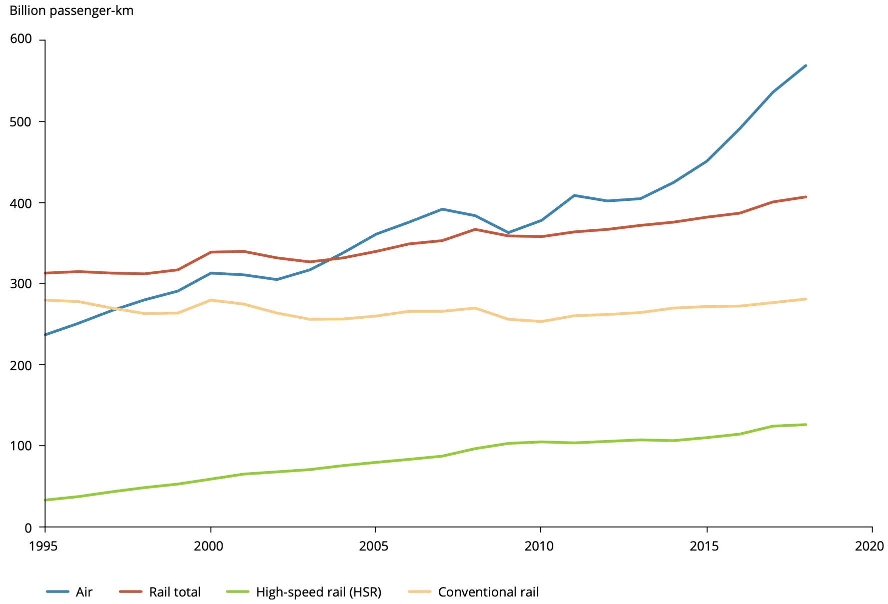


Figure 1.1: Passenger-km travelled by rail and air, EU-27, 1995-2018 (Source: European Environment Agency, 2021)

1.1.3. The role of rail

Nowadays, rail is increasingly taking centre stage in the transport debate, attracting the attention of many scholars and policymakers. An important reason for this growing trend relates to its crucial role in the fight for the reduction of GhG emissions in the transport sector. In fact, rail is widely regarded as one of the most environmentally friendly and energy-efficient transport modes, especially when compared to the air and road competitors on the long-distance market (European Commission, 2020a). In this regard, the results of a study commissioned by the EEA, illustrated in Figure 1.2, show that rail is the most efficient transport mode in the EU, featuring considerably lower GhG emissions per passenger-km

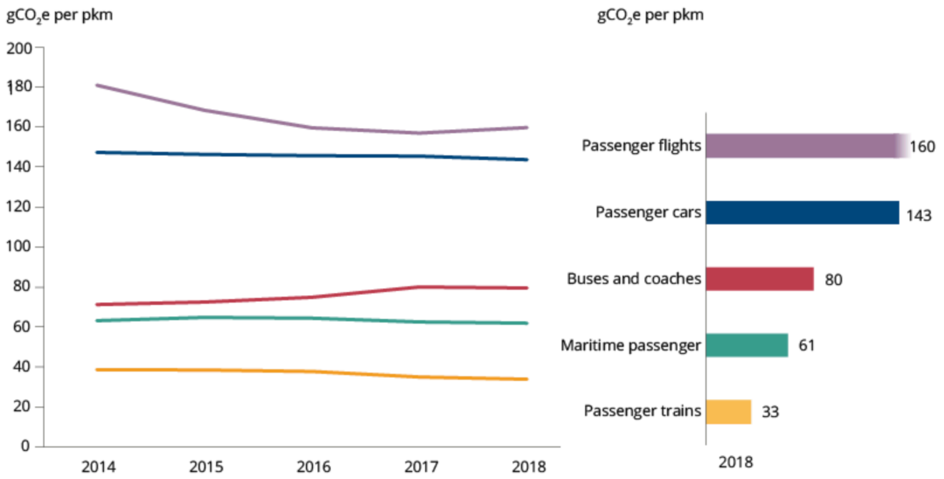


Figure 1.2: Average GHG emissions by motorised mode of passenger transport, EU-27, 2014-2018 (Source: Fraunhofer ISI and CE Delft, 2020)

(pkm) compared to all other modes, and especially in relation to aviation and road transport, the two main competitors on long-distance transport (Fraunhofer ISI & CE Delft, 2020). Furthermore, the EC also highlights that rail is the only transport mode that over the last decades has consistently allowed to limit the impact of transport-related emissions in the EU (European Commission, 2021b). Thus, the European Union is determined in jointly working to enhance the attractiveness of railways, aiming to increase its modal shares at the expenses of the aforementioned air and road transport modes European Parliament and Council of the European Union (2020). The European Parliament has, in fact, recently approved a set of policy initiatives named European Green Deal (EGD), which aim to make Europe the first climate-neutral continent by 2050 (European Commission, 2020c). In order to reach these ambitious goals drastic changes in the fields of transport and mobility are required. In particular, the EC in the Sustainable and Smart Mobility Strategy (SSMS) highlights that rail represents one of the core pillars of this revolutionising process of the transport sector. In fact, in order to achieve the objectives set in the EGD, the EC set specific goals for the rail market. If a 90% reduction in transport-related greenhouse gas emission is to be obtained, traffic on high-speed rail ought to double by 2030 and triple by 2050, whilst rail freight traffic figures should increase by half by 2030 and double by 2050 (European Commission, 2020b).

1.1.4. The impact of Covid-19

The outbreak of the Covid-19 pandemic in 2020, despite causing severe setbacks in both passenger and freight transport, only renewed and reinforced the crucial role of rail in the transport landscape of the European continent. In 2020 rail travel has seen considerable losses in patronage, with ridership sinking in every European country, recording reductions in comparison to 2019 that ranged from the 74% of Ireland to the 22% of Bulgaria (Eurostat, 2021b). Although the pandemic clearly burdened the European rail market, some scholars argue that the current situation could be the occasion for rail transport to improve its competitiveness, especially in terms of medium- and long-distance travels, where the direct competitor is air transport. In this regard, S. Yang and Chen (2022) found that whilst the service frequency of both rail and air transport dropped significantly after the outbreak, only the former has managed to rapidly recover after the relaxation of the Covid-19 related measures and restrictions. Furthermore, Tardivo et al. (2021) state that during the early phases of the pandemic rail services have experienced a lower number of disruptions in comparison to other transport modes, arguing that rail transport will prove paramount in ensuring the availability of Public Transport during times of pandemic. The authors, in fact, highlight that core characteristics of railway systems allow rail services to adapt more easily than other modes to the safety requirements introduced to prevent the spread of the virus (Tardivo et al., 2021). On a different note, Rothengatter et al. (2021) sustain that some time will probably be required to re-establish the lost ridership confidence in public transport, arguing that stimulus funding packages could actually help to improve the attractiveness of public transportation modes, restoring confidence in more sustainable transport modes and contributing to the growth of transport systems' resilience.

1.1.5. The European response

In this context, on the 23 December 2020 the European Parliament and the Council of the European Union approved the Decision 2020/2228, making of 2021 the European Year of Rail (European Parliament & Council of the European Union, 2020). Thus, the European Union has developed and launched an Action Plan that aims to boost long-distance and cross-border passenger rail transport, by overcoming the obstacles that still seem to represent an hindrance for the flourishing of such services (European Commission, 2021a). In particular, the Action Plan identifies the lack of connectivity of the networks and their sub-optimal use among the many barriers to the development of this kind of rail services (Serafimova et al., 2022). Alongside the aforementioned Action Plan the EC is also revising the Trans-European transport network (TEN-T) to increase high-speed rail capacity, and is assuring increased support from the European Investment Bank (EIB) towards targeted public and private investment in rail projects to accelerate the recovery of the sector after the advent of the pandemic (European Commission, 2021a).

1.2. Problem Definition

The considerable efforts of the European Union directed at making of railways the main pillar of European transport do not represent a novelty of the last years, but are rooted in a process that started more than three decades ago. Following the success of the deregulation of the airline market in the 1980s, the EU since 1991 has started to develop a new political and legal framework aimed at the revitalisation of the heavily subsidised and largely inefficient rail market (Feuerstein et al., 2018). Between 2001 and 2016, the EU has adopted four legislative packages targeted at gradually opening up the rail market for competition, at assuring the interoperability between different national railway systems and at developing a single European railway area (European Commission, 2022b). Through the Decision 2020/2228, the European Union has stipulated that the rules agreed under the Fourth Railway Package, approved in 2016, are to be implemented throughout the Union (European Parliament & Council of the European Union, 2020). The Fourth Railway Package is a set of six legislative texts designed to project towards the creation of a single European market for Rail services, the Single European Railway Area (European Commission, 2022a). In particular, its overarching goal was to revitalise the rail sector and to increase its competitiveness with other modes, such as road transport on the short-medium range (0-400 km) and air transport on the medium-long range (400-1000 km). In fact, high-speed and night trains are widely deemed capable of becoming a sustainable alternative to flights with a range of less than 800-1000 kilometres if appropriate and thorough political support is provided (Seidenglanz et al., 2021; Serafimova et al., 2022; Witlox et al., 2022). The study from Zhu et al. (2018) even pushes this boundary further, proving that in China High Speed Rail is dominantly preferred over air alternatives for

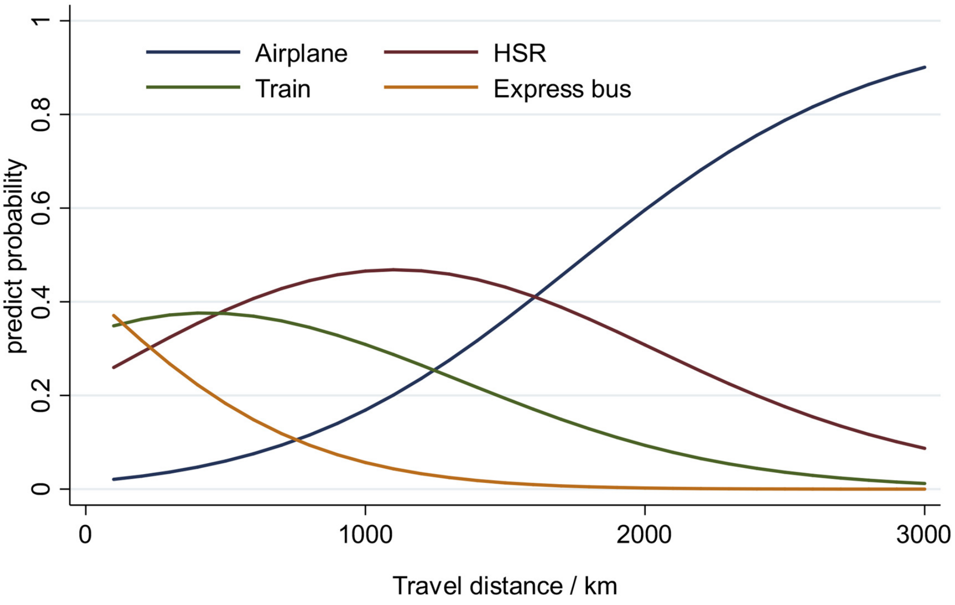


Figure 1.3: Evolution of inter-modal choice probability against distance (Source: X. Li et al., 2020)

routes up to 1300 km. Similarly, X. Li et al. (2020), investigating the factors influencing passengers' intercity choice behaviour, argue that the probability that high speed rail is chosen starts to decrease for travel distances of around 1100 km, as shown in figure 1.3. Given that the average flight distance in 2020 for intra-continental European flights is 981 km, it appears clear that high speed rail could potentially compete with air services on a considerable share of routes (Eurocontrol, 2020). Furthermore, a wide range of studies analysing the interaction between air transport and High-Speed Rail in France, Spain, and Japan generally agrees in concluding that it is rather complicated for the former mode to compete effectively with the latter on shorter distances under 500 km, over which high speed rail appears to have a considerable edge (Clewlow et al., 2014). Finally, from a behavioural perspective, the study from Malichová et al. (2022) further restates the potential for rail to substitute air services, finding that train passengers are more likely than car or air passengers to judge long-distance trips as

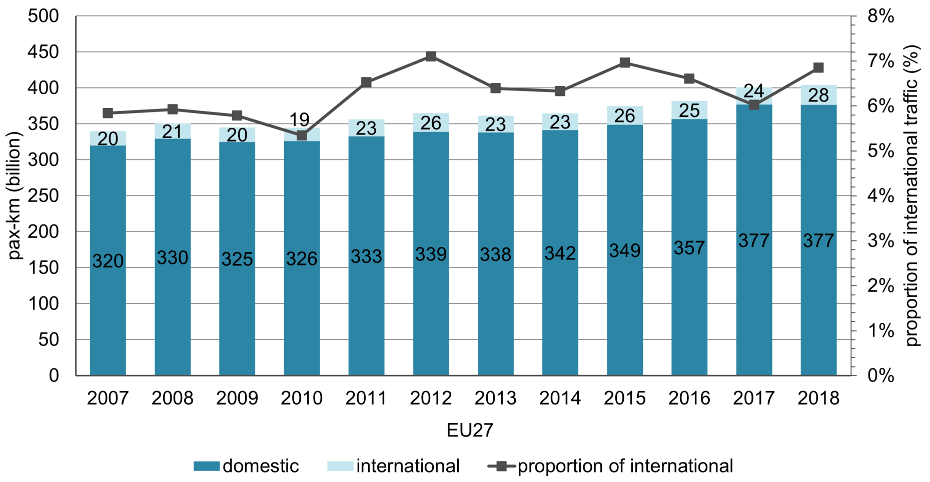


Figure 1.4: Evolution of rail passenger traffic volumes (Source: RMMS, 2020)

worthwhile.

However, despite the robust and comprehensive set of measures and initiatives taken by the European Union, especially in terms of the harmonisation of technical systems and operational regulations, and in relation to the opening up of the international rail market to competition, data shows that the modal shift in the market is still not taking place at the desired pace. In 2019, in terms of modal shares, cross-border travel by public transport within the EU mainly consisted of air passengers, with rail and bus only accounting for around 10% of the cross-border Public Transport (PT) users each (European Commission & Directorate-General for Mobility and Transport, 2021). Furthermore, in 2020 national rail traffic is still largely predominant, with international transport representing less than 10% of the total rail passenger traffic in every EU country except Luxembourg (Eurostat, 2021c). In this regard, the figures, provided in Figure 1.4, show that over the last decade the shares of international traffic have remained stationary at around 7%, with domestic traffic still representing the vast majority of the total rail passenger traffic in the EU (European Commission, 2021b). These figures prove that passengers' travel behaviour is not really changing at the expected pace, suggesting that other modes are currently more appealing for cross-border long-distance travel. Many scholars have investigated the reasons for this failed increase in the figures of international travel, providing a wide set of perspectives to answer the aforementioned question. These different perspectives are analysed in more detail in Chapter 2, and the results are briefly elaborated upon in order to identify the research gap provided in Section 1.3.

1.3. Research Gaps

The problem defined in Section 1.2 features both a scientific and a practical side, the former referring to the academic dimension of the problem and the latter relating to the more practical implications from the perspective of European policy-making and governance. Thus, the research gaps have been subdivided in these two categories and are summarised in the next two Sections.

1.3.1. Scientific Gap

The problem defined has also been widely addressed throughout the scientific literature. Over the last years, many scholars have delved into this matter, highlighting the wide range of factors that are still preventing rail from becoming a key player in the European long-distance cross-border transport market. However, the literature review in Chapter 2 has revealed a twofold research gap, consisting of an empirical and a methodological facet.

In regards to the former, a first important gap relates to the potential for rail to compete and even substitute rail. A large part of literature, whilst agreeing that rail has the potential to compete with air services, fails to address the concerns related to the actual performances of rail. Some studies, such as the one from Witlox et al. (2022), have identified some bottlenecks from a governance perspective, however the network dimension of the problem appears to be rather unexplored. Some researchers, such as Givoni et al. (2012) and Kroes and Savelberg (2019), have partially filled this gap, the former employing a GIS-based methodology to examine the potential for air-rail substitution and the latter assessing the potential for high-speed rail to substitute air services at Amsterdam Schiphol Airport using a forecasting model based on a modal split model and demand growth factors. However, both studies have a few important limitations. The former manages to capture wider geographical context, providing a global overview of the routes with potential substitution. However, considering only the geographical distance and travel demand the study fails to consider the service supply on those corridors, which is a fundamental aspect to assess substitution and competition potential (Avogadro et al., 2021; Luo et al., 2019; Zhu et al., 2018). The latter study, on the other hand, manages to take into account service supply but, focusing on a limited scope, does not allow to capture the continental dimension of the problem, which is particularly important, especially in light of the EC goals in terms of the creation of a single railway area. In this regard, Clewlow et al. (2014) highlight the paramount importance of considering wider trends using a system perspective in developing policies aimed at reducing the environmental impact of the transport sector.

Another important gap is found in the connectivity domain, in particular in relation to the rail sector. The literature review has, in fact, revealed the poignant scarcity of studies that assess the connectivity of rail transport systems. Comparative research on the connectivity of rail and air appears to be even more limited, despite Zhu et al. (2018) and Zhu, Zhang and Zhang (2019) proposed an interesting connectivity index that can be employed for both rail and air and consequently for the comparison between the two modes. However, despite the added value of having an aggregated measure of connectivity in terms of clarity and simplicity of communication, this approach also has some limitations. Connectivity, in fact, is a rather broad term that covers multiple definitions and important aspects. In this regard, the review of the literature on network science applications in transport has provided an interesting perspective on the matter, showing that interesting insights on connectivity can be obtained through the use of its mathematical tools. However, network science and connectivity have only recently started to merge, and literature on the matter has just began to flourish. In this regard, some scholars highlight how many studies have failed to obtain insightful results for practice due the missed integration of transport characteristics within such models (Dupuy, 2013; Zanin et al., 2018). Despite a wide range of studies in the field of PT and aviation have successfully filled this gap, in the case of rail research appears to be lagging behind.

Finally, a third gap relates to the specific case study considered in this research. As pointed out by the EC, the connectivity of the European rail network has neither been quantified nor thoroughly studied yet (European Parliament & Council of the European Union, 2020). Furthermore, the literature review has shown that most studies in the field focus on the Chinese networks. A notable exception is the study from Calzada-Infante et al. (2020), which, focuses mostly on analysing the network properties of the European rail network, failing to include important factors (i.e. travel time) and consequently providing a limited set of implications for policy-making. Furthermore, despite the characteristics and performances of the European and Chinese rail networks are compared and the differences analysed, no research appears to have bench-marked rail and air connectivity in an effort to provide insights into the competitiveness between the two modes from a network perspective. It is not clear what is the actual potential of competition and substitution across the continent. Furthermore, the role of different nodes and links and their performances within the network has not been analysed. In particular, the literature does not provide an overview of the current state of the market, including the identification of critical components. In conclusion, at the moment of the writing a comprehensive comparative study that analyses the performances of European long-distance air and rail transport service networks does not appear to be available yet.

This relates to the aforementioned methodological gap, which appears to be twofold. On the one hand, a thorough analysis of the literature has revealed that whilst researchers have approached the potential of rail to compete with air from a wide range of perspectives such as governance, policy-making, technical barriers and modal choice behaviour, a general lack of attention surrounds the network dimension of the problem. On the other hand, it has emerged that network science applications in the field of rail have generally focused on the network analysis, rather than on their performances, often providing limited added value to practice and policy-making. Most existing studies apply the network science indicators either disregarding the core characteristics of transport services and focusing on topological properties or failing to develop a framework on the practical implications of the results on governance. In summary, throughout the literature a methodology to quantify and compare the connectivity across the rail and air networks focused on providing a set of data-driven tools for policy-making does not appear to be available yet. This is especially true for the case study of the European continent, as further highlighted in Section 1.3.2.

1.3.2. Practical Gap

The practical gap focuses on the specific case study of the European rail and air networks focusing on the implications for practice and policy-making from the perspective of European governance bodies. As previously argued in Section 1.2, despite the strenuous efforts, the expectations relative to performance improvements and modal share growth for rail across the European continent have not been met yet. Thus, the EC deem of pressing urgency and of fundamental importance to tackle the problem from different perspectives. In particular, assessing the degree of rail connectivity and comparing it to values of the aviation sector is believed to provide important insights on the matter (European Parliament & Council of the European Union, 2020). Insights which could provide a clear overview of the actual context to policy-makers, allowing them to create more accurate forecasts and make a more efficient use of scarce resources. The Commission further highlights that the railway sector is currently lagging behind the air transport sector in terms of the development and availability of connectivity indices. In fact, the EC notes that, whilst in the aviation field a number of connectivity indices (e.g. most notably, the Eurocontrol connectivity index), currently no comprehensive railway connectivity index is available in the European market.

1.4. Research Aims

The goal of this study is to provide a comprehensive comparative overview on the performances of the rail and air service networks in the long-distance European market. In particular, this study aims to reconcile the scientific and academic cutting edge research with practice, applying the theoretical innovations in networks science to a real-life problem. Thus, this research has a dual set of goals which are explored in the next Sections.

1.4.1. Scientific Aim

Building upon the empirical and methodological gap identified in Section 1.3, this research aims to partially fill them employing an approach targeted at quantifying connectivity and identifying critical components using the tools provided by network science. The decision to approach the question through the network lens represent and substantiate the methodological aim of this study. This thesis, in fact, aims to propose a novel and reproducible approach that, ensuring the comparability between rail and air, could enrich the research in the field. Secondly, merging and intersecting different fields of study, this research aims to provide a broad overview of the current state of the competitiveness between air and rail on the long-distance transport market in Europe. Thirdly, integrating transport system characteristics, service supply data and passenger behaviour this study aims to show how network science can practically contribute to enhance the current knowledge on the connectivity of rail transport systems. Finally, this study aims to bring state-of-the-art research closer to practice, providing some data-driven tools to tackle the practical matters that are currently gripping the policy-making sector.

1.4.2. Practical Aim

Following, the practical research gap highlighted in Subsection 1.3.2, the main practical aim of this thesis relates to contributing with a novel and unexplored layer to the analysis that the European Commission is currently developing in relation to the assessment of the European railway network connectivity. The idea, is thus to support and complement rather than substitute the index developed in the feasibility study for a European “Rail connectivity index” commissioned to Ecorys under DG MOVE’s Framework Contract MOVE/A3/2017-257. To do so, this research will have a different approach, focusing on the network dimension of the problem rather then on the spatial and economical dimensions, at the heart of the index developed for the EC.

1.5. Research Questions

Based on the research gap and the research aim, previously highlighted in Chapter 1, the main research question for this study is formulated and is provided below.

“From a network connectivity perspective, what is the potential of rail to compete with air transport in the European long-distance market?”

This research plans to provide an answer to the main research question answering the following sub-questions:

1. What are the factors determining and influencing the connectivity of rail and air transport networks in relation to modal competitiveness?
2. What are the properties and performances of the European rail and air transport networks, and how do they compare?
3. What is the connectivity of the European rail and air transport networks, and how do they compare?
4. What are the critical components in terms of potential and performances?

1.6. Research Relevance

The literature review has clearly shown the considerable societal relevance of connectivity measures (Zhu et al., 2018). First of all, connectivity metrics provide important insights in terms of benchmarking and comparing the performances of transport networks (Burghouwt & Redondi, 2013). Insights which can aid policy makers and industry professionals to design coherent strategies, make evidence-informed policy decisions and plan a future-proof development of infrastructure and service developments. The crucial role of rail in the current European long-distance transport scenario further supports the relevance of this research. Using data-driven methodologies could help to better direct the substantial investments that the European rail sector is going to receive in the foreseeable future. Sensible and informed decision-making could, in turn, boost the modal share of rail providing considerable improvements in terms of sustainability to the transport sector and to society as a whole.

From a scientific and academic standpoint, the relevance of this research lies in the development of a novel methodology that blending network science with connectivity and focusing on air-rail competition and substitution could aid policy-making and governance. Intersecting these different research fields could enrich each of them singularly, providing an overview on the synergies and on the possibilities of further collaboration. Furthermore, linking state-of-the-art theory to practice could provide interesting insights in how academia and research could practically aid policy-making and governance practices.

1.7. Approach

This research aims to assess and compare the connectivity of long-distance air and rail transport services in Europe by modelling these systems as complex networks, graphs characterised by a large collection of nodes or vertices interconnected through links or edges (Newman, 2010). Such networks can be analysed and compared employing Network Science, the emerging field of study that focuses on the analysis of "network representations of physical, biological, and social phenomena leading to predictive models of these phenomena" (National Research Council, 2005, p.28). Analysing the network structure and characteristics provides, in fact, interesting insights into the topological features and the performances of the networks (Paleari et al., 2010). In particular, when applied to the context of long-distance rail and air transport systems, Network Science allows to quantify the connectivity of the network and to identify eventual bottlenecks enabling to answer recurring questions from a relatively unexplored perspective (Psaltoglou & Calle, 2018; W. Wang et al., 2020). Thus, this approach has been chosen to answer the research questions defined in Section 1.5.

As a final note, it seems important to highlight that from now on the network science notations "network", "node" and "link" are generally going to be preferred to the graph theory counterparts "graph", "vertex" and "edge", which will only be employed when talking about the mathematical definitions of graph theory. This decision follows the fact that this research relates to the practical and empirical application (network science) rather than to the theoretical and mathematical conceptualisation (graph theory) and aims to avoid any sort of confusion related to the use of distinct terms which are often used as synonyms across the scientific literature (Barabási & Pósfai, 2016).

1.8. Research Scope

First and foremost, this study focuses on analysing passenger traffic, thus any consideration about freight traffic is to be reckoned as out of scope. Furthermore, as previously touched upon, this research follows a network science approach, analysing graphs and focusing on the network dimension of the problem. Consequently, other important aspects, such as travel demand, travel behaviour and railway or aviation governance, do not represent a core focus area for the development of this study and, despite being discussed and considered, are not going to be structurally analysed. Moreover, this research focuses on the competition between rail and air travel, and thus excludes both the inter-modal cooperation between the two aforementioned modes and the competition of rail with any other mode, like road based modes such as coach or car. Finally, this research mainly tackles questions related to international traffic, to medium- and long-distance traffic (routes between 100 and 1.500 km) and to cross-border traffic. Thus, in modelling the rail networks greater attention is going to be put on high-speed lines, main national/international lines and night train lines. In this instance, it is important to highlight that regional lines are not going to be taken into account as they are deemed to be out of scope. This will also impact the level of aggregation of the geographical scope, which will reflect the broad interest of the research, and will be characterised by a number of approximations, in terms of links and nodes included. This is also going to be reflected by the number of flight connections modelled, which will follow the characteristics of the modelled rail network in order to maximise their comparability. Finally, the focus of this study is limited to the spatial relationships in the networks, thus temporal coordination is disregarded. In fact, considering the temporal dimension would considerably increase the complexity of the model, which would become dynamic. Thus, considering the broad scope of the research, this dimension is not considered. However, it is important to point out that this is an important dimension to consider, especially when analysing transfers, which represent a crucial dimension of long-distance travel. More details on the characteristics of the model can be retrieved in Chapters 3 and 4 and the geographical scope is described in Section 4.1.

1.9. Report Structure

This report is structured in four main Chapters. Chapter 2, provides a review of the literature on connectivity definitions, network science applications in the transport field and European long-distance transport market, answering research sub-question 1. Next, in chapter 3 the methodology is described and explained. In particular, the model setup is outlined, answering research sub-question 2, and the connectivity measures employed in networks science are reviewed and selected, answering research sub-question 3. The methodology is thus applied to the case study of the European Air and Rail Transport Networks in Chapter 4. Then, the results, answering research sub-questions 4, 5 and 6, are illustrated and discussed in Chapter 5. Finally, the conclusions are presented in Chapter 6.

2

Literature Review

Three main research areas of interest have been identified and reviewed. A first one concerns the broader context of this research, referring to the literature on the European long-distance transport market. A second one focuses on the term connectivity and its various definitions. Finally, a third one draws from the research on network science and its applications on complex transport networks. This review aims to tackle and analyse state-of-the-art literature on these three topics, shedding light on the current state of the academic research, grasping possible knowledge gaps and highlighting the position of this study within the literature. Firstly, Section 2.1 focuses on the long-distance transport market highlighting the broad thematic context of the research. Next, some definitions of connectivity are provided in Section 2.2, introducing the concept at the core of this thesis, highlighting the determinant factors and clarifying the definition that is given in this study. Finally, Section 2.3 is devoted to review the literature on network science applied to transport systems from a methodological perspective.

2.1. Rail and Air Transport in the European Long-Distance Market

As highlighted in Section 1.1, rail transport is increasingly becoming a central topic in the transport debate, gaining a widespread attention among scholars. The developments in the rail sector in Europe have steeply risen over the last decades, revolutionising its role in the wider transport context (Seidenglanz et al., 2021). Rail, in fact, currently represents an extremely promising field due to its potential to contribute to the decarbonisation of the transport sector, limiting greenhouse gas emissions and energy consumption. In particular, governmental bodies, such as the EC, saw this as a possible solution to the recurrent and persistent environmental issues embedded in the structure of the current transport system, pushing towards the substitution of short haul flights with rail alternatives. Thus, researchers have devoted much attention to this topic, developing a large body of literature and widely covering the many matters and questions that arose over the last decades. In an effort to shed light on the feasibility of such a project and on the the possible challenges that this process could bring about, the research on the topic has been pointed towards two main directions, following two diverging approaches. On the one hand, some efforts have been targeted at identifying the crucial barriers and bottlenecks that are still hindering this process. On the other, further endeavours have been directed at assessing the eventual potential for competition, substitution and cooperation between the two modes. The former efforts focus on allowing to remove the obstacles that are still preventing the aforementioned modal shift, pushing towards a more rail-centric long-distance transport market and revitalising the momentum of this process. Conversely, the latter endeavours attempt to more radically question if, and under which conditions, railways are actually able to compete or cooperate with air transport in this specific market.

2.1.1. Potential for inter-modal competition, cooperation, integration and substitution

From a first overview, literature seems to generally share a rather positive outlook on the potential of rail, and especially high-speed rail, to compete with air transport. Adler et al. (2010), assessing the rail modal shares in a competitive long-distance transport market through a game theory approach, find that the rail system could manage to attract almost 25% of the passengers in the below 750 km range, figure that sinks to 9% of the market in case of longer distances. Following these results, the authors suggest that the considerable potential of this mode might even justify the extremely high infrastructural costs of the Trans-European high-speed rail projects, concluding that in order to maximise the social welfare the development of the high-speed rail network should be encouraged (Adler et al., 2010). Behrens and Pels (2012) also recognise the great potential of high-speed rail to compete with airline services, deeming that the former could likely prove to be a viable substitute for the latter. The authors analysing the travel behaviour by means of multinomial and mixed logit model estimations, conclude that competition in the London–Paris corridor is expected to fall in the long run leaving high-speed rail as the main competitor (Behrens & Pels, 2012). Another important type of competition is the intra-modal one, which is widely regarded to help lower the prices of rail services, thus making it more attractive and increasing its modal shares (Beria et al., 2019).

A different approach is followed by Albalate et al. (2015), who through a supply oriented empirical analysis explore the potential for cooperation and competition between rail and air transport, identifying the impacts of High-Speed rail on air service frequencies and capacities. Their findings show that air transport capacity for a specific OD pair tend to reduce once a high-speed rail alternative enters the market, thus confirming that high-speed rail services can directly compete with air transport alternatives. However, the authors also highlight that there is a noteworthy potential for cooperation between the two modes, with high-speed rail providing feeding services to long-haul flights, especially in hub airports (Albalate et al., 2015). In this regard, Román and Martín (2014) argue that, despite the fact that air and rail passenger transport are often considered merely as substitutes, they could be seen as complementary services.

Similarly, Socorro and Viecens (2013) highlight that, especially in the European Union, a widespread increase in interest towards inter-modality is slowly replacing the traditional view that sees airlines and high-speed transport only as competitors. This study finds that in general air-rail integration is beneficial for the companies involved, as it implies a spike in their market power (Socorro & Viecens, 2013). In regards to the potential implications of integration on social welfare the results show that the overall welfare is generally enhanced when the integration is focused on capacity-constrained airports. Furthermore, another interesting perspective on the possible impacts of inter-modality on the transport market is provided by A. Zhang et al. (2019). The authors state that, as a result of the increased availability of inter-modal services due to competition, high-speed rail could trigger traffic redistribution phenomena on air traffic, possibly provoking a raise in congestion levels in major airports with high degrees of air connectivity, whilst causing drops in traffic in smaller airports. Socorro and Viecens (2013) share this view, suggesting that in many cases the integration of the two services might cause an increase in demand, and consequently arguing that from an environmental perspective the substitution of air transport with rail services is to be preferred where possible. Avogadro et al. (2021) also supports this statement, arguing that substitution will probably lead to a reduction in demand following the decrease in supply and the worsening of transport conditions.

Concerning the substitutability of air routes with a rail alternative, Avogadro et al. (2021) estimate at 26.5 million (or 3.02% of total intra-European) the currently offered seats that could be replaced assuring similar travel times for travellers. In this regard, Behrens and Pels (2012) suggest that a complete substitution of air connections with high-speed rail alternatives will most likely only take place if and when full integration of high-speed routes and airlines' networks is going to be realised. On a different note, Avogadro et al. (2021) foresees that, due to the recent adoption of policies aimed at reducing the emissions of the aviation sector, the cancellation of all the air routes where an effective alternative exists is likely to be forthcoming. However, many authors highlight that there might be some underlying risks under this process of substitution. Socorro and Viecens (2013), for example warn that substitution, as much as it can positively impact the environment, also leads to the elimination of inter-modal competition from the market. Avogadro et al. (2021), agrees that inter-modal competition is bound to be reduced as a consequence of the cancellation of air routes, arguing that this will inevitably lead to increased fares and lower service quality. Thus, Socorro and Viecens (2013) highlight that a paramount impact of rail-air integration is the fostering of inter-modal competition, which means lower prices, a wider service offer for passengers and overall higher levels of social welfare.

2.1.2. Modal Competitiveness Factors

As previously mentioned in Section 1.2, despite all the efforts made by the European Union, and despite literature widely supports the idea that rail can and will substitute short-haul flights, the modal shares of railway in the European long-distance cross-border market are not really increasing at the expected pace, and the substitution of short-haul flights with rail services still looks far from becoming reality. Literature widely agrees that rail is current unable to compete on a level playing field with the other transport markets, especially with road in the medium-distance and air in the long-distance, with the exception of a limited number of corridors. In particular researchers often argue that rail transport has not managed to really affect road and air transport shares yet due to its weak "modal competitiveness". Román and Martín (2014), for instance, argue that more specific policies targeted at removing some important barriers are required in order to make inter-modal options concretely attractive for passengers. A large number of researchers has consequently analysed the possible causes for these failed expectations, identifying the factors influencing the modal competitiveness of rail in the current European scenario through a diverse set of perspectives. Notable approaches include Game Engineering, Transport Policy, Transport Governance, Travel Behaviour, Transport Supply Analysis, Spatial Analysis and Transport Economics.

The first two crucial elements that heavily hinder the growth of the train market relate to the higher travel times and travel costs. Analysing the matter from a governance approach, Witlox et al. (2022) suggest that from a passenger's perspective the generally noncompetitive journey times and journey costs are still two paramount obstacles that need to be overcome in order to see rail flourish. From an econometric perspective, Clewlow et al. (2014) also corroborates these findings suggesting that reduced journey costs and times in the train sector would allow for the growth of its market shares to the detriment of air transport. In particular, the study finds that the presence of high-speed rail services, characterised by lower journey times, negatively influence air travel demand, whilst highlighting that low-cost carriers contribute to an increase in air traffic (Clewlow et al., 2014). In fact, literature generally finds that longer door-to-door travel times lead to a reduction of the likelihood that a certain mode is chosen (Witlox et al., 2022). In this regard, Dobruszkes et al. (2011), through a supply-oriented study, maintain that travel time is a crucial factor for HSR to successfully compete with air services. Furthermore, many authors point at the much higher travel costs that rail services feature in comparison to air transport, suggesting that these could play an important role in passenger choice, representing a key bottleneck (Clewlow et al., 2014; Witlox et al., 2022). In particular, Witlox et al. (2022) identify the differences in business model as one of the causes of this, suggesting that the rail market needs to shift its focus on passengers' needs to make rail more attractive. This need to shift the focus towards passengers' requirements and needs is also shared by Bergantino and Madio (2020), who underline the importance of revenue management practices, arguing that demand for high-speed rail could be captured through the introduction of new behavioural-based pricing schemes. Another interesting element relates to the role of inter-modal competition in lowering ticket prices. In this regard, Bergantino et al. (2015), analysing the impact of competition on airline pricing behaviour, find that, in case of direct competition between air and HSR services air fares have demonstrated to significantly drop.

Despite appearing as the most important factors, travel costs and travel times are not the only ones. Bergantino and Madio (2020) and Witlox et al. (2022) highlight that travel comfort is another crucial factor influencing modal choices. In general, higher comfort is found to reduce the importance of travel time whilst lower comfort is found to be often preferred by price-sensitive travellers when paired with lower prices. This term, thus, evokes another important dimension of modal choice, the personal perception of the journey as experienced by the passengers. The subjective psychological and attitudinal perspective, in fact, also influences the aforementioned travel time and travel costs. For instance, air journeys tend to look shorter than comparable rail journeys to the eye of a passenger because the in-vehicle time is generally used as a metric. However, the actual journey time also includes, access/egress and waiting times, which are generally much higher for air. Similarly, when accounting for the costs of a journey travellers more easily perceive the price spent in the moment of the travel rather than some fixed costs such as subscriptions. Thus, a passenger using the car will account for fuel or tolls, but probably will not consider the impact of all the fixed costs (car price, maintenance, taxes, etc.). Analysing passengers' preferences through a discrete choice experiment, Román and Martín (2014) argue that ticketing, fares integration and schedule coordination, aimed at facilitating the purchase of through tickets, are also important factors. In fact they could guarantee passengers rights and reduce connecting times, in turn making rail more attractive.

Another factor literature widely deems influential in modal competitiveness is frequency. In this regard, Behrens and Pels (2012) find that travel time and frequency are the main determinants of travel behaviour. Results, which are also supported by the findings of Bergantino et al. (2015), who argue that service capacity and frequency are two core strategic variables. However, currently the supply and service capacity of rail in the European long-distance market are retained to be lacking, in particular in terms of High-Speed lines and their frequencies. Seidenglanz et al. (2021), in fact, analysing the evolution of the rail service supply trends between 1990 and 2019, argue that there are considerable limitations due to the supply of cross-border long-distance rail services. In particular the two main trends are found. The former regards a large increase in supply over the main railway corridors for short or medium-distance, focusing on enhancing the frequency of the connections. The latter in parallel shows a constant decrease in the supply of long-distance, night and periodic train services, which have largely been curtailed or cancelled (Seidenglanz et al., 2021). However, other than service capacity, literature also stresses the importance of infrastructure capacity. The coveted growth in railway shares, in fact, will probably lead to a growth of frequencies which will in turn limit the availability of infrastructure capacity (Rotoli et al., 2016). Serafimova et al. (2022), from a policy perspective, argue that this represents a crucial barrier preventing an increase in the modal share of rail passenger traffic, adducing as causes of this the high track access charges, the frequently lacking infrastructures and the overly congested network. In fact, the increase in frequencies and service capacity is not possible without a parallel improvement of the infrastructure conditions. In this regard, many researchers highlight the need for investments directed at enhancing the infrastructure capacity, building new infrastructure and enhancing the use of existing capacity through improved capacity allocation and traffic management (Serafimova et al., 2022). In general, what is widely deemed to be missing in this regard is an international comprehensive and coherent vision and planning for European Rail. From a governance perspective, many underline the importance of moving away from the current focus on national strategies to shift towards a common European vision. Literature, in fact, points out that the lack of coordination in planning and operations between stakeholders and the lack of international harmonisation of traffic control are important obstacles to the realisation of the ambitious goals for European rail. In terms of possible solutions, Witlox et al. (2022), highlight that the multilayered structure of the railway sector, especially in terms of the separation from the national and the international level, represent a crucial barrier to the success of European railways. In this instance, Serafimova et al. (2022) suggest to develop an overarching European Masterplan for rail infrastructure, resembling the solutions adopted for air traffic management that could allow to coordinate the efforts directed at resolving the main bottlenecks.

Finally, C. Chen and Vickerman (2017) approach the question from a spatial planning and economics perspective, identifying among the key obstacles that negatively influence rail modal choice in medium- and long-distance trips the failure to understand the importance of connectivity between the high-speed railway network and local networks. In this regard, also Martín et al. (2014) suggest that accessibility, despite being often overlooked, is a factor that greatly influences passengers' behaviour. In particular the authors highlight that spatial competitiveness, in terms of accessibility and connectivity of airports and high-speed rail stations, plays a crucial role in modal choice, arguing that improving terminal accessibility is a key strategy to increase the attractiveness of modes and to influence the market shares of competing modes (Martín et al., 2014).

2.1.3. Conclusions

In conclusion, this first part of the literature review has shown that many factors influence the modal competitiveness of rail with air services on the long-distance European market. To summarise and highlight the relationships between the different factors a causal loop diagram of the system dynamics in inter-modal competition is provided in Figure 2.1. System dynamics is a methodology often employed to model complex systems (Fontoura & Ribeiro, 2021). In particular, by reproducing their structure and patterns of behaviour this approach allows to simulate the interactions among the different elements within the system inferring possible outcomes and future scenarios (Bala et al., 2017). In this specific case, however, this methodology is particularly useful to highlight the causal relationships among the different factors summarising the findings of this section. In the diagram, causal relationships are indicated by a directed arrow going from the influencing variable towards the influenced one. Each arrow is associated either to a positive or negative sign, which represent respectively directly or inversely pro-

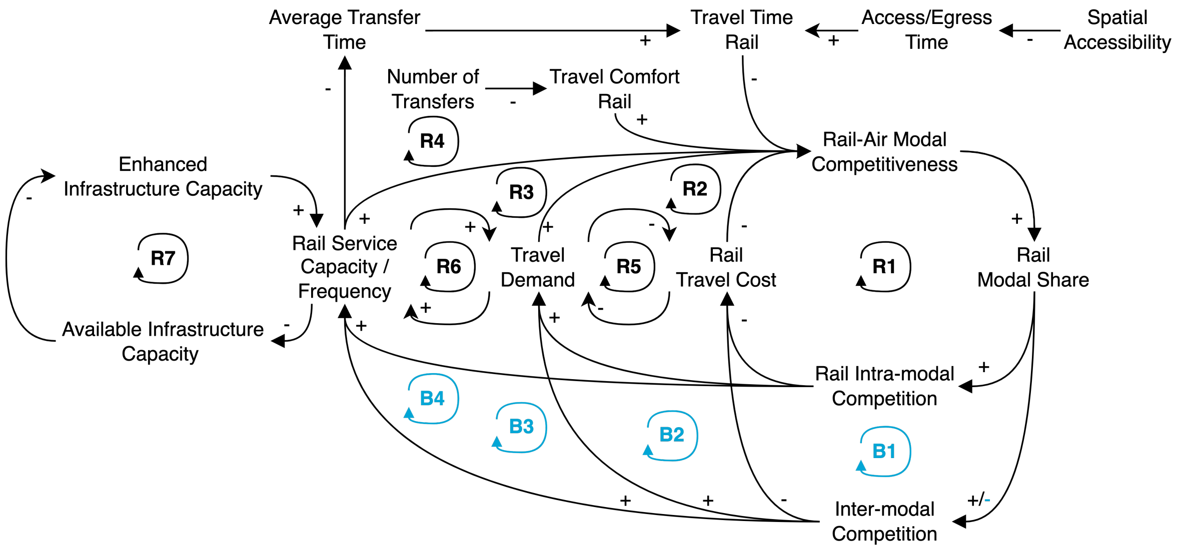


Figure 2.1: System Dynamics of Modal Competitiveness

portional relationships between variables. The only exception is the connection between rail modal share and inter-modal competition, which presents both positive and negative signs. This is due to the fact that these two variables have a curvilinear (bell-shaped function) rather than a linear relationship. In fact competition is believed to reach a maximum when modal share is equally distributed between modes and to progressively and symmetrically decrease when moving towards higher and lower values of modal share. This also impact the dynamics of the system. Multiple consecutive connections create, in fact, feedback loops that can be either positive or negative depending on the presence of an even (including zero) or uneven number of negative relationships within the loop. Positive feedback loops are also defined reinforcing loops as the growth of a variable would feedback into even more growth of the same variable. On the other hand, negative feedback loops are also described as balancing loop, as these tend to drive the system towards an equilibrium. The twofold sign between rail modal share and inter-modal competition, thus, implies that Feedback Loops B1, B2, B3, B4 are complemented by a parallel series of Reinforcing Loops R8, R9, R10, R11 that are not shown in the diagram for sake of simplicity. In fact, these have exactly the same effects of loops R1, R2, R3, R4 respectively and are thus assimilated to them and treated as one. A complete overview of the Feedback Loops depicted in Figure 2.1 including a brief description is provided in Table 2.1. Finally, it seems important to highlight the assumptions made in developing the causal-loop diagram:

- It is assumed that the findings from literature, which are often related to specific case studies, can be universalised into a general overview for the entire European market.
- It is assumed that intra-modal competition is present in the rail market. As this is not always true, in its absence, the related variable should be disregarded considering inter-modal competition only.
- It is assumed that inter- and intra-modal competition share the same set of affecting factors and influenced variables. This in reality is not true, as many elements affecting the latter are disregarded in this representation for sake of simplicity. On the other hand, the literature review presented in this chapter highlights that the effects of both are widely deemed similar.
- The diagram focuses only on air and rail, considering competition and competitiveness only between these two modes. Thus, other shared means of transportation such as bus or car-pooling are not

considered. Literature, in fact, finds that a shift towards rail is less likely to come from these modes (Bergantino & Madio, 2020).

- Private modes are also excluded as they are assumed to be attracting a different portion of the market. Research, in fact, generally agrees in believing that a shift from car to rail is particularly complex to achieve (Borsati & Albalate, 2020).
- It is assumed that rail and air can compete on a level playing-field. However, currently this is true only in few cases.
- Hard infrastructural limitations are not considered as it is assumed that infrastructure capacity can always be improved, either by optimising it or by building new infrastructure. However, in reality there are considerable economic, temporal and, in some cases, even geographical constraints related to the possibility of building new capacity. Furthermore, the present efforts of the EU, clearly prove that capacity optimisation often requires considerable technical know-how, harmonisation across systems and collaboration among stakeholders, which is not always possible or straightforward.
- For sake of simplicity the causal relationships between frequency and travel demand and between travel demand and travel cost are assumed to be direct. In reality there are other elements in between them, more details are provided in the description of R5 and R6 (see Table 2.1).
- Following the scope of this chapter the diagram takes the perspective of rail and assesses how its competitiveness relates to the system. It is assumed that rail is currently less competitive and aims to increase its competitiveness.

Table 2.1: Feedback Loops Overview

Code	Name	Description
R1	Travel Cost Attenuation Loop	Higher intra-modal competition leads to lower travel costs, increasing the modal competitiveness of rail and consequently enhancing intra-modal competition
B1	Travel Cost Balancing Loop	After the maximum level of inter-modal competition is reached, further enhancements of rail modal shares start triggering a process of reduction of inter-modal competition, consequently causing an increase in travel costs and a decrease of rail modal competitiveness
R2	Travel Demand Enhancement Loop	Increased intra-modal competition leads to higher travel demand, positively affecting the modal competitiveness of rail and consequently enhancing intra-modal competition
B2	Travel Demand Balancing Loop	After the maximum level of inter-modal competition is reached, further enhancements of rail modal shares trigger a process of reduction of inter-modal competition, consequently causing a decrease in travel demand and a decrease of rail modal competitiveness
R3	Frequency Enhancement Loop	Increased intra-modal competition leads to higher frequencies, positively affecting the modal competitiveness of rail and consequently enhancing intra-modal competition
B3	Frequency Balancing Loop	After the maximum level of inter-modal competition is reached, further enhancements of rail modal shares trigger a process of reduction of inter-modal competition, consequently causing a decrease in frequencies and a decrease of rail modal competitiveness
R4	Travel Time Enhancement Loop	Higher intra-modal competition leads to a growth in frequencies, reducing transfer times and consequently also travel times and thus enhancing the competitiveness of rail
B4	Travel Time Balancing Loop	After the maximum level of inter-modal competition is reached, further enhancements of rail modal shares trigger a process of reduction of inter-modal competition, consequently leading to lower frequencies and higher transfer times. Thus, positively affecting the growth of travel times and decreasing the competitiveness of rail
R5	Law of Demand	Lower travel costs lead to an increase of travel demand, which in turn increase the supply/-frequency enhancing competition and lowering the travel costs
R6	Mohring Effect	Increasing travel demand implies a growth in frequencies, which reducing the transfer times lead to lower generalised travel costs, further increasing demand
R7	Infrastructure Capacity Enhancement Loop	Higher frequencies imply a reduced amount of capacity available leading to capacity enhancements targeted at further increasing the frequencies

First of all, this diagram, portraying how the different elements interact in the system, allows to infer which factors determine modal competitiveness and how they influence it. In particular, the causal loop diagram developed highlights that the main factors directly influencing the modal competitiveness of rail appear to be travel costs, travel demand, travel time, travel comfort and service frequency. Thus, these are deemed to be the crucial factors that should be taken into account when defining the connectivity index. A more in-depth description of their characteristics and on their use in this study is further provided in Section 3.1.2.

Furthermore, what appears to be particularly interesting from this schematisation is the dual dynamics that follow from the aforementioned twofold relationship between modal share and inter-modal competition. To better analyse that, two different scenarios are going to be considered. A first one where both inter- and intra-modal competition are present and a second one where only inter-modal competition is present, as, at present, the European rail market does not widely support and allow intra-modal competition in the sector. The former scenario, is characterised by two parallel sets of feedback loops. Whilst the internal one (consisting of R1, R2, R3, R4) is always characterised by reinforcing loops, the external one features a set of reinforcing loops (consisting of R8, R9, R10, R11) when rail has a lower modal share than air, which turn into a set of balancing loops (consisting of B1, B2, B3, B4) when the modal share of rail overtakes the one of air. This implies that, in this case, that after reaching the maximum inter-modal competition, competitiveness of rail should keep growing thanks to the reinforcing loop of intra-modal competition neutralising the balancing loop of inter-modal competition. This process could allow rail to progressively replace air services until reaching a state of complete substitution. On the other hand, the latter scenario, is constituted only by the external dual loop. This means that without intra-modal competition in the market, the shares of rail would keep growing until reaching a state of equilibrium and stabilising around that. Thus, this diagram suggests that a state of inter-modal competition in the market, which relies on an effective modal competitiveness, generally tends to maximise inter-modal competition. This is jeopardised only by the presence of intra-modal competition, which through its reinforcing loop tends to foster more and more internal competition improving the competitiveness of rail and thus reducing inter-modal competition. Thus, it is possible to conclude that in the current state, without external interventions (e.g. fuel prices regulation, emission charges application, short-haul flight ban) rail can successfully substitute air transport only in the presence of intra-modal competition in the rail market. In its absence the system appears to be naturally tending to a market equilibrium with modal share being equally split between the two modes.

Finally, to validate and to evaluate the sensitivity of the choices made in this Section similar case studies, analysing rail and air competition within the European long-distance market, are analysed. The selected influencing factors appear to be in line with the explanatory variables employed by Kroes and Savelberg (2019) to model air-rail modal choice behaviour in the European long-distance market. The authors, in fact, argue that the most important variable in determining passenger's choice behaviour is travel time, followed by the number of daily departures (frequencies), ticket prices and travel comfort. Furthermore, the crucial role of travel demand and its intertwined relationship with passenger's choice behaviour are highlighted throughout the study. The approach from Kroes and Savelberg (2019) appears to be in line with the system dynamics developed in this Section. in particular, it is interesting to highlight how the influencing factors Travel Time, Travel Comfort, Travel Costs and Frequency could be also defined as passengers choice determinants, whilst travel demand could be viewed as the outcome of such passenger behaviour.

2.2. Connectivity Definitions

Science aims to explain reality by connecting ideas and evidence employing critical or scientific thinking. The former concept represent the core element of human thinking. An idea is, in fact, defined as a structured shape that can be experienced in thought and expressed in language (Fisher, 1976). Logic is the science that connects thinking and language and that allows to coherently and successfully relate different ideas to evidence. Applying logic in complex reasoning processes allows the thinking to be factual and to represent reality in the most reliable way possible. Thus, in order to develop a logically sound reasoning it is fundamental to clearly and unequivocally relate words and definitions to ideas and facts. Especially because words can often have different meanings and interpretations depending on the context and field in which they are used. The following paragraphs aim to provide a clear understanding on the central complex idea that is at the core of this research and on its definition.

2.2.1. Connectivity

The complex idea that represents the backbone of this thesis is the one expressed through the term “connectivity”. Across the literature this term is used in a range of different ways, and it is hardly possible to find a commonly accepted definition. According to the Oxford English Dictionary, connectivity can represent both “the state of being connected or interconnected” and “the degree to which two things are connected”. The former definition is employed in the field of transport by some authors, such as Clewlow et al. (2012), who describe with air-rail connectivity the state where the two modes are mutually interconnected. In the same way, many authors employ the latter definition, measuring this static state of being connected by ascribing a wide range of weighting factors (Bombelli, 2020; Derrible & Kennedy, 2010a; Luo et al., 2019; Paleari et al., 2010; Psaltoglou & Calle, 2018; Z. Xu et al., 2020). Given that this research aims to analyse and compare the degrees of connectivity of the rail and air networks, this section will focus on reviewing and analysing the latter group of studies, highlighting which factors are employed across the literature to measure connectivity.

When defining connectivity, it is important to distinguish it from accessibility, as often these two terms are most often employed by researchers either as synonyms or with overlapping significance. In particular, Ortega Pérez et al. (2011), employ the term connectivity to describe the state of being connected and the term accessibility to describe the degree of connectivity. It could be indeed useful to differentiate between the aforementioned two definitions using connectivity for the former and accessibility for the latter. However, it was decided to avoid the use of accessibility, as this term is deemed to have an enduring relation to the land-use and spatial dimensions, which do not represent the core focus of this research. In fact, following the definition of Geurs and van Wee (2004), accessibility represents the “extent to which land-use and transport systems enable (groups of) individuals to reach activities or destinations by means of a (combination of) transport mode(s)”. This is also demonstrated by the mathematical formulation of accessibility indicators, which almost always include demographic or economic parameters of some sort. Furthermore, it is noticeable that the term accessibility is used more often for more limited geographical scopes, especially at the regional or urban level, and for modes such as car and PT. On the other hand, the term connectivity is a far more commonly used indicator in long-distance transport, being especially diffused in the field of aviation. In conclusion, the main difference between the two terms appears to be the fact that accessibility has a relation to the attractiveness of the destination and/or to the spatial geography and land use whilst connectivity mostly takes into account links, connections and their quality.

2.2.2. Connectivity in Graph Theory

Due to the relative vagueness of its definitions and underlying set of meanings, it could be argued that the term connectivity is rather broad, assuming a diverse set of shapes depending on the context within which is employed. For example, in graph theory, the mathematical scaffold behind network science, the term connectivity is associated to graphs (or networks), describing the degree to which a certain graph is (inter)connected. A graph is defined by Wilson (1972) as a non-empty finite set of elements named vertices (or nodes), and a finite set of distinct unordered pairs of distinct elements of vertices named edges (or links). Thus, each vertex represents a component in the system, and each edge represents the interaction between specific pairs of edges (Barabási & Pósfai, 2016). Such graph is said to be connected when each pair of distinct vertices is connected by a path, an alternating sequence of distinct vertices and distinct edges (Trudeau, 1993; van Steen, 2010; Wilson, 1972). In this context, connectivity is thus defined by Trudeau (1993) as the extent to which a graph is connected. Building upon this, the vertex connectivity of a graph is defined as the minimum number of vertices whose removal results in the disconnection of a previously connected graph, whilst the edge connectivity is defined as the minimum number of edges whose removal results in the disconnection of a previously connected graph (Trudeau, 1993; van Steen, 2010; Wilson, 1972). These metrics, however, are traditionally employed more to evaluate graph robustness and vulnerability rather than connectivity defined as the strength or degree of interaction between two nodes (Wei et al., 2014).

2.2.3. Connectivity in Transport Systems

The concept of connectivity is a crucial element in transport network analysis and is widely studied especially for air and public transport networks, whilst for rail and road networks research on connectivity appears to be more limited (Zhou et al., 2019). Rodrigue (2020) highlights that this concept is especially important in analysing transport systems due to its threefold effect on the economic, network and spatial dimensions. However, also in this case it was not possible to find a common and universally agreed definition across the literature. In fact, a wide range of definitions of the transport systems' connectivity can be found, depending in particular to the specific mode to which it is applied. (Zhu et al., 2018) highlights that in the transport field this concept was first introduced to assess the importance of an airport in relation to the degree to which this is connected to the rest of the airport network. On the other hand, a generic definition is given by Rodrigue (2020) who, drawing from graph theory, broadly defines connectivity in transport systems as the extent to which flows of passengers (or freight) from a certain node can reach other nodes directly or indirectly via another node or a series of nodes. Similarly, Handley et al. (2019) define public transport connectivity as the metric that measures the quality of the supplied transportation service. In this regard, Zhou et al. (2019) point out that public transport connectivity generally focuses on the quality and number of transfers and spatial accessibility. Despite the abundance of literature on transport connectivity, an important gap was found in the rail sector, and especially in the European research context, probably due to the fragmentation of the continental rail network into separate national components. The only definition of rail connectivity retrievable across the literature is, in fact, provided by the study from Z. Xu et al. (2020). The authors employ the term connectivity to describe the critical performance assessment of railway systems which measures how well stations are interconnected in relation to vehicular movements and commodity flows. However, none of these definitions does really clarify how this extent or degree of connectivity should be measured. Slightly more precise in this regard are the definitions given throughout the aviation sector, which, following the analysis from Burghouwt and Redondi (2013), appears to be characterised by two main perspectives on connectivity. On the one hand, the accessibility perspective (in-direct connectivity), which measures “the number and quality of direct and indirect air travel connections available to the consumer at a certain airport” (Burghouwt & Redondi, 2013). On the other, the centrality perspective (hub connectivity perspective), which considers “the number of transfer opportunities available via a specific airport” (Burghouwt & Redondi, 2013). It is interesting to notice how in this case the word accessibility assumes a different connotation due to the different field of application of the term, proving once again its dynamic and multi-layered essence.

2.2.4. Connectivity Metrics

Zhu et al. (2018) highlight that travellers generally show a wide range of preferences in terms of destinations, travel distance, travel time and other dimensions, which are captured by specific metrics. These are generally case specific and tend to vary considerably depending on the precise scope of the study and on the goal of the indices. As previously mentioned, a considerable amount of literature on connectivity is present in the aviation sector, as opposed to the rail field which features a very limited number of papers on the topic. Despite the factors used to calculate connectivity generally appear to be mode dependant, from a thorough analysis of literature a general basis common to all could be identified. Thus, to overcome the lack of research on rail networks connectivity and to provide a broader perspective on the topic, some of the elements and metrics employed in defining connectivity metrics for other networks, such as Air, PT and Road, are taken into account. An overview of the metrics and components most commonly employed across the literature to measure connectivity is provided in Table 2.2. In particular, each metric is defined and briefly described, their aggregation level and the network components they relate to are highlighted and some references in the literature are provided. The aggregation level is useful to discern between independent metrics and dependent indicators, whose definition relates to multiple independent metrics. Moreover, the network component shows on which level the metric or indicator is generally applied.

Across the literature it was possible to identify distinct layers of metrics based on their level of aggregation. Starting from very basic components, to more complex and refined metrics, which put in relation different elements and variables. A basic overview of important factors is provided by Rodrigue (2020), who underlines the influence that factors such as generalised travel costs, capacity, frequency and spatial accessibility have on connectivity. This is also confirmed by this review, which widens the perspective including more specific metrics. Overall, the common basic elements most frequently employed to compute connectivity for different modes are found to include travel time, distance, velocity, capacity, frequency and travel demand. These are widely employed across most connectivity indicators, if not independently combined together to create more complex variables. For instance, Luo et al. (2019) measure connectivity using a metric rooted in closeness centrality, which including travel time, number of transfers and headways aims to measure the travel impedance across the stops of different tram networks. Takebayashi (2015) employ a similar metric defined as generalised travel costs to take into account the resistance against inter-modal services, measuring the extra travel time and travel costs incurred by passengers having to transfer from one mode to another. Moreover, analysing PT networks, Handley et al. (2019) employ a connectivity index rooted in travel times and in the number of opportunities per stop, where the latter term is an aggregated accessibility index representing the cost, quality and quantity of opportunities available at a certain stop. Similarly, Kaplan et al. (2014), assessing the equity in public transport provision, proposes to measure PT connectivity as a comprehensive impedance measure comprising door-to-door travel time, frequency, service reliability and the possibility of “seamless” multi-modal transfers and further weighting the individual components in relation to their relative importance to passengers. Also (Zhu et al., 2018) employ the travel time between each OD pair as the basic element of the connectivity indicator, transforming the classical shortest path length accessibility model into a quickest path length accessibility model. In this case, the usage of travel time stresses the importance of this factor as a component of the more articulated connectivity metric developed. In particular, (Zhu et al., 2018) develop the Connectivity Utility Model (ConnUM), a model that aims to represent connectivity as the aggregate utility for passenger to choose specific terminals, including many connectivity sub-components, such as the availability of seats (service capacity), the trip duration (travel time) and the quality of transfers. This model is particularly interesting for this study due to its focus on the long-distance connectivity of multiple modes, as it is used to analyse both rail and air transport (Zhu et al., 2018) and inter-modal trips (Zhu, Zhang & Zhang, 2019). In fact, most of the existing research focus on analysing the connectivity for single transport modes, not allowing to compare the connectivity across different modes. However, to do so, it is paramount that connectivity is defined in a homogeneous way, taking into account the different characteristics of each mode and

Table 2.2: Overview of the Determining Factors/Components of Connectivity across the Literature

Metric	Short Definition	Aggregation Level	Network Component	References
Travel time	Function of in-vehicle time on the route, in some cases considering also transfer and waiting times, with or without specific Value of Time	Disaggregated	Link	Handley et al. (2019), Luo et al. (2019), Nieße and Grimme (2015), Paleari et al. (2010), Psaltoglou and Calle (2018), Zhu et al. (2018, 2019), Zhu, Zhang, Zhang et al. (2019)
Distance	Function of the distance, in some case taking into account minimum euclidean distance	Disaggregated	Link	Allroggen et al. (2015), Calzada-Infante et al. (2020), Paleari et al. (2010), Psaltoglou and Calle (2018) and Z. Zhang et al. (2015)
Velocity	Function of speed of the connection, sometimes compared to the average speed	Disaggregated	Link	Nieße and Grimme (2015), Psaltoglou and Calle (2018), Williams and Musolesi (2016), Zhu et al. (2018, 2019), Zhu, Zhang, Zhang et al. (2019)
Service Capacity	Function of the number of available seats on the route (or maximum tonnage in case of freight)	Disaggregated	Link	Bombelli (2020), Psaltoglou and Calle (2018), Rodrigue (2020), Zhu et al. (2018, 2019), Zhu, Zhang, Zhang et al. (2019)
Frequency	Function of frequency or headway	Disaggregated	Link	Allroggen et al. (2015), Calzada-Infante et al. (2020), Luo et al. (2019) and Rodrigue (2020)
Travel Demand	Function of the travel demand on the route	Disaggregated	Link	Z. Xu et al. (2020)
Number of transfers	Function of the number of stops for each route	Disaggregated	Node	Allroggen et al. (2015) and Luo et al. (2019)
Number of direct connections (Degree)	Number of incoming and outgoing links	Disaggregated	Node	Paleari et al. (2010), Psaltoglou and Calle (2018), Soh et al. (2010) and J. Wang et al. (2011)
Service reliability	Function of the reliability or punctuality of the service	Disaggregated	Network	Handley et al. (2019) and Kaplan et al. (2014)
Generalised Travel Costs	Function of the generalised travel costs, generally including also travel time	Aggregated	Link	Rodrigue (2020) and Takebayashi (2015)
Travel Impedance	Function of different indicators such as travel time, frequency, distance, cost (similar to generalised travel costs)	Aggregated	Link	Luo et al. (2019)
Node Strength (Weighted Degree)	Function of weight (frequency, travel time) aggregated per node	Aggregated	Node	Calzada-Infante et al. (2020), Feng et al. (2021), Soh et al. (2010) and W. Wang et al. (2020)
Number of neighbours	Function of the number of direct and indirect connections ranking them based on their distance in steps	Aggregated	Node	Zhou et al. (2019)
Temporal coordination/ Transfer Time	Temporal coordination of schedules to assess the feasibility of transfers	Aggregated	Node	Burghouwt and de Wit (2005), Calzada-Infante et al. (2020), Danesi (2006), Paleari et al. (2010) and Williams and Musolesi (2016)
Routing Factor	Ratio between the actual travel distance and the theoretical minimum travel distance	Aggregated	Link	Burghouwt and de Wit (2005), Danesi (2006) and Paleari et al. (2010)
Directness	Function of travel time, detours and transfers	Aggregated	Link	Allroggen et al. (2015)
Connection quality	Function of transfer and detour times, Value of Time	Aggregated	Node	Allroggen et al. (2015), Zhu et al. (2018, 2019), Zhu, Zhang, Zhang et al. (2019)
Importance of destination (Eigenvector)	Function of demographic, economic or connectivity indicators	Aggregated	Node	Allroggen et al. (2015), W. Wang et al. (2020), W. Xu et al. (2018) and Z. Zhang et al. (2015)
Centrality (Closeness)	Function of average shortest path or average quickest path	Aggregated	Node	Luo et al. (2019), Malighetti et al. (2008) and W. Wang et al. (2020)
Hub potential (Betweenness)	Function of inbound and outbound frequencies and centrality indicators	Aggregated	Node	Dennis (1998), Derudder et al. (2010), Paleari et al. (2010) and Psaltoglou and Calle (2018)
Network Connectivity	Function of the number of nodes and links within a specific range from a node (beta index)	Aggregated	Node	Z. Xu et al. (2020)

consequently allowing to compute comparable values across them.

A first important dimension, often considered when assessing the connectivity in the air sector relates to the temporal coordination between sequential services and consequently to transfer times. The fact that an incoming service and an outgoing one connect in a certain hub does not necessarily imply that this connection is either viable or attractive to passengers. This measure is particularly important for air transport, due to the low service frequencies and the wider distribution in transferring times. In this regard, the literature review on air transport networks connectivity from Burghouwt and Redondi (2013) show that in studying air transport networks there are two main currents of thought in regards to the inclusion of temporal coordination within connectivity measures and connection models. Some employ min/max transfer time thresholds to consider the timetable coordination and assess the feasibility of the transfers (Calzada-Infante et al., 2020; Danesi, 2006). Others more simply deem a certain indirect connection viable when two legs transferring in a certain hub, one departing from the origin and the other arriving in the destination, are available. In particular, Danesi (2006), following the first current, employs timetable coordination to assess the viability and quality of indirect connections in the computation of the hub connectivity of some European airports. This measure could also be employed for international rail services, as the frequencies are also in this case quite limited, and transfer times represent an important variable in making international rail travel really attractive. In fact, Calzada-Infante et al. (2020), analysing European rail, consider indirect services only when the transfer is actually possible, depending on the synchronisation of the schedules. Thus, the availability of indirect services is conditional on waiting times in the stations between arrival of incoming services and departure of outgoing ones being lower than a hour. Also Williams and Musolesi (2016), following a similar approach, employ spatio-temporal networks to consider the temporal dimension and schedule coordination. Finally, Zhu, Zhang and Zhang (2019) take into account of the temporal dimension to assess the quality of transfers and inter-modal connections by considering the difference between the actual transfer time of a service and the maximum connection time.

Temporal coordination expressed as transfer times is often used as a measure of transfer and/or connection quality. However, measures considering the quality of connections do not take into account only this parameter. Allroggen et al. (2015), for instance, measure air connectivity taking into account the total quantity of all connections available and the relative quality of every connection in terms of the effort required for passengers to travel using that connection. In particular, link quality is defined as a measure of frequency and directness, where the latter measures travel time and a detour factor, and is paired with destination quality. In this regard, other authors, such as Z. Zhang et al. (2015), employ network science tools, such as eigenvector centrality, to take into account of the importance not only of connections but also of connected nodes. Simpler metrics focusing on connections and destinations include the number of neighbours, as employed by Zhou et al. (2019) and the number of direct connections as employed by Danesi (2006).

Furthermore, it appears clear that other than the importance of the destination, when considering indirect connections, also layover time and the overall quality of the stopover and of transfer hubs gain importance. Paleari et al. (2010) other than taking into account the number of direct connections, the OD distances, also focuses on the relative position of intermediate airports and on the waiting times spent there as a measure of transfer quality. Transfer quality also depends on less quantifiable elements like the comfort of transfer hubs, which are rarely included in connectivity metrics. What is often included, on the other hand is the relative importance of nodes. In particular, Soh et al. (2010) defines the node strength as “the sum of the weights on the edges incident upon it”, which considering network science definitions could be translated as the weighted degree of a node. Similarly, Feng et al. (2021) defines local transfer strength as “the convenience of arriving in and departing from one node”, measuring it as the connection frequency and the distance between nodes. Furthermore, in the field of network science, betweenness centrality is often employed to assess the centrality of a certain node in terms of the number of shortest/quickest paths transiting through it. On the other hand, Derudder et al. (2010) defines hub connectivity as absolute hub intensity, accounting for the total number of transferring passengers, and relative hub intensity, identified by the ratio between absolute hub intensity and the total number of passengers arriving or departing from the airport.

Another important dimension, relates to the attractiveness of possible routes and on the eventual limitation of less attractive ones. To do this, in the aviation sector, the so-called routing factor is often employed (Burghouwt & Redondi, 2013). Paleari et al. (2010), define the routing factor as “the ratio between in-flight distance and potential direct flight distance” , thus the ratio between the actual flight distance of a certain route and the minimum distance between origin and destination. Routing factors are often employed as thresholds to impose maximum acceptable detour values and discard less direct and attractive routes(Danesi, 2006; Paleari et al., 2010). However, these routing factors could also be included as components of connectivity metrics accounting for the directness of routes. The directness is another component often employed in connectivity measure, and is closely related to the routing factor. In particular, Allroggen et al. (2015) define the directness of a connecting flight as a function of detours and layovers, arguing that both are elements that cause disutility to the traveller. In their study, directness is used together with frequencies to assess the “connectivity value” of each connecting service at the “link-quality-level”.

Other ways to implicitly take into account detours relate to the shortest and quickest path methodologies, where the former consist of finding the shortest path between origin and destination on a certain transport service network whilst the latter add to it the travel time dimension (Burghouwt & Redondi, 2013). These measures are employed to identify the minimum number of transfers and the minimum travel time respectively and can be used either to select the most attractive(s) (shortest/quickest) path(s) between each OD-pair discarding all others or to rank all possible paths based on the shortest possible connection. The former approach, is employed for instance by Malighetti et al. (2008), who computes the connectivity of a certain airport as sum of the number of links in the shortest paths to reach every other node in the network divided by the number of nodes in the network minus one. This formulation closely resembles the closeness centrality as defined in network science.

2.2.5. Conclusions

This second part of the literature review shows the importance of connectivity in transport systems. Zhu et al. (2018), for instance, highlight the strong relationship between transport connectivity and regional economic development. However, connectivity is a versatile and multi-faceted term that does not have a clear, univocal, and truly comprehensive definition. This concept, in fact, acquires precise meanings depending on the specific context into which is employed. Many are the definitions and metrics associated to this term employed across the literature. The common thread that appears to connect every definition and metric is that end goal of connectivity is to measure the strength or degree wherewith two elements are connected. However, this degree of connectivity can be measured at different levels, especially when working with graphs and networks, which are constituted by many elements. Handley et al. (2019), for instance, make a distinction between stop/node connectivity and network connectivity, where the former represents the specific connectivity of the local node and the latter represents the global connectivity of the entire network. Furthermore, Z. Xu et al. (2020) make a distinction between edge connectivity and node connectivity, where the former represents the size of the minimum edge cut set, whilst the latter represent the size of the minimum node cut set required to disconnect a specific pair of nodes in a network. Similarly, Zhu et al. (2018), distinguishes the connectivity between two urban areas, which measures the convenience of transportation on the specific corridor, from the connectivity between an urban area and all other urban areas in the network, which measures the importance of the city within the network. The former definition could be related to edge connectivity, whereas the latter could be equated to node connectivity. Finally, Zhou et al. (2019), using percolation theory, defines global connectivity as the degree to which the entire network is connected and local connectivity as the extent to which each node is connected to its neighbours. Thus, throughout this thesis the following definitions of connectivity are used as aggregation levels:

- Link Connectivity: connectivity of a specific corridor (link)
- Node Connectivity: connectivity of a specific urban area (node)
- Network Connectivity: connectivity of the entire air/rail transport system (network)

And the following definitions are used to distinguish between the two levels of analysis:

- Local Connectivity: connectivity of direct connections
- Global Connectivity: connectivity of all possible connections including transfers

Furthermore, from this brief but comprehensive review it can be also concluded that a wide range of connectivity measures is available in the literature, from very disaggregated levels to more aggregated ones. Zhu, Zhang and Zhang (2019), in particular, point out that a wide range of connectivity metrics captures elements related to four main components: travellers, transport system, land use and temporal. Given the scope of this study, the first two dimensions represent the core focus of this research, whereas the latter two are only marginally going to be addressed. Zhu, Zhang and Zhang (2019) also describe the importance of carefully choosing or designing metrics that can be applied to more than one transport mode, in order to compare them. More details about the metrics chosen and the level of analysis are provided in Section 3.3. Finally, comparing the basic disaggregated metrics, provided in Table 2.2, with the factors influencing the modal competitiveness between rail and air, as illustrated in Section 2.2.4, an important overlap between these two is noticeable. Thus, it can be concluded that assessing the connectivity of links, nodes and overall networks provides insight on the modal competitiveness between the two modes, which consequently allow to answer the main research question, as defined in Section 1.5.

2.3. Network Science applications in Transport

Network science is a research field rooted in graph theory, which aims to study the relationship and interconnection between the elements of complex systems (Newman, 2010). Despite being a rather novel field of study, Network science has polarised the attention of scholars in the last two decades (Barabási & Pósfai, 2016). Lin and Ban (2013) highlight that this trend has been reflected in the transport sector, with an increasing number of studies investigating the topological properties of transport networks, especially in the fields of Public Transport and Aviation (Derrible & Kennedy, 2010a; Kotegawa et al., 2014). In fact, the recent developments in Network Science have enabled a more systematic quantification of specific network topological properties, providing researchers with powerful tools to analyse transport networks and the collective dynamics arising from the interactions between different elements of the system (Cats, 2017). Network Science has been widely employed to broadly study the specific characteristics, features and structure of networks (Derrible & Kennedy, 2010a; P. Sen et al., 2003; Sienkiewicz & Hołyst, 2005; von Ferber et al., 2007, 2009; J. Wang et al., 2011), to assess their impact on pricing (Silva et al., 2022) and on performance efficiency (Kotegawa et al., 2014) and to compare these values across different modes (Feng et al., 2021; Gattuso & Miriello, 2005; Majima et al., 2007). Furthermore, scientists have also employed Network Science to identify critical components within the network (Feng et al., 2021; Psaltoglou & Calle, 2018), and to study the evolution (Bombelli, 2020; Calzada-Infante et al., 2020; Cats, 2017; C. Chen et al., 2018), reliability (Z. Zhang et al., 2015), robustness (Derrible & Kennedy, 2010b), accessibility (Luo et al., 2019; W. Xu et al., 2018) and connectivity (Bombelli, 2020; Paleari et al., 2010; W. Xu et al., 2018) of Transport Networks. A first overview of some meaningful papers related to the topic is provided in Table 2.3,

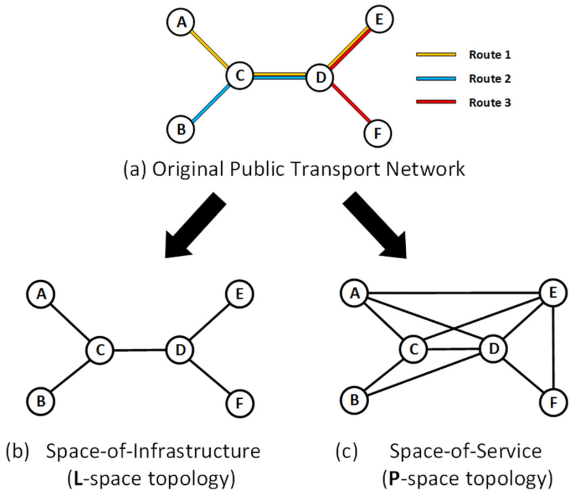
Table 2.3: General Overview of Relevant Papers Related to the Scope of this Research

Study	**Transport Mode**	**Aggregation Level**	**Network Representation**	**Research Scope**
Bombelli (2020)	Air	Global Level	N/A	Analysis of the global networks of integrators (air cargo service providers)
Cats (2017)	Metro	Urban Level	L-Space	Analyse the topological evolution of a metropolitan rail transport network
C. Chen et al. (2018)	Rail	National Level	L-Space	Analyse the development of the Chinese high-speed rail network over time
Derrible and Kennedy (2009)	Metro	Urban Level	L-Space	Investigate the role of network design on ridership
Derrible and Kennedy (2010a)	Metro	Urban Level	L-Space	Characterise the network feature of metro systems
Derrible and Kennedy (2010b)	Metro	Urban Level	L-Space	Analyse the complexity and robustness of metro networks
Gattuso and Miriello (2005)	Metro	Urban Level	L-Space	Analyse comparatively metro networks
Guo et al. (2022)	Air	Regional Level	N/A	Detect delay propagation in regional air transport systems
Kotegawa et al. (2014)	Air	N/A	P-Space	Examine impact of service network topology on performance efficiency metrics in air transportation systems
W. Li and Cai (2007)	Rail	National Level	P-Space	Analyse empirically the China Railway Network (CRN)
Majima et al. (2007)	Rail, Metro & Waterbus	Urban Level	L- and P-space	Investigate five transport networks in Japan
P. Sen et al. (2003)	Rail	National Level	P-Space	Analyse the small-world properties of the Indian Railway network
Sienkiewicz and Hołyst (2005)	Public Transport	Urban level	L- and P-Space	Analyse statistically PT networks in Poland
Silva et al. (2022)	Air	National level	N/A	Investigate how the air transport network structure affects ticket prices
Soh et al. (2010)	Public Transport	Urban level	P-space	Analyse the dynamical properties of the PT system in Singapore
von Ferber et al. (2007)	Public Transport	Urban Level	L- and P-space	Analyse statistically PT networks
von Ferber et al. (2009)	Public Transport	Urban Level	L-, P-, B- and C-Space	Analyse empirically and model PT networks
Z. Xu et al. (2020)	Rail	National Level	L-Space	Analyse the Chinese railway system as a complex network

2.3.1. Network Topology

Rodrigue (2020) defines network topology as the specific arrangement of nodes and links, in terms of their location and the nature of their connection. In other words, representing the physical and logical relationships intercurring between the nodes of a network, this term describes the configuration of the links and nodes of a certain network. Consequently, network topology does not depend only on the type of network (e.g. social network, transport network) that is to be represented but also on the translation of the specific system characteristics into a network structure. In particular, given the considerable impact that network structure has on the properties of the system, Barabási and Pósfai (2016) argue that the choices made in representing systems as networks determine the ability to use network science successfully. In fact, depending on the selection of the network specifications very distinct representations of reality can be developed allowing to answer diametrically different questions. In transport networks nodes can represent a wide set of different functional entities, ranging from urban areas, to airports, ports, train stations or PT stops. At the same time, also the links of a network may represent different elements such as transport services, infrastructural corridors or transfer stations. Furthermore, links can be either directed or undirected and weighted or unweighted, influencing the topology of the graphs and consequently its characteristics.

Luo et al. (2019) argues that in order to unravel the topological characteristics of a transport network it is fundamental to consider the key features of that specific mode. In relation to this, an important concept is the one of graph planarity. In graph theory, a planar graph is a graph that can be embedded in the Euclidean plane, so that its edges do not intersect in any place other than the vertices to which both are incidental (Wilson, 1972). Lin and Ban (2013) highlight how, generally, aviation and maritime networks present more non-planar characteristics compared to other ground transport networks, such as PT or rail. This relates to the specific characteristics of the two group of modes. In fact, despite both share a service dimension, in the case of the former that is independent from a physical infrastructure, whilst for the latter the service dimension is superimposed on the infrastructure dimension. Thus, whilst air networks are consistently represented with airports as nodes and with the direct routes among them as links, representations of PT or rail networks can follow a wider set of paradigms. In this regard, von Ferber et al. (2007), identifies four main topological representation for PT networks, namely L-space, P-space, C-space and B-space. The latter two focus on transfers, placing an important focus on the line dimension. In the case of rail this appears to be less influential, and for this reason they are not going to be further considered nor described. The L-space and P-space topological representations of Public Transport networks are visually rendered in Figure 2.2.



(a) Original Public Transport Network

(b) Space-of-Infrastructure (**L**-space topology)

(c) Space-of-Service (**P**-space topology)

Figure 2.2: Topological Representations of Public Transport Networks (Source: Luo et al., 2019)

L-space, or infrastructure space, is a straightforward representation of ground networks that aims to capture the infrastructure dimension of the system maintaining the actual configuration of the transport network (von Ferber et al., 2009). Public transport stops are modelled as nodes, whilst links represent the physical adjacency of the stops in terms of infrastructure segments. L-space representation can also be employed to represent rail networks with a few adaptations. Z. Xu et al. (2020), models railway system as a L-space network, representing unique stations as nodes and the existence of physical connections (i.e. the existence of rail infrastructure) between two nodes as links. Due to its resemblance to physical networks L-space is widely used by researchers to model PT and rail networks (Cats, 2017; C. Chen et al., 2018; Derrible & Kennedy, 2010a; Gattuso & Miriello, 2005; Majima et al., 2007; Psaltoglou & Calle, 2018; Sienkiewicz & Hołyst, 2005; Z. Xu et al., 2020; Z. Zhang et al., 2015).

P-space, or service space, is a representation introduced by P. Sen et al. (2003) solely focused on the service dimension, as it does not capture any information related to the infrastructure layer. In the case of PT nodes represent stops whereas links designate the availability of a direct service connection. As said for L-space also this network representation can be adapted to rail, by substituting stations to stops (W. Wang et al., 2020). It is important to note that P-space does not contain any information about infrastructure, and consequently follows the same representations paradigms that are widely used in aviation. This consequently allows to create comparable air and rail network representations, which can be analysed comparatively. Across the literature on PT, P-Space is generally employed to assess transfer possibilities, as it enables to assess the number of transfers and the total transfer times (Sienkiewicz & Hołyst, 2005; von Ferber et al., 2007, 2009). Furthermore, P-space is often employed to evaluate the topological characteristics of both PT and rail networks (P. Sen et al., 2003; Soh et al., 2010).

2.3.2. Complex Network Analysis of Transport Systems

Transport networks characteristics, including connectivity as defined in Section 2.2, can be assessed and compared by representing transport systems as networks and by employing the tools provided by network science. This approach is supported by the fact that transport systems, featuring a compound set of dependencies, relationships and interactions between elements, can be regarded as complex systems (Zanin & Lillo, 2013). To highlight and study the aforementioned features and the underlying properties, complex systems can be further represented as complex networks. Complex networks, relating to the structure of complex systems, are characterised by non-trivial topological features, including patterns and relationships between nodes that do not generally occur in random graphs (Barabási, 2009). Complex networks, thus, do not share the characteristics and properties of neither purely regular or purely random networks but feature specific properties that can provide interesting insights on the structure of the system. In particular, two important network structure properties relate to the concepts of small-world and scale-free networks. A small-world network, as defined by Watts and Strogatz (1998), is a quasi-regular network characterised by a small average path length (similarly to a random graph) and a large clustering coefficient (larger than a random graph). Small-world networks are decentralised, highly connected and very efficient, allowing to easily connect every node to the rest of the network. On the other hand, Barabási and Albert (1999) proposed the scale-free property to describe the World Wide Web and all the networks whose node degree distribution conform to a power-law. This property captures the tendency of the nodes to connect preferentially to already well-linked nodes, which represents a shared characteristic among some types of complex networks. As a consequence,these network are characterised by few hubs with many connections and by many nodes with a few connections only. With growth these differences in node degree generally tend to be further exacerbated.

Many scholars have focused on analysing the topological characteristics of transport networks by means of global and local indicators and by identifying specific network structure properties (Derrible & Kennedy, 2010a; Gattuso & Miriello, 2005; W. Li & Cai, 2007; P. Sen et al., 2003; Sienkiewicz & Hołyst, 2005; von Ferber et al., 2007; Z. Xu et al., 2020). In particular, W. Wang et al. (2020) highlight the consistent number of studies which applied the two aforementioned concepts to analyse air transport networks. Guimerà et al. (2005) first observed that the worldwide airport network is a scale-free and small-world network. In particular, comparing the node betweenness and the node degree Guimerà et al. (2005) discover that the most-connected nodes (i.e. high degree) are generally not the most-central (i.e. high betweenness centrality). The study further suggests that the anomaly of centrality values is related to the existence of communities in the network. The topological analysis from Paleari et al. (2010) further confirms that also at a continental level in Europe, the US and China air systems follow the behaviour of small-world networks. In particular, the authors find that despite the degree distributions are comparable to the scale-free power law, they are better fit by a double Pareto law. This is further proven by J. Wang et al. (2011), who, analysing the topological characteristics of the air transport network of China, find that it is a small-world network at the same time not featuring scale-free characteristics. In fact, the degree distribution appears to be best captured by an exponential function, indicating a more considerable dominance of large hubs. Finally, Luo et al. (2019) highlight that also in the field of PT many studies assessed the topological characteristics of the networks, especially focusing on scale-free and small-world properties. Given that Public Transport networks are not part of the scope of these research these are not further described within this review. However, a notable example is the study from Soh et al. (2010) and a critical review on the topic is provided by the paper from Zanin et al. (2018).

As shown in the previous paragraph, the topological analysis of transportation networks is a widespread methodology employed to study PT and air transport networks. However, it has been rarely used to assess rail systems. In particular, P. Sen et al. (2003) first analysed the Indian railway network, discovering that the system displays small-world properties. Calzada-Infante et al. (2020) comes to similar conclusions analysing the European International Railway Network (EIRN) by means of Complex Networks Analysis. In particular, the study concludes that, despite its average path length is higher than the one of a random graph, the EIRN behaves like a small-world network. The authors, however, argue that the higher average path length was expected as the international rail routes are geographically constrained. Furthermore, the methodology developed by Calzada-Infante et al. (2020), describing the network properties of international rail services, enables to compare the European case study with the Chinese one, observing their similarities and differences in terms of network structure. The results show that the clustering coefficient of both networks is larger than random networks, even though in the Chinese case node strength (measured in terms of frequencies) grows faster and the clustering coefficient decreases faster than in the European case. Finally, Calzada-Infante et al. (2020) compares the case of direct services only to a network with transfers discovering that in the latter case the number of potential destinations reachable from well connected origins quite literally “explodes” due to the higher assortativity.

Assortativity represents the preferential attachment tendency of nodes to connect to similar nodes. Thus, a network is said to be assortative when nodes are connected mostly to other nodes with similar node degree. This indicator is also used by W. Wang et al. (2020) to analyse the Chinese Railway Network, which appears to be disassortative, meaning that tend to show hub-and-spoke tendencies with low degree nodes being often connected with high degree ones. This is further proved by the negative degree-degree correlation. In this instance it is interesting to notice that also the Chinese Air Network appears to be largely disassortative (J. Wang et al., 2011). These results are obtained through a weighted k-core decomposition, which partition of the nodes based on their relative importance on a weighted (i.e. number of runs/frequency) network. In particular, the Chinese Rail Network, comparably to the case of the Chinese Air Network described by J. Wang et al. (2011), exhibits a small-world behaviour with a two-regime power-law degree distribution with different exponents for low and high degree nodes. Furthermore, analysing the degree-clustering correlation W. Wang et al. (2020) find that low-degree cities, due to the lower connectivity in the network, tend to feature higher clustering coefficients, whilst high-degree cities show the converse tendency. Finally, other important properties of the network are derived from the node degree-strength positive correlation, which shows that nodes with a higher connectivity (i.e. node degree) also feature higher traffic (i.e. node strength), and from the cumulative distribution of edge weight, which following a power law, implies that most trips are generated by a limited number of major corridors with most corridors having low frequencies.

Despite the great insights that both the aforementioned papers provide, neither of them compares the characteristics and performances of the rail network to the ones of the comparable air network. In this regard, Feng et al. (2021) represents an exception modelling both modes and employing network science to study the inter-modal high-speed rail and aviation transfer network in China using a P-space representation. An interesting finding relates to the fact that such network exhibits small-world properties, which the authors argue is counterintuitive. The paper, however, focuses mostly on the assessment of vulnerable elements rather than assessing or compare the connectivity of the two modes.

2.3.3. Network Science and Connectivity

Over the last years network science has been increasingly employed to analyse the connectivity of transport networks. However, researchers do not share a universally agreed methodology to assess connectivity, as many different indicators are employed across the literature. As previously highlighted in Section 2.2.2, physical connectivity metrics as defined in graph theory are not able to correctly represent the connectivity of transport systems, such as High-speed Rail Networks (Z. Xu et al., 2020). Furthermore, the connectivity measure generally used in Network Science (gamma index) can only be employed for planar graphs, thus excluding P-Space. In fact, the gamma index measures the relationship between the number of links in a network and the number of possible links. Moving towards transport geography Rodrigue (2020) proposes the so-called beta index, which measures the relationship between the number of links and the number of nodes to compute the level of connectivity in a network. However, this measure is bounded to the network level and does not allow to compute the connectivity at the node and link levels. Thus, W. Xu et al. (2018) adapts the aforementioned beta index to assess the implications of high-speed rail for Chinese cities. In particular, the connectivity of cities is computed as a ratio of the HSR lines passing through a certain city (links) to the number of HSR stations within that city (nodes). The extension of the beta index proposed by W. Xu et al. (2018) allows to capture connectivity at node level, but is still limited to that level, failing to address and determine link connectivity. Furthermore, it takes into account only network dimensions, completely disregarding most of the important connectivity factors identified in Section 2.2.4.

To deal with the former limitation, researchers usually assess connectivity employing the concept of centrality, as centrality indicators can be computed at node, link, and network level. For instance, Calzada-Infante et al. (2020) employs centrality indicators to evaluate the degree of connectivity between European cities, both in terms of direct and transfer connections across the European international rail network. Furthermore, C. Chen et al. (2018) analyse the evolution of the Chinese high-speed rail (HSR) network examining the variations and developments in terms of network accessibility measured by node degree, strength, closeness, and betweenness. However, these studies are still limited to the analysis of the network characteristics, failing to take into account other important factors, such as the ones determining rail-air modal competitiveness, identified in Section 2.1.3. In the Public Transport domain, some researchers have successfully managed to include also characteristics of the system that are not directly related to the network. Luo et al. (2019), for instance, applies weights on the links to include relevant metrics to the network. In particular, the connectivity of PT stops is measured by computing the closeness centrality of a weighted network, where link weights represent the travel impedance between nodes measured in terms of generalised travel cost. Similarly, but through a different approach, Psaltoglou and Calle (2018), in order to assess the connectivity urban public transportation networks, considers network metrics (i.e. node degree and betweenness centrality) together with the operational features of the transport system, including frequency, velocity, vehicle capacity and distance. The study proposes a methodology to identify the critical nodes within a PT network combining network science principles, public transport features and urban planning indicators into a connectivity index.

Also Z. Zhang et al. (2015) uses centrality indicators to explore the relative importance of the elements of the Chinese railway network and consequently to determine the level of priority of specific nodes and links. In particular, the local and global effects on the network are assessed using degree centrality and betweenness centrality respectively. Furthermore, the relative importance of the nodes is assessed through eigenvector centrality, and the likelihood of experiencing more directly the effect of failures is investigated through closeness centrality. In relation to the identification of critical elements within networks another interesting study is the one from Feng et al. (2021), who evaluate the performance of the transportation system in terms of the global transfer connection strength to assess the importance of the transport system components.

2.3.4. Conclusions

This third part of the review has shed light on the many applications of network science within the transport system domain, focusing in particular on the studies related to connectivity, air and rail network characteristics and the identification of critical components. A complete overview of the reviewed papers, summarising their contents and characteristics, is provided in Figure 2.4. In this regard, Derrible and Kennedy (2011) argue that studies employing network science to analyse transport systems have often been led by networks scientists in search of a case study rather than by transport engineers and planners looking to answer actual problems. Consequently, Luo et al. (2019) highlight that much of the literature only aimed at capturing topological characteristics and properties of the networks failing to embed some important characteristics of the transport system, such as the ones identified in the connectivity metrics review in Section 2.2.4. This has drawn the critiques of some researchers, such as Dupuy (2013) and Zanin et al. (2018), who point out some crucial limitations of network science in providing recommendations to policy-makers and transport planners. However, this review has shown that Network science can be successfully employed to analyse transport system. Many studies, in fact, assessing the characteristics of transport networks through network science, managed to provide in-

Table 2.4: Overview of the Papers Reviewed

Study	Transport Mode	Aggregation Level	Network Representation	Network Type	Link Type	Link Label	Research Scope
Calzada-Infante et al. (2020)	Rail	Continental Level	P-Space	Directed	Weighted	Frequency	Analyse topologically the European international railway network
Feng et al. (2021)	Air & Rail	National Level	L- and P-Space	Directed	Unweighted	N/A	Identify structure and critical components of high-speed rail and aviation networks in China
Luo et al. (2019)	Tram	Urban Level	L- and P-Space	Undirected	Weighted	Travel impedance	Assess comparatively PT accessibility
Paleari et al. (2010)	Air	Continental Level	N/A	Undirected	Weighted	Annual offered seats per OD airport	Investigate and compare the connectivity of the airport networks in China, Europe and US
Psaltoglou and Calle (2018)	Bus	Urban Level	L-Space	Directed	Weighted	Number of routes	Identify critical points in urban PT networks
J. Wang et al. (2011)	Air	National Level	N/A	Undirected	Unweighted	N/A	Examine air transport network structure and nodal centrality in China
W. Wang et al. (2020)	Rail	National Level	P-Space	Directed & Undirected	Weighted	Counts of trips	Analyse the connectivity of China's HSR
W. Xu et al. (2018)	Rail	National Level	N/A	Undirected	N/A	N/A	Assess the connectivity and accessibility of cities on the Chinese HSR network
Z. Zhang et al. (2015)	Rail	National Level	L-Space	Undirected	Weighted	Estimated flows of passengers	Assess the critical nodal and linear elements of railway infrastructure in China
This Study	Rail & Air	Continental Level	P-Space	Directed & Undirected	Weighted	Travel times, frequency	Assess comparatively the connectivity of long-distance transport services in Europe

teresting and useful recommendations, from both a theoretical and practical perspective, in relation to network properties (Calzada-Infante et al., 2020; Paleari et al., 2010; J. Wang et al., 2011; W. Wang et al., 2020), connectivity (Luo et al., 2019; W. Xu et al., 2018) and to the identification of critical components of the network (Feng et al., 2021; Psaltoglou & Calle, 2018; Z. Zhang et al., 2015). This is largely due to the inclusion of the aforementioned transport-related factors, whether through the use of weighted networks (Calzada-Infante et al., 2020; Luo et al., 2019; Paleari et al., 2010; Z. Zhang et al., 2015), by combining network-related and transport-related indicators (Psaltoglou & Calle, 2018; W. Xu et al., 2018) or by employing gravity-based indicators for nodal strength (Feng et al., 2021). The crucial importance and added value of using weights to analyse complex networks was already proven by Barrat et al. (2004), who highlight the increased significance of the findings compared to the use of unweighted networks. Thus, as already argued by (Cats, 2017), analysing networks properly weighted with meaningful factors allows to remove a consistent part of the limitations of this methodology and to provide more insightful conclusions and recommendations.

This section has also demonstrated that a P-space topological network representation allows to compare the supply of direct services across rail and air transport systems. In particular, throughout the literature Network Science applications have mostly been employed for the analysis of Public Transport and Air Systems, whilst the rail sector appears to be lagging behind (Calzada-Infante et al., 2020). Furthermore, it is clearly noticeable that China represents the geographical scope of most studies on the matter, especially in relation to High-speed rail and Air connectivity. The focus on the Chinese case clearly follows the relentless rise of the Chinese High-speed network, which over the last decades has increasingly established itself as a clear competitor for air on most national routes. However, whilst a similar process has arguably also happened in Europe, this has not translated into an equivalent amount of research on the network connectivity dimension of the problem. A probable reason for this disparity is the fragmentary nature of the European rail system, both in terms of service supply and of data availability, compared to the centralised Chinese system.

3

Methodology

Within this chapter the model characteristics and specifications required to answer the research questions, as defined in Section 1.5 are identified and discussed. In particular, as highlighted in Section 1.7, the supply of air and rail services is modelled in this study using complex networks. In order to construct realistic networks and consequently to ensure the validity of the results it is paramount to accurately translate the characteristics of the rail and air networks to the model. Complex networks represent, in fact, a simplification of reality used to translate the characteristics of a complex system into an abstract structure called topology (Newman, 2010). In this instance, the literature review provided useful insights into the approaches employed by researchers to tackle comparable problems, and in particular on the characteristics of the different types of network topology. It seems important to highlight that the networks analysed in this research still represent merely a model of reality, being able to only partially capture their characteristics and properties. Every model aims to approach reality in the most accurate way, however it is required to restrict the level of detail in order to limit the cumbersomeness and to guarantee the technical feasibility of the analysis, ensuring that the validity of the model is preserved. Thus, the design assumptions and constraints employed to define the aggregation level that best allows to clearly and thoroughly tackle the problem defined in Section 1.2, without being overly dispersive or excessively cumbersome, are highlighted throughout this chapter.

An overview of the methodology employed for the development of the model is schematised and presented in Figure 3.1. The geographical scope, defined in Section 4.1 is employed as an input together with the criteria highlighted in Section 3.1.3 to select the urban areas to include in the model. The selection of cities is consequently employed to select the airports and train stations to include in the model using the methodology described in Sections 3.1.4, 4.6 and 4.7. During this first phase of the methodology, highlighted in blue in Figure 3.1, all the characteristics of the model are specified and discussed. In particular, the network specifications and all the assumptions made during the model definition process are illustrated in Section 3.1. The output of this first step serves as input for the collection of the data required for the network analysis. During this second phase the data is collected at the terminal level and cleaned following separate processes for rail and air as described in Section 3.2.2 and 3.2.1. Furthermore, to allow the comparability of the results across the two modes, the data is re-aggregated at the urban area level. The outputs from the re-aggregation process are then employed to construct

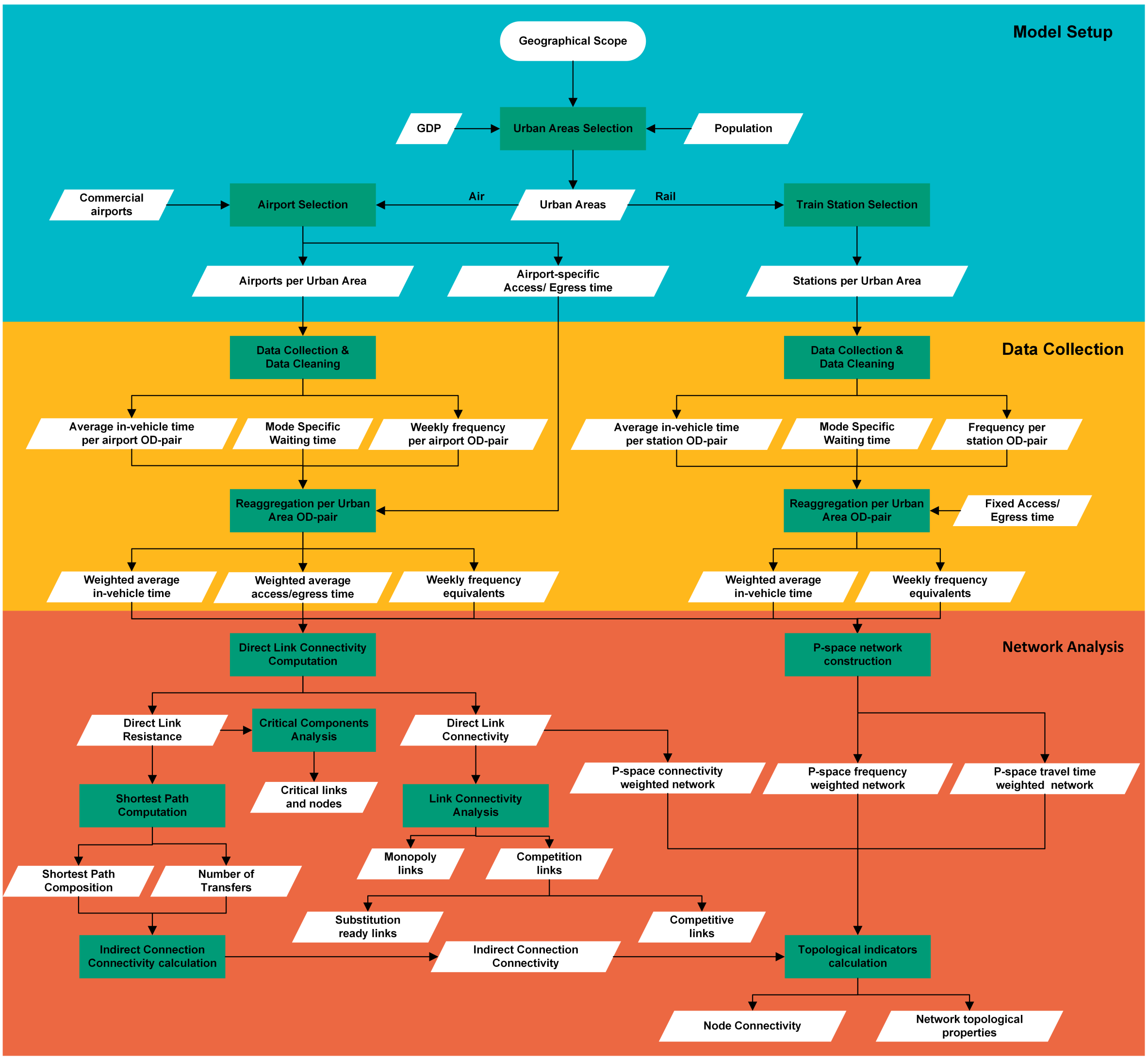


Figure 3.1: Methodology Overview

the P-space networks and to compute the degree of connectivity of all the available direct connections. The specific processes employed during this third phase and used to analyse the network characteristics and performances in terms of centrality, connectivity and modal competitiveness are further described in Section 3.3.

3.1. Model Setup

As mentioned in Section 2.3 some important representation decisions are required in order to translate reality into complex networks and to be able to use Network Science to derive the expected results. Thus, an in depth explanation of the network representation choices and on the network specifications selected for this research based on the research questions defined in Section 1.5 is provided in this section. Furthermore, more details about the specific case study and the Urban Area Selection processes can be found in Chapter 4.

3.1.1. Network Specification

Following the definition of graph provided in Section 2.2.2, all the networks analysed in this study are represented as undirected weighted graphs (network) $G = (N, L)$, composed of a set of vertices (nodes) $N = \{n_1, n_2, ..., n_{|N|}\}$ and of a set of edges (links) $L = \{l_1, l_2, ..., l_{|L|}\}$ using a space-of-service or P-Space topology as described in Section 2.3. Thus, given the scope of the research, nodes N represent a set of European urban areas connected by international train services and with specific characteristics, as further specified in Section 4.4, whilst links L represent the eventual existence of at least a direct connection between each pair of urban areas. Following what discussed in Section 2.3, the use of P-Space was preferred to other topologies, such as L-Space, to ensure the comparability between modes in terms of network representation. In fact, air networks, missing an infrastructure layer, cannot be represented in L-Space. Furthermore, weighted networks are employed to include specific information on the quality of the service supply other that the mere presence or absence of direct services (unweighted network). In particular, the weights employed on the links consists either of the factors influencing modal competitiveness between rail and air, selected in Section 3.1.2, or of derived metrics (e.g.connectivity). Thus, for each link weight a squared adjacency matrix, whose size is the number of urban areas, is created and used as input to generate the networks. Finally, undirected links are used on all networks to simplify the model and make the results clearer and easier to read. This entails assuming symmetrical conditions between all OD pairs, which appears to be a simplification of reality as on long-distance travel this is not always true. The number of services (frequencies) traversing a link and the amount of time (travel time) required to connect two points are not always the same in each direction. For instance, air services do not only fly back and forth between a specific OD-pair following a tour-based planning but might undertake more complicated triangular routes in the case of hub-and-spoke airlines or fly schedule-optimised routes connecting different airports within the same daily route in the case of point-to-point airlines. Using the available data, an examination of the correlation of the in- and out- node degree on the frequency-weighted networks is undergone, leading to Pearson's coefficients greater than 0.99 for both rail and air. The almost perfect linear relationship between in- and out-degree of the nodes allows to conclude that frequencies can be assumed to be symmetrical avoiding important losses of detail which would prevent an accurate depiction of reality. In the case of travel time, the differences can also be considered negligible, as they are generally limited to a few minutes

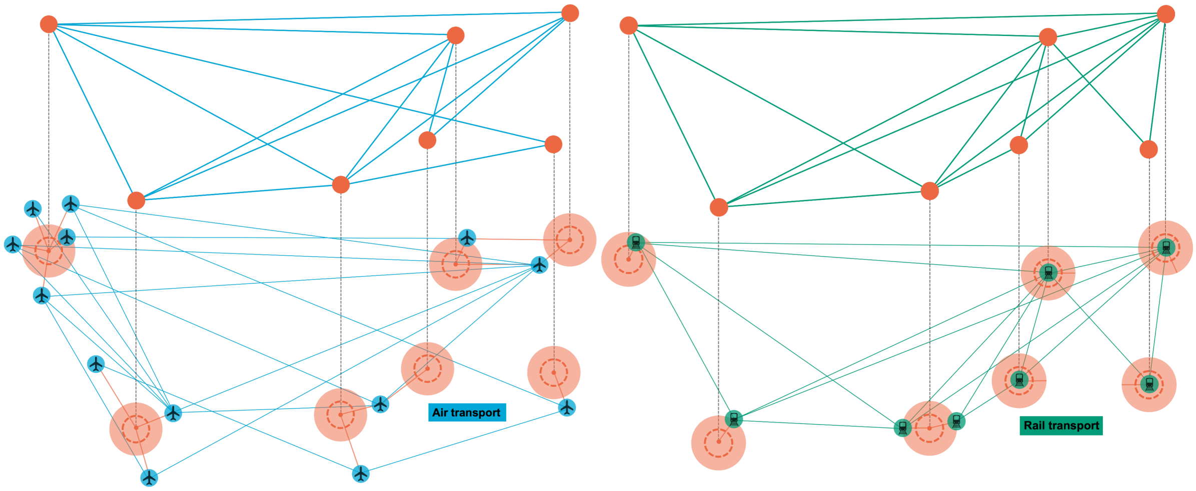


Figure 3.2: Basic representation of the modelisation of reality (lower layer) into networks (upper layer)

out of travel time magnitudes in the hundreds. As for the case of frequency this is confirmed by the Pearson's coefficients obtained from the available data.

In modelling the networks, particular attention is placed into assuring the comparability between the rail and air networks. In particular, figure 3.2, represents how reality has been modelled (lower layer) and translated into networks (upper layer). Each urban area, represented with wide light orange circles, is associated to a specific set of terminals, including rail stations, represented as green dots, and airports, represented as blue dots, and is connected with them by an orange line representing the access/egress leg. The connections between all the terminals, representing the in-vehicle leg, are shown as thin lines, blue for air and green for rail. Urban areas are thus translated at the graph level as nodes, shown as dark orange circles. Furthermore, all the connections between the respective terminals are aggregated into unique lines in the network, shown as thick lines, blue for air and green for rail. The two networks are represented separately and the data is retrieved at the level of terminals connections, as further explained in Section 3.2. The urban areas are chosen, following the criteria described in Section 3.1.3, and consequently the sets of associated airports are defined using the methodologies described in Section 3.1.4. Finally, the travel time composition based on this representation is described in Section 3.1.5, and some important considerations about travel time sensitivity per mode are provided in Section 3.1.6. In relation to what is discussed in this section the following assumptions are made:

- It is assumed that air and rail networks are completely distinct and disconnected. Thus, air-rail integrated routes and inter-modal connections are not taken into account. Given that the aim of this thesis is to compare the performances of the rail and air networks, it was deemed necessary to keep the two networks separated.
- For sake of simplicity the temporal element is not incorporated. Schedules, arrival/departure times and transfer time are consequently not considered. Thus, the connection quality and feasibility are not checked and it is assumed that it is possible to transfer from one flight to the next or from one train to the next if the services are available.

3.1.2. Connectivity and Factor Selection

Following the conclusions of the first part of the literature review, provided in Section 2.1.3, the determining factors influencing the modal competitiveness between rail and air are identified. These are further specified and described in Table 3.1.

Table 3.1: Overview of the Factors directly influencing Rail-Air modal competitiveness

Name	**Description**	**Influence on Modal Competitiveness**	**Included/Excluded**
Travel Time	Total duration of the door-to-door trip, including access/egress, waiting, in-vehicle and transfer time	Negative	Included
Travel Cost	Total economic value of the trip, including travel fares and access/egress cost	Negative	Excluded
Travel Demand	Total demand between each urban area OD pair	Positive	Excluded
Frequency	Number of services available on the main leg (terminal-to-terminal)	Positive	Included
Travel Comfort	Comfort level of the door-to-door trip (including transfers)	Positive	Some Components Included

The travel time element in this study is going to be taken into account as door-to-door travel time, where the specific components are further specified in Section 3.1.5. It was decided to exclude travel costs due to their inherent volatility and dynamic nature. In fact, prices for long-distance travel tend to fluctuate considerably based on many different factors, such as advance purchase and travelling period (i.e. seasonality). Furthermore, given that this study approaches the problem from a service supply perspective, travel demand is also not directly modelled. However, travel demand is arguably indirectly modelled throughout this study in the urban area selection process and in the identification process of connections with substitution potential. In the case of the former process, the decision to select only major urban areas follows, in fact, the assumption that, in order to be practically sustainable, international long-distance services require a considerable demand, which can only be assured under precise economic and demographic conditions. In particular, in order to ensure the long term sustainability of a supplied service the economic profitability of the specific route, and consequently the presence of enough travel demand, is fundamental. In fact, the supply of international long-distance services, especially in the aviation and high-speed markets, is generally liberalised and not subject to Public Service Obligations (PSOs) with a few exceptions only (mostly services used to serve remote and isolated areas, such as islands). In the latter process demand is indirectly modelled in terms of air frequencies, as they are employed as a minimum threshold in the urban areas selection process. This follows the assumption that the supply of flights in the free market implies the presence of enough demand in the market to guarantee the economic feasibility of such connections.

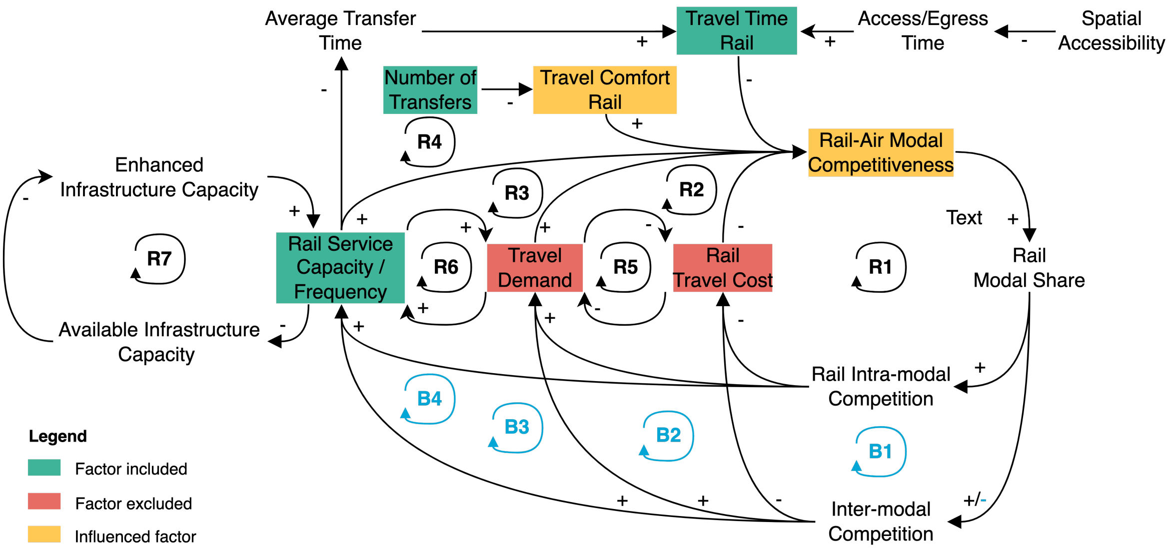


Figure 3.3: Factors Influencing Rail Modal Competitiveness

The frequency component in this study is included as weekly frequencies. The decision to have a weekly time-frame rather than a daily one refers to the inherent characteristics of long-distance transport, where service scheduling and planning is generally developed on a weekly rather than daily basis as the frequencies are rather low, in some extreme cases sinking to one or two weekly connections. On the other hand, travel comfort is not included as a variable as assessing the quality of transfers, due to the extent of the scope, was not deemed practically feasible. Although it can be argued that also this component is, at least partially, taken into account as the number of transfers/transfer time, which directly influences travel comfort, is included in the study. Furthermore, the fact that travel time is weighted

based on passengers' perception, as per explanation in Section 3.1.6, further adds to the consideration of the travel comfort component. In this study frequency is treated as a measure of service capacity. However, it seems important to point out that service capacity can widely vary across different modes. In fact, narrow-body aircraft (i.e. Airbus A320, Boeing 737-800), usually employed on European international routes, can hardly offer more than 200 seats per flight, whilst French TGV trains offer 430 seats per flight on average and the average train capacity of German ICE is 380 (Janic, 2003). Thus, the amount of offered seats at the same frequency levels can vary depending on the specific aircraft or train model and configuration employed by airlines and railway undertakings. However, due to the extreme complexity of the rolling stock fleets currently in use and to the possibility of coupling and uncoupling them, it was decided to use frequency as a measure of service capacity.
Finally, the included and excluded factors are contextualised within the system dynamics of rail-air modal competitiveness, provided in Section 2.1.3, as illustrated in Figure 3.3. Related to what discussed in this section the following assumptions are made:

- It is assumed that air and rail networks are completely distinct and disconnected. Thus, air-rail integrated routes and inter-modal connections are not taken into account. Given that the aim of this thesis is to compare the performances of the rail and air networks, it was deemed necessary to keep the two networks separated.
- The model represents a static condition of the system and does not take into account any dynamic element. Thus, it was excluded to include travel costs as a graph label, despite this was considered an important determining factor of connectivity. On a similar fashion traffic and airport congestion are excluded and not taken into account.

3.1.3. Urban Area Selection Criteria

The decision to select to use specific nodes rather than entire regions or grid-based systems follows the assumption that, in order to be practically sustainable, international long-distance services require a considerable demand, which can only be assured by major urban centres with specific characteristics. In fact, the supply of international long-distance services, especially in the aviation and high-speed markets, is generally liberalised and not subject to PSOs with a few exceptions only (mostly services used to serve infrastructurally isolated areas such as islands). Consequently, in order to assure the long term sustainability of a supplied service the economic profitability of the specific route is fundamental. Thus, to select the urban areas to include in the network some parameters to define their attractiveness in terms of long-distance international transport have been identified. Among these, population, GDP, labour market attractiveness, touristic attractiveness, and intra-EU migrations. Furthermore, two strict criteria, the former related to the presence of at least one international airport within the catchment area, as defined in Section 4.6 and the latter related to the presence of inter-regional rail services, such as intercity were placed. Based on the resulting cities a final geographical coverage criteria is going to be applied to ensure that the selected cities represent comprehensively and coherently every country in the continent. Finally, the selection is skimmed based on data availability. The actual process of selection applied to the case study is further described in Section 4.4.

This research focuses on the supply side, thus demand is not directly modelled. However, as previously mentioned, in selecting the urban areas to include in the network it is important to consider the feasibility of such connections. And thus to consider only urban areas that are able to generate and attract enough international, long-distance trips. In general in basic transport models, such as the 4-step one, travel demand is determined in relation to trip generation and trip attraction between respective OD pairs. In case of urban or regional transport models the demand is thus forecast based on residential and employment characteristics, such as the job numbers and type, and the household numbers and type. This follows the assumptions that the trips made by the passengers mostly relate to outbound home-borne trips in the morning, directed to work places or other attractors, such as commercial areas, and to inbound trips that bring the agents back to their house in the evening. Thus, the standard travel behaviour modelled is the one of daily commuters. However, in the case of long-distance international trips the travel behaviour that needs to be captured is radically different, mostly due to the distinct trip purposes and drivers leading to the necessity of moving. Long-distance international passengers rarely complete tours in one day (leaving home in the morning and coming back home in the evening) as the longer travel times means that it would be hardly efficient in terms of balance between the time lost travelling compared to the time spent at the final destination. And in the few cases when that might happen (i.e. daily business trips) the behaviour will not repeat on a daily basis, but rather on a weekly or monthly basis. In this instance, Rich and Mabit (2012) state that holiday, business and personal are the main trip purposes for long-distance international travel, consequently highlighting the importance of tourism and labour market in the definition of travel demand.

Following this, the first five criteria are defined. Population and GDP aim to capture the general production and attraction of international cross-border trips, following the assumption that demand levels are higher in urban areas with more population and higher GDPs. Furthermore, labour market attractiveness, touristic attractiveness, and intra-EU migrations aim to capture the specific travel behaviour for long-distance trips by reflecting the three main trip purposes as defined by Rich and Mabit (2012). Labour market attractiveness has been chosen as a measure to represent the intensity of business trips, touristic attractiveness the intensity of holiday trips, and intra-EU migrations the intensity of personal trips (following the assumption that most passengers moving internationally within Europe for personal reasons are expats). In relation to what discussed in this section, the main assumptions are summarised in the following points:

- It is assumed that in order to be practically sustainable, international long-distance services require a considerable demand, which can only be assured by major urban centres with specific characteristics.
- It is assumed that these characteristics relate to the statistical indicators selected, meaning that demand levels are assumed to be higher in urban areas with higher population, GDP, labour market attractiveness, touristic attractiveness and population of EU expats.

Using the methodology explained in this section urban areas included in this study are selected, as further explained in Section 4.6. The following inputs are used throughout this process:

- Population per urban area
- GDP per urban area
- Labour market attractiveness per urban area
- Touristic attractiveness per urban area
- Population of EU expats

The output of this process is summarised in the following point:

- List of selected urban areas

3.1.4. Airport Selection Criteria

As previously mentioned this research aims to compare the performances of air and rail transport. Thus, to ensure compatibility and comparability between the two networks it was deemed fundamental to employ the same nodes in both instances. These nodes are set to represent urban centres as determined through the methodology outlined in Section 4.4. However, whilst railway stations are generally located within urban centres, airports due to space constraints are typically located in the outskirts, sometimes even dozens or hundreds of kilometres away from the heart of the city. Furthermore, in more densely inhabited areas airports might serve more than one urban centre and at the same time larger urban areas might feature more than one airport. Thus, in order to select the commercial airports to include in the research and to assign them to the urban areas selected, it was made use of catchment areas.

Literature highlights that there are two dominant methods to measure catchment areas of specific locations, be them airports or urban centres (Augustyniak & Olipra, 2014; Lieshout, 2012; Pavlyuk, 2009; Rothfeld et al., 2019). Distance-radius maps, employing isodistances to represent the area within a certain distance from a location, and time-radius maps employing isochrones to represent the area reachable within a certain time. Literature generally agrees in considering the latter of the two as the more relevant indicator, in fact, as Rothfeld et al. (2019) note, small spatial distances do not always reflect in short temporal distances (Pavlyuk, 2009). A third approach, from Dziedzic et al. (2020), assumes NUTS2 regions to be a reliable approximation of airport catchment areas. A similar method is used by Dobruszkes et al. (2011) who associate each FUA to the airports in close proximity. In this regard, it is important to highlight that all these methods represent an important simplification of reality. In fact, many authors agree that catchment areas of airports are not really static being affected by a wide range of variables and factors, such as the drivers of airport passenger choice (Lieshout, 2012). In this regard, Dziedzic et al. (2020) highlight that airports' attractiveness is influenced by the ground side accessibility, the available offer and the passengers type. Dobruszkes et al. (2011) add to this suggesting that airports catchment areas also depend on the territorial morphology and the presence of high speed services. Furthermore, the study from Lieshout (2012) has shown that catchment areas are also heavily influenced by the final destination. Another important factor highlighted by many authors is competition and consequently the overlap between different airport catchment areas (Augustyniak & Olipra, 2014; Dobruszkes et al., 2011; Lieshout, 2012; Lieshout et al., 2016; Pavlyuk, 2009). Dobruszkes et al. (2011) further underline that airport catchment areas heavily depend on fares and airline types. All these factors, come together influencing passengers' airport choice and consequently causing contradictory impacts on catchment areas. Given the extreme complexity of the matter is deemed to be outside of the scope of this research, only the time variable has been considered in computing the catchment areas. In fact, Welch et al. (2015) and Sun et al. (2017) argue that access travel time is one of the most important factors in determining the attractiveness of an airport.

To guarantee a sufficient level of simplicity, it was decided to fix a standard time distance. Also in this case, literature is rather divided about which value represent an accurate estimate. However, researchers widely agree in considering airport catchment areas to be generally ranging from 1 up to 3 hours by car. In this regard, Suau-Sanchez and Voltes-Dorta (2019) employ 1h, 2h and 3h radii to identify catchment areas. On the other hand, Lieshout (2012) and Lieshout et al. (2016) argue that in some cases airport catchment areas might expand even further, highlighting that despite this the market shares on those distances are usually very low. Marcucci and Gatta (2011) assert that the approximated catchment area of an airport is "whatever resides within a 2h drive by car", value which is also defined by

Dziedzic et al. (2020) as the traditionally used one. Augustyniak and Olipra (2014) state that a common approach is to define catchment areas as geographical radii of 1h or 2h travel time around the airport. Also Malina (2010) develops different catchment areas of up to 2h by car from the airport differentiates between segments of passengers and traffic types. However, the European Commission believes that a 1h time-radius is a rather conservative estimate (European Commission, 2013). Thus, the scope of the model being limited to international and continental traffic, an estimate of 90min or 1,5h was chosen as a reference limit for the catchment areas' range. It is interesting to note that Poelman (2013) also employs 90min polygon-shaped areas to measure the accessibility to passenger flights in Europe.

Using this methodology airports are associated to urban areas, as further explained in Section 4.6. The following inputs are used throughout this process:

- List of commercial airports
- List of selected urban areas
- Catchment areas

Furthermore, the output of this process is summarised in the following point:

- List of airports associated per Urban Area

3.1.5. Travel Time Composition

Given the considerable differences in terms of travel time characteristics between the two modes due to the distinct procedures at the terminals, air having generally shorter in-vehicle times and larger waiting, access and egress times compared to rail, it is deemed crucial to translate them into the model to allow full comparability between the actual performances of the two modes. Thus, rather than simply considering the in-vehicle time, the total door-to-door travel time is taken into account. That is, in fact, the travel time measure upon which passengers generally rely in order to make their travel decisions (Peer et al., 2013). To model the door-to-door trip some extra travel time components are added, as specified in Figure 3.4. In particular, the total travel time tt_{ij} between an average house located in urban area i to an average house located in urban area j is given by the sum of five separate travel time components. The time required to access the origin terminal (airport or station) o (1), the waiting time before departure in the origin terminal o (2), the in-vehicle time between origin o and destination d terminals (3), the waiting time at destination terminal d (4), and the egress time from destination terminal d (5), as illustrated by formula 3.1.

$$tt_{ij} = tt_{io}^{acc} + tt_{o}^{waitdep} + tt_{od}^{in-veh} + tt_{d}^{waitarr} + tt_{dj}^{egr} \tag{3.1}$$

To further simplify the model it was decided to fix some of these components to standardised values based on practical experience and on the recommendations provided by airports, airlines, train stations, and train companies. In regards to access/egress times, it was decided to keep the values for air case-specific, as airports are generally located at various distances from urban areas. These are computed following the distances identified within the catchment areas as defined in Section 3.1.4. On the other hand, rail stations are generally located in highly accessible areas, oftentimes in the proximity of city centres. Thus, the values for access/egress time of rail stations have been fixed at 20 min as it was assumed that in most urban areas that is a fair estimate of the average time required to reach the main station(s). The decision to fix the waiting time at departure for air at 120 min is based on the minimum requirements suggested for international European flights to be in the airport at least 90 min in advance to allow for check-in, luggage loading, security controls and boarding procedures. Further 30 minutes are added to take into account of eventual congestion and the longer times required in many cases to access airports, as access/egress times are computed based on the fastest modal alternative. The waiting at departure for rail is fixed at 25 minutes, time which is generally suggested as the standard time to reach stations when travelling with international trains within Europe. Waiting time at arrival are much shorter for both modes, being fixed at 5 min for rail, considering the time to disembark the train and exit the station and at 30 min for air as the processes are longer due to the bigger size of terminals and the longer disembarkment and luggage unloading procedures. Finally, in-vehicle times are case-specific and are obtained from actual data of scheduled transport services, as further explained in Section 3.2.

The values have thus been checked and refined based on the assumptions made by different studies across the literature. Avogadro et al. (2021), for instance, approach for the specific case of Bruxelles-Paris link, assuming an average access/egress time of around 15 min for high-speed rail and of around 35 min for air, a departure waiting time of 60 min for air and 15 min for rail, and a transfer time at destination of around 5 min for both modes. It is important to point out that Avogadro et al. (2021) retrieves the door-to-door travel times directly from a journey planner app. Thus these values are generally deemed to be slightly underestimating the actual conditions of transport and do not really take into account an average house within the urban area but rather the city centre, which works well for smaller centres but is less reliable for major urban areas. Slightly higher figures are employed by Zhu, Zhang and Zhang (2019), who studying the connectivity of air and rail networks in China, assume waiting times of 90 or 120 min (domestic and international flight) for air and of 30 min for rail at the origin and of 30 min for air and of 15 min for rail at the destination. The values first assumed by this study generally lie between the figures identified by Avogadro et al. (2021) and the ones defined by Zhu, Zhang and Zhang (2019) and are thus deemed sensible enough. The final travel time components per mode are illustrated in Figure 3.4.

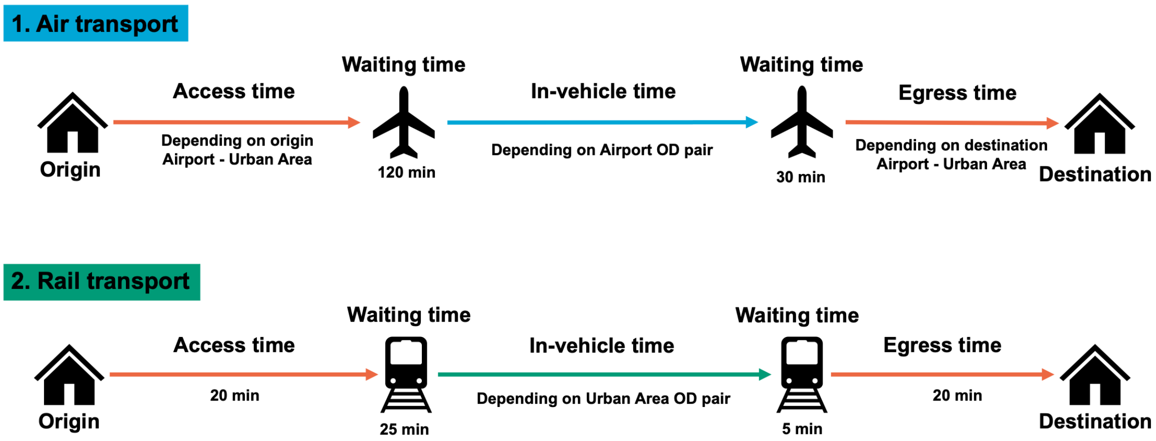


Figure 3.4: Overview of the Travel Time Composition per Mode

Given that travel time is employed as link label within the networks, and considering that, being represented in P-Space topology, these only include direct services, travel time composition is modelled primarily for direct services. For indirect connections, the shortest weighted path in terms of in-vehicle time is substituted to the in-vehicle time of the direct connection. This implies that transfer times are not directly included within the in-vehicle time components. Thus, an additional transfer time parameter is added for indirect connections. In relation to what discussed in this section, the main assumptions are summarised in the following points:

- It is assumed that travellers take into account door-to-door travel time rather than in-vehicle time when planning their journeys.
- It is assumed that travellers' preferences are independent from the time of the day of departure and arrival (peak/off-peak), and the access/egress mode.
- The access and egress times for rail are given fixed values and thus are assumed to be comparable across all urban areas.

3.1.6. In-Vehicle Time Modal Sensitivity

On the one hand, travel time represents an objective measure as it can be precisely measured as the interval of time elapsing between the departure at the origin and the arrival at the destination. On the other, there is an important subjective component, the individual perception of travel time, which over the last decades has been increasingly attracting the attention of scholars , especially through the lens of behavioural sciences, economics, transport geography, public health and social psychology (Cornet et al., 2022; van Acker et al., 2010). This is an important paradigm in long-distance travel, due to the considerable differences in the total travel time and in the specific travel time components across different transport modes. Travellers, in fact, appear to have different sensitivity to travel time changes when travelling with different transport modes, due to the different level of comfort that characterise each mode and to the specific period of the day when the trip takes place. In particular, Witlox et al. (2022) highlights the relationship between comfort level and the perception of travel time arguing that higher level of comfort reduce the sensitivity of passengers to travel time. In this regard, Malichová et al. (2022) argue that from a passenger perspective, the train, enabling a wider range of activities (e.g. working, browsing the internet...), allows a more beneficial use of travel time compared to air services. This, in turn, implies that the passenger's sensitivity to travel time changes is generally lower for train compared to planes. Similarly, it could be argued that night trains, enabling passengers to sleep during the entire journey and to arrive rested at the final destination, feature a lower travel time sensitivity compared to day trains (Heufke Kantelaar et al., 2022). Thus, in order to capture the specific travel time sensitivity of the different travel modes some weights are defined for day train, night train and plane. These are consequently employed to discount the in-vehicle travel times of the modes that are less sensitive to travel time changes. To derive these values some ratios between the travel time sensitivity are taken from the literature. The study from Heufke Kantelaar et al. (2022) provides travel time parameters for night train (−0,651) and plane (−1,45), allowing to compute the ratio between the sensitivity of night trains and plane (0,44). Furthermore, the study from Román et al. (2010) provides the willingness to pay for travel time savings for high speed rail (25,68) and plane (34,22) allowing to compute the ratio between the sensitivity of plane and day trains (1,33). Thus, the ratio between night trains and day train is obtained by multiplying the two aforementioned ratios (0,58). The three obtained weights, 0,58 for night train, 1 for day train and 1,33 for plane are then re-scaled to the max-

Table 3.2: Overview of the selected sensitivity weight per mode

Mode	Night Train	Day Train	Plane
Weight	0,44	0,75	1

imum value obtaining 0,44 for night train, 0,75 for day train and 1 for plane. The selected sensitivity weight per mode are highlighted in Table 3.2.

In relation to what discussed in this section, the main assumption is summarised in the following point:

- It is assumed that the behaviour of all passengers, and consequently the individual perception of travel time, is comparable, thus differences based on trip purpose (i.e. business, personal, leisure) or personal characteristics (e.g. age, gender, car ownership) are not taken into account.

3.2. Data Collection

Following the definition of the model, as explained in Section 3.1, the specific data requirements are determined. Throughout Section 3.1.2 the factors to include in the analysis have been identified in Weekly Frequency, Average Travel Time and Number of Transfers. The specific travel time components have been further defined and specified in Section 3.1.2. The data required for the analysis refers to the three aforementioned parameters that are going to be used as link weights of the networks and within the connectivity indicator. In particular, the number of transfers can be retrieved using the networks, leaving only the out-of-vehicle and in-vehicle time and weekly frequency. The former are fixed for rail whilst for air the access/egress times are retrieved manually using Google Maps, in order to take into account of the specific conditions of each airport-urban area pair. Furthermore, the latter two are retrieved through air and rail service schedules, as described in Section 3.2.1 and 3.2.2. Finally the data is re-aggregated from a terminal level to a urban area/node level using the methodology explained in Section 3.2.3. All the data retrieved has been cleaned using the NumPy and Pandas libraries for Python.

3.2.1. Air Network Data

As highlighted by Bombelli (2020), data on air service schedules is scarce, not openly accessible and difficult to retrieve. Thus, the data has been manually collected through the use of a publicly available real-time aviation data provider, Flight Aware. In particular, the flight schedules between each airport OD pair are retrieved through its “search by route” engine. This tool reports all the flights departing from a certain origin airport and arriving at a certain destination airport for a week time, from five days before to two days after the current date. In particular, the data was collected on July 22nd 2022. To make the data collection process faster and less cumbersome the airport list was defined based on the requirements of this specific research, as explained in Section 4.6. Thus not all European airports have been included. For each origin airport, the data has been collected by iterating over all the possible destination included in the airport list using the ICAO codes of the two airports. An extract of the data concerning the departures from Wien airport (VIE/LOWW) is reported in Table 3.3. Each entry represents a scheduled service and is characterised by an airline, a flight number, an origin airport, a destination airport (both IATA and ICAO), a departure time, and an arrival time (estimated time of arrival in case of flights still due to depart).

Despite the output data had already been pre-filtered using the filters made available by Flight Aware, it is possible to notice how entries do not represent only commercial services, including some general aviation flights. Given that this study focuses on commercially available air services, private flights need to be removed. Thus, the data is further cleaned using an exhaustive list of all commercial airlines in Europe to filter the entries. Moreover, the timestamps are converted to a timedate format to allow to retrieve the in-vehicle time. Time zones are then matched across all the dataset and the in-vehicle

Table 3.3: Extract of the data concerning the departures from Wien airport (VIE/LOWW)

Airline	Flight Number	Origin_ICAO	Destination_ICAO	Destination_IATA	Departure Time	Arrival Time
Pink Sparrow "Sparrow"	SOW6	LOWW	LOWL	LNZ	Tue 19:12 CEST	19:26 CEST Tue
Aero-Jet "Swissjet"	AOJ37R	LOWW	LOWL	LNZ	Sun 16:25 CEST	17:10 CEST Sun
GlobeAir "GlobeAir"	GAC515V	LOWW	LOWL	LNZ	Fri 13:52 CEST	14:24 CEST Fri
Salzburg Jet Aviation "Mozart"	MOZ118	LOWW	LOWS	SZG	Fri 17:03 CEST	17:29 CEST Fri
Austrian Airlines	AUA975	LOWW	LOWG	GRZ	Fri 22:45 CEST	23:20 CEST Fri
Austrian Airlines	AUA977	LOWW	LOWG	GRZ	Fri 10:05 CEST	10:40 CEST Fri
Austrian Airlines	AUA975	LOWW	LOWG	GRZ	Thu 22:45 CEST	23:20 CEST Thu
Austrian Airlines	AUA977	LOWW	LOWG	GRZ	Thu 10:05 CEST	10:40 CEST Thu
Austrian Airlines	AUA975	LOWW	LOWG	GRZ	Wed 22:45 CEST	23:20 CEST Wed
Austrian Airlines	AUA977	LOWW	LOWG	GRZ	Wed 13:00 CEST	13:40 CEST Wed

time is calculated in minutes as the difference between arrival and departure time. An extract of the cleaned data concerning the departures from Wien airport (VIE/LOWW) is provided in Table 3.4. It is possible to notice how only commercial services (i.e. austrian airlines) are now present in the data. it seems important to point out that entries with the same flight number are not filtered because, the identifications are provided per route scheduled on a daily basis and consequently do not differ across different days of the week.

Table 3.4: Extract of the cleaned data concerning the departures from Wien airport (VIE/LOWW)

Airline	Flight Number	Origin_ICAO	Destination_ICAO	Destination_IATA	Departure Time	Arrival Time	In-vehicle Time
Austrian Airlines	AUA975	LOWW	LOWG	GRZ	20:45	21:20	35
Austrian Airlines	AUA977	LOWW	LOWG	GRZ	08:05	08:40	35
Austrian Airlines	AUA975	LOWW	LOWG	GRZ	20:45	21:20	35
Austrian Airlines	AUA977	LOWW	LOWG	GRZ	08:05	08:40	35
Austrian Airlines	AUA975	LOWW	LOWG	GRZ	20:45	21:20	35
Austrian Airlines	AUA977	LOWW	LOWG	GRZ	11:00	11:40	40
Austrian Airlines	AUA963	LOWW	LOWG	GRZ	13:40	14:20	40
Austrian Airlines	AUA963	LOWW	LOWG	GRZ	13:40	14:20	40
Austrian Airlines	AUA901	LOWW	LOWI	INN	15:15	16:10	55
Austrian Airlines	AUA903	LOWW	LOWI	INN	07:35	08:30	55

Finally, each unique OD-pair is grouped, averaging the in-vehicle time and counting the number of occurrences of that OD-pair in the dataset (weekly frequencies of flights).

3.2.2. Rail Network Data

The data used to create rail networks has been collected through the UIC MERITS (Multiple East-West Railways Integrated Timetable Storage), a database containing the integrated timetable data of many European countries which is used to provide information to journey planners and ticket-selling websites. The access to this commercial UIC-owned database is provided by Hacon within the framework of the “Rail connectivity index” project developed in consortium with Ecorys for the Directorate-General for Mobility and Transport (DG MOVE) of the European Commission. The decision to use this commercial database, rather than manually retrieve the data from journey planner websites, relates to the many inconsistencies and considerable discrepancies detected in the data available throughout the different journey planners. Notable journey planners that have been queried and whose results have been compared include DB (Deutsche Bahn), Trainline, Rail Europe, Eurail/Interrail, Omio, Rome2Rio, and ÖBB (Österreichische Bundesbahnen).

Table 3.5: Extract of the data concerning rail scheduled services

Origin_ID	Destination_ID	Origin Date	Origin Time	Destination Date	Destination Time	Duration	Products	Operators	Service_ID
5100750	5103060	2022-07-03	16:14:00	2022-07-03	19:19:00	PT3H5M	IC,Intercity	PKP Intercity S.A.	MXwxNjE5OTV8MHw4M3wzMDcyMDIy
5100750	5103060	2022-07-03	18:15:00	2022-07-03	21:32:00	PT3H17M	IC,Intercity	PKP Intercity S.A.	MXwxNjIwNDI8MHw4M3wzMDcyMDIy
5100750	5103060	2022-07-03	19:52:00	2022-07-03	23:07:00	PT3H15M	IC,Intercity	PKP Intercity S.A.	MXwxNjIwOTB8MHw4M3wzMDcyMDIy
5100750	5103060	2022-07-03	22:38:00	2022-07-04	02:04:00	PT3H26M	TLK,Yours Rail Lines	PKP Intercity S.A.	MXwxNTkyODJ8MHw4M3wzMDcyMDIy
5100750	5103350	2022-06-27	05:56:00	2022-06-27	08:45:00	PT2H49M	EIP,High-speed train	PKP Intercity S.A.	MXwxNTg4MzZ8MHw4M3wyNzA2MjAyMg==
5100750	5103350	2022-06-27	07:00:00	2022-06-27	09:45:00	PT2H45M	EIP,High-speed train	PKP Intercity S.A.	MXwxNTg4NTJ8MHw4M3wyNzA2MjAyMg==
5100750	5103350	2022-06-27	07:56:00	2022-06-27	10:49:00	PT2H53M	EIP,High-speed train	PKP Intercity S.A.	MXwxNTk0NzB8MHw4M3wyNzA2MjAyMg==
5100750	5103350	2022-06-27	08:25:00	2022-06-27	11:35:00	PT3H10M	IC,Intercity	PKP Intercity S.A.	MXwxNTg0MDF8MHw4M3wyNzA2MjAyMg==
5100750	5103350	2022-06-27	09:02:00	2022-06-27	11:45:00	PT2H43M	EIP,High-speed train	PKP Intercity S.A.	MXwxNzI1MjI8MHw4M3wyNzA2MjAyMg==
5100750	5103350	2022-06-27	09:44:00	2022-06-27	12:32:00	PT2H48M	EIP,High-speed train	PKP Intercity S.A.	MXwxNTk0NzZ8MHw4M3wyNzA2MjAyMg==
5100750	5103350	2022-06-27	10:56:00	2022-06-27	13:45:00	PT2H49M	EIP,High-speed train	PKP Intercity S.A.	MXwxNzI4Njh8MHw4M3wyNzA2MjAyMg==
5100750	5103350	2022-06-27	12:57:00	2022-06-27	15:44:00	PT2H47M	EIP,High-speed train	PKP Intercity S.A.	MXwxNTg4ODI8MHw4M3wyNzA2MjAyMg==
5100750	5103350	2022-06-27	13:54:00	2022-06-27	16:45:00	PT2H51M	EIP,High-speed train	PKP Intercity S.A.	MXwxNTk0ODI8MHw4M3wyNzA2MjAyMg==
5100750	5103350	2022-06-27	14:02:00	2022-06-27	17:41:00	PT3H39M	TLK,Yours Rail Lines	PKP Intercity S.A.	MXwxNzIyNTZ8MHw4M3wyNzA2MjAyMg==
5100750	5103350	2022-06-27	14:56:00	2022-06-27	17:49:00	PT2H53M	EIP,High-speed train	PKP Intercity S.A.	MXwxNTg5MDI8MHw4M3wyNzA2MjAyMg==

The data collected includes all the train services available between each station OD pair in the week between June 27th 2022 and July 3rd 2022. Furthermore, data for the previous week, from June 20th 2022 to June 26th 2022 is retrieved to check that there are no gaps in the data. Each entry represent a scheduled train service from an origin station (with UIC code ORIGIN_ID) to a destination station (with UIC code DESTINATION_ID) and is characterised by an origin date, an origin time, a destination time, a destination date, a duration, a type of service (PRODUCTS), an operator and a unique service_ID. An extract of the data concerning rail scheduled services is reported in Table 3.3. The data includes also

Table 3.6: Data manually added

Origin	Destination	Origin Name	Destination Name	Average In-vehicle Time	Frequency
9431039	7160000	Lisbon	Madrid	585	7
7160000	9431039	Madrid	Lisbon	585	7
9402006	7131400	Porto	Santiago de Compostela	217,5	14
7131400	9402006	Santiago de Compostela	Porto	217,5	14
7216052	7300415	Belgrade	Thessaloniki	422,4	7
7216052	7872480	Belgrade	Zagreb	282,75	14
7216052	7942300	Belgrade	Ljubljana	380,15	14
7300415	7216052	Thessaloniki	Belgrade	422,4	7
7872480	7216052	Zagreb	Belgrade	282,75	14
7942300	7216052	Ljubljana	Belgrade	380,15	14
5103350	2412310	Warsaw	Kaunas	355,5	7
2412310	5103350	Kaunas	Warsaw	355,5	7
7111511	8727100	San Sebastián	Paris	255	42
7111511	8758100	San Sebastián	Bordeaux	157,5	70
7111511	8761100	San Sebastián	Toulouse	232,5	7
8727100	7111511	Paris	San Sebastián	255	42
8758100	7111511	Bordeaux	San Sebastián	157,5	70
8761100	7111511	Toulouse	San Sebastián	232,5	7
7300415	5216000	Thessaloniki	Sofia	382,5	7
5216000	7300415	Sofia	Thessaloniki	382,5	7

regional and suburban services which are filtered out, given that the scope of the study is limited to long-distance transport. Furthermore, given that some urban areas featured more than one station was included, it was required to make sure that train services stopping in multiple stations within one city are counted only once. Thus, the data was filtered using the service_ID to limit the entries to unique train services only. Finally, also in this case the data for each unique OD-pair is grouped, averaging the in-vehicle time and counting the number of occurrences of that OD-pair in the dataset (weekly frequencies of train services). Due to some gaps in the available data, some cross-borders connections have been manually added to ensure that the entire network in interconnected. A complete list of the data manually added is reported in Table 3.6. It is important to note that some of these services have been suspended due to the Covid-19 pandemic and have currently not resumed (Lisbon to Madrid and Belgrade to Thessaloniki, Zagreb and Ljubjana). Others, on the other hand, are not direct connections but include a transfer at the border, Warsaw to Kaunas in Białystok, San Sebastián to Paris, Bordeaux and Toulouse in Irun-Hendaye, and Porto to Santiago de Compostela in Vigo. In these cases the total travel time from origin station to destination station, including the specific transfer time based on schedule synchronisation is employed as average in-vehicle time. It is important to note that average in-vehicle times provided in table 3.6 already consider the type of service (day train, night train) being weighted using the in-vehicle time sensitivity parameters defined in Section 3.1.6. The basic data for these extra options is retrieved either through the aforementioned journey planners or using the websites of the national railway undertakings.

3.2.3. Re-aggregation

As mentioned in the previous Sections, the model used to represent reality employs unique urban areas as nodes of the networks. However, the input data (travel time and frequency) is collected at the level of single terminals (station/airport). It is thus required to re-aggregate the travel times and frequencies between each terminal OD-pair od to each urban area OD-pair ij. To do so, urban areas are first associated to the accessible airports within the catchment area using the methodology highlighted in Section 4.6 and to the accessible rail stations as specified in Section 4.7, and then the data collected at the terminal level is re-aggregated at the urban area one. The process is graphically rendered in Figure 3.5. This step differs across the two modes given that, as explained in Section 3.1.5, the access/egress times are fixed for rail, whilst in the case of air they are dependent on the specific airport - urban area. For rail the process is consequently quite simple and straightforward as it is not needed to take into account any specificity of individual stations when aggregating the data. Thus, in-vehicle time tt^{in-veh} of every direct available route r_{od} is averaged and the frequency f of each route r_{od} is summed, as per Formulae 3.2 and 3.3.

$$\overline{tt}_{ij}^{rail} = \frac{\sum_{r_{od} \in R} tt_{r_{od}}^{in-veh}}{|R|} \tag{3.2}$$

$$f_{ij}^{rail} = \sum_{r_{od} \in R} f_{r_{od}} \tag{3.3}$$

where,

$i, j \in I$: set of unique nodes in the network, each representing a specific Urban Area (represented in the graphs with an orange dot).

$o, d \in T$: set of terminals (airports/stations) associated to each urban area.

$r_{od} \in R$: set of all available direct routes r connecting origin i and destination j through the respective

terminals o and d.
$\overline{tt}_{ij}^{rail}$: average in-vehicle time of rail services between origin i and destination j.
f_{ij}^{rail} : total frequency of rail services between origin i and destination j.
$tt_{r_{od}}^{in-veh}$: average in-vehicle time of services between origin station o and destination station d.
$f_{r_{od}}$: average frequency of services between origin station o and destination station d.

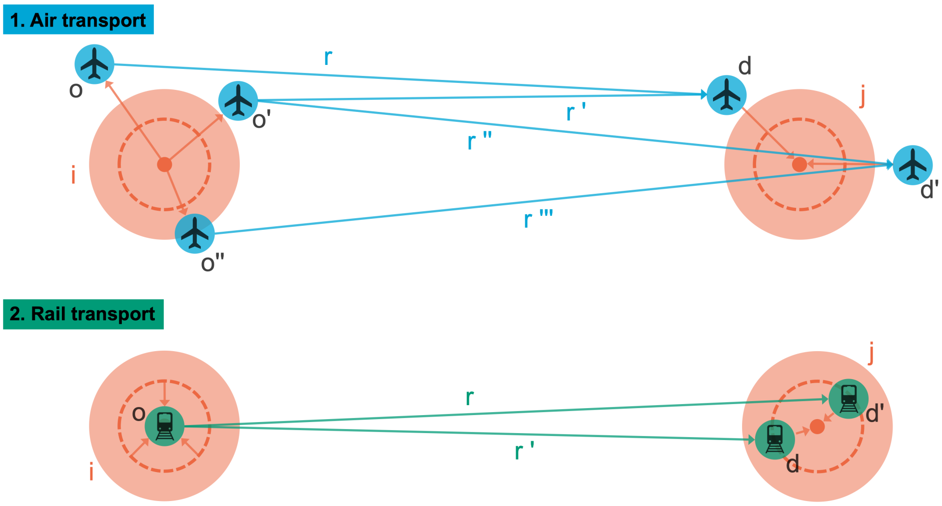


Figure 3.5: Basic representation of re-aggregation variables

In the case of air it was decided to take into account the attractiveness of specific terminals when re-aggregating the data, in an effort to capture the influence of different airports characteristics (i.e, size and location/accessibility) on route choice. In fact, differently from train stations, airports have variable characteristics in terms of traffic (frequency) and accessibility (access/egress times), which passengers take into account when selecting a specific route. Thus, this step aims to capture the inconvenience of using routes that are longer and less frequent. To model this, frequencies on a specific route are used as weights to average in-vehicle time, as shown in Formula 3.4 and the total travel time of each route is used as a measure to discount the frequencies of longer routes, as illustrated by Formula 3.5. This implies that the among all the direct available routes between i and j, the route r_{od}, using origin terminal o and destination terminal d, with the shortest total travel time will have all the frequencies included, whilst longer routes will add to that only a up to an extent. This is particularly useful to penalise routes that use distant terminals when terminals located closer to the urban area are available.

$$\overline{tt}_{ij}^{air} = \frac{\sum_{r_{od}\in R} f_{r_{od}} tt_{r_{od}}^{in-veh}}{\sum_{r_{od}\in R} f_{r_{od}}} \tag{3.4}$$

$$\tilde{f}_{ij}^{air} = \sum_{r_{od}\in R} \hat{tt}_{r_{od}} f_{r_{od}} \tag{3.5}$$

where,
$\overline{tt}_{ij}^{air}$: average in-vehicle time of air services between origin i and destination j.
$\tilde{f}_{ij}^{air}$: total frequency equivalents (not necessarily integer) of air services between origin i and destination j.

In particular, the travel time discount $\hat{tt}$ per route represents the ratio of the minimum total travel time between origin i and destination node j to the total travel time of a specific route. Total travel times, as per definition provided in Section 3.4, are used to include the travel time on the entire root.

$$\hat{tt}_{r_{od}} = \frac{\min_{r_{od} \in R} tt^{total}_{r_{od}}}{tt^{total}_{r_{od}}} \tag{3.6}$$

Furthermore, the total weekly frequency f^{total} per airport o is used to weigh the average access/egress time per urban area, as illustrated by Formula 3.4.

$$\overline{at}^{air}_i = \frac{\sum_{o \in T} f^{total}_o at_{io}}{\sum_{o \in T} f^{total}_o} \tag{3.7}$$

where,
$\overline{at}^{air}_i$: weighted average access/egress time of urban area i.
f^{total}_o : average frequency per airport o.
at_{io} : average access/egress time between urban area i and airport o.

In conclusion, the following inputs are used throughout this process:

- List of terminals (airports/stations) associated per Urban Area
- Airport-Urban Area specific access/egress times
- Average in-vehicle travel time per terminal OD pair
- Weekly frequency per terminal OD pair

Furthermore, the outputs of this process are summarised in the following points:

- Weighted average in-vehicle time per Urban Area OD pair (rail and air)
- Weekly frequency/frequency equivalents (rail and air)
- Weighted average access/egress time per Urban Area (air only)

3.3. Network Analysis

The collected data is thus employed to construct the OD matrices and P-space undirected networks required to analyse the network performances, connectivity and significant components. First of all, in Section 3.3.1, link connectivity is first defined and a novel indicator to measure it is presented. Secondly, based on the literature review provided in Section 2.3, some topology indicators employed by Network Science to analyse network structure are briefly introduced and discussed in Section 3.3.2. Next, the metrics employed to capture connectivity at the node level are presented in Section 5.4.3. Then, the metrics employed to identify significant components in the networks are highlighted in Section 3.3.4. Finally, the relationships between OD pairs are analysed and categorised using the methodology developed and explained in Section 3.3.5. It is worth noting that every step of the network analysis is developed using Python and its library NetworkX is employed to construct the networks and compute some indicators, whilst other are manually computed.

3.3.1. Link Connectivity

Following the analysis of connectivity metrics described in Section 2.2.5, a novel connectivity indicator is proposed. In particular, the link connectivity LC_{ij} between an origin i and a destination j is defined, within this study, as the ratio of connection intensity to travel inconvenience, as per Formula 3.8, where the former is a function of travel time and the latter is a function of frequency. Furthermore, link resistance LR_{ij} is defined as the reciprocal of link connectivity, as per per Formula 3.9. The two components CI_{ij} and TI_{ij} are put into relation using a ratio in order to provide a measure of the relationship incurring between their magnitudes. Using a perhaps simpler linear combination of the elements would have posed some issues due to the different relationships incurring between the different components and connectivity. In fact, connection intensity and connectivity are positively related, meaning that an increase in frequency leads to an increase in connectivity, whilst travel inconvenience and connectivity have a negative relationship, as the latter tend to decrease with higher travel times. Furthermore, the use of a ratio allows to enhance the influence of the former component on connectivity, and of the latter on resistance. In fact, an important implication of the use of a ratio regards the different distribution of values. When the numerator is greater than the denominator the ratio is distributed between 1 and infinite, whilst in the opposite case it distributes between 0 and 1. This implies that on a linear scale greater importance is given to the cases where the numerator is greater than the denominator. In this specific case, this means that the parameter that influences the most the link connectivity index is connection intensity and consequently frequencies, as opposed to travel times. This appears to be in line with the characteristics of many connectivity indicators found throughout the literature, most notably the one defined by Zhu et al. (2018). In regards to resistance the opposite holds, as travel time is at the numerator. This is deemed sensible as across the literature shortest paths appear to be computed mostly on networks weighted with travel times (Luo et al., 2019).

$$LC_{ij} = \frac{CI_{ij}}{TI_{ij}} \tag{3.8}$$

$$LR_{ij} = (LC_{ij})^{-1} = \frac{TI_{ij}}{CI_{ij}} \tag{3.9}$$

The connection intensity component CI_{ij}, used as numerator in the connectivity formula, represents the intensity of the connection between two urban areas and is measured by the link frequency, in case of direct connections, or by the effective frequency, in case of indirect connections. Effective frequency represents the frequency of indirect connections and is measured by the ratio of the minimum frequency among all the shortest path's legs to the number of transfers. In particular, the effective frequency of indirect connections is assumed to be restrained by the minimum frequency, as lower frequencies on a certain link limit the opportunities to take full advantage of the higher number of frequencies available on the remaining legs. Furthermore, including the number of transfers allows to penalise frequencies of indirect connections compared to direct equivalents.

$$CI_{ij} = \theta_f \; f_{ij} \tag{3.10}$$

The travel inconvenience component TI_{ij}, used as denominator in the connectivity formula, represents the degree of inconvenience that passengers incur into when using a certain link, in case of direct connections, or a series of links, in case of indirect connections. This parameter, which is somehow comparable to the travel impedance metric employed by Luo et al. (2019) for PT, is measured by three main components, namely in-vehicle time, out-of-vehicle time and transfer time (only for indirect connections), as shown by Formula 3.11. For direct connections the in-vehicle time is measured using the data reaggregated using formulae 3.2 and 3.4 for rail and air respectively whilst for indirect connection it is calculated by summing the in-vehicle time for each service leg. The out-of-vehicle time is computed for both direct and indirect connections by summing access, egress and waiting time as defined in Section 3.1.5. Indirect connections are thus modelled as direct connections with the exceptions of the transfer time parameter that is included to provide a measure of the inconvenience that transfers procure on passengers.

$$TI_{ij} = \theta_{in-veh}\, tt_{ij}^{in-veh} + \theta_{out-veh}\, (tt_{ij}^{acc} + tt_{ij}^{wait} + tt_{ij}^{egr}) + \theta_{trf}\, tt_{ij}^{trf} \tag{3.11}$$

Indirect connections, requiring one or more transfers along the route, imply that passengers must spend additional time to reach their destination. Other than the total in-vehicle time (sum of the in-vehicle time for all travel legs), the access and egress times from the origin and final destination terminals, and the waiting time at both terminals, passengers spend time also in transfer terminals. In particular, they need to egress the incoming service, access the outgoing service and wait for its departure, time which, depending on the schedule synchronisation of the two services, can vary considerably. To model this, a specific transfer time component is included for every stop along the route. For this scope the formula defined by A. K. Sen and Morlok (1976) to estimate the average waiting times for indirect connections of intercity public transport systems is used as a reference. The formula is opportunely adapted to take into account weekly, rather than daily, frequencies. This is done by changing the value of the numerator, which represents the period taken into account in minutes. To account for the fact that long-distance services rarely depart or arrive after 00:00 in the evening and before 06:00 in the morning 18 hour operational days are considered rather than complete 24 hour days. Thus, the period is obtained by multiplying the number of daily operational hours (18), by the number of days in a week (7) and by the minutes per hour (60), obtaining the formulation illustrated by equation 3.12.

$$tt_{ij}^{trf} = \sum_{k \in K} \frac{T}{f^{inc} + f^{out}} \tag{3.12}$$

where,
$k \in K$: set of transfers on the shortest path between i and j.
T : constant representing the period, in this case 7560.
f^{inc} : frequencies on the incoming link.
f^{out} : frequencies on the outgoing link.

It seems important to point out that this formulation is usually employed in the Public Transport domain to compute the waiting time at stops between services scheduled at regular intervals with multiple daily frequencies. Consequently, there are some important implications to account for when applying it to the case of long-distance services. In fact, by extending the period and by reducing the frequencies the risk is to get to unrealistically high transfer times. In fact, long-distance transport services are often scheduled to allow changes and to optimise transfer time even with very low frequencies (e.g. hub-and-spoke model in the aviation sector). Thus, computing transfer times using the aforementioned formula could result in consistent biases. Let's assume, for instance, that going from A to B the shortest path implies using two services with a daily frequency and transferring from the former to the latter at C. Using the aforementioned formula the waiting time at C would be of 540 minutes or 9 hours. However, in reality the transfer time might as well be of only 30 minutes, in case the two services are planned to allow smooth and quick on-going trips. It is possible to conclude that this formula does not always guarantee an accurate estimation of transfer times. Nonetheless, it was decided to employ it as the indicator TI_{ij} does not aim to precisely measure the travel time of the connection, but rather to give an indication on its travel impedance. And it is assumed that lower frequencies tend to negatively impact the ease of transferring between consecutive services even when transfer times are optimised (e.g. in case of cancellations or delays).

To each sub-component of both connection intensity and travel inconvenience is associated a specific coefficient θ, which represents the degree of influence that each element has on passengers' behaviour for long-distance travel. The parameters considered are not mode specific as in-vehicle time modal sensitivity, as defined in Section 3.1.6, is already taken into account in the data set. Furthermore, differences across modes for the remaining parameters (i.e. out-of-vehicle time, transfer time and frequency) are not deemed as marked and are thus neglected. The chosen parameters are derived from ratios between model parameters provided across the literature. An overview of the studies considered for the parameter selection is provided in Table 3.7. To guarantee good compatibility between the different parameters it was decided to select them, as far as possible, from a single body of research. Thus, the study from Moeckel et al. (2015) has been employed as a base to set the parameters for in-vehicle time, out-of-vehicle time and transfer time. What is particularly interesting is that the study, rather than estimating econometrically the parameters for a R3 Logit long-distance mode choice model, derives them heuristically using a solid four-fold methodology. The use of this method is especially interesting considering that data for long-distance travel is generally scarce, as already mentioned in Section 3.2. This scarcity might implicate a limited accuracy of econometric approaches, which is supported by the fact that ratios between econometrically estimated parameters appear to widely diverge across the literature, as demonstrated by a thorough review. In terms of limitations it seems important to highlight that the parameters employed by Moeckel et al. (2015) are specifically derived for a case study in the United States (i.e. North Carolina), context which clearly has different characteristics compared to the European market. However, the parameters, following the ratios generally employed in the Public Transport domain, appear to be sensible also for the European case. Furthermore, it is important to highlight that due to the aforementioned limitations in the data availability a comprehensive estimation of long-distance mode choice model parameters considering both rail and air in Europe was not easily retrievable across the literature.

Table 3.7: Parameters used across the literature for long-distance mode choice

Parameter /Study	Bhat, 1995	Koppelman and Wen, 2000	Wen and Koppelman, 2001	Moeckel et al., 2015
In-vehicle time	-0.011	-0.0076	-0.0031	-0.023
Out-of-vehicle time	-0.0362	-0.0321	-0.011	-0.046
Transfer time	-	-	-	-0.1
Frequency	0.0741	0.0651	0.0288	-
Ratio in-vehicle time	1	1	1	1
Ratio out-of-vehicle time	3.291	4.224	3.548	2
Ratio transfer time	-	-	-	4.348
Ratio frequency	6.736	8.566	9.290	-

The study from Moeckel et al. (2015), however, does not take into consideration any parameter for frequency. Thus, in this case three other papers, estimating parameters for the same case study (i.e. inter-city travel mode choice in the Toronto-Montreal corridor) using different choice models, are analysed and compared. The study from Bhat (1995) employs a heteroscedastic extreme value model (i.e. random utility model with independent, but non-identical error terms distributed with a type I extreme value distribution), comparing its results the the ones obtained using a simple Multinomial Logit (MNL) model. The results of the heteroscedastic model reject the MNL formulation of mode choice suggesting

Table 3.8: Overview of the θ coefficients' value per component

Component	In-vehicle time	Out-of-vehicle time	Transfer time	Frequency
Parameter	θ_{in-veh}	$\theta_{out-veh}$	θ_{trf}	θ_f
Value	1	2	4	9

that the former is likely to be superior to the latter model. However, the authors highlight that Nested Logit (NL) model may be superior when identical-nonindependent components are randomly distributed. In this regard, Koppelman and Wen (2000) employs a paired combinatorial logit (PCL), which relaxing the restriction related to covariances between each pair of alternatives, allows to estimate the differential competitive relationships between them. The application of this model to the aforementioned case study demonstrates that the PCL model feature higher likelihood, significantly rejecting both MNL and NL models. Finally, Wen and Koppelman (2001) employs a Generalized Nested Logit (GNL) model, proving the statistical superiority of this model to the aforementioned models. In particular, the authors underline the consistent differences between parameter estimation results across different modes, demonstrating that other than data scarcity also the specific choice model selection has a consistent impact on the validity of the parameters. For this reason the frequency parameter is set based on the latter study. In conclusion, θ_{in-veh} is set to 1, $\theta_{out-veh}$ is set to 2 and θ_{trf} is set to 4 following Moeckel et al. (2015), whereas θ_f is set to 9 following Wen and Koppelman (2001), as per Table 3.8.

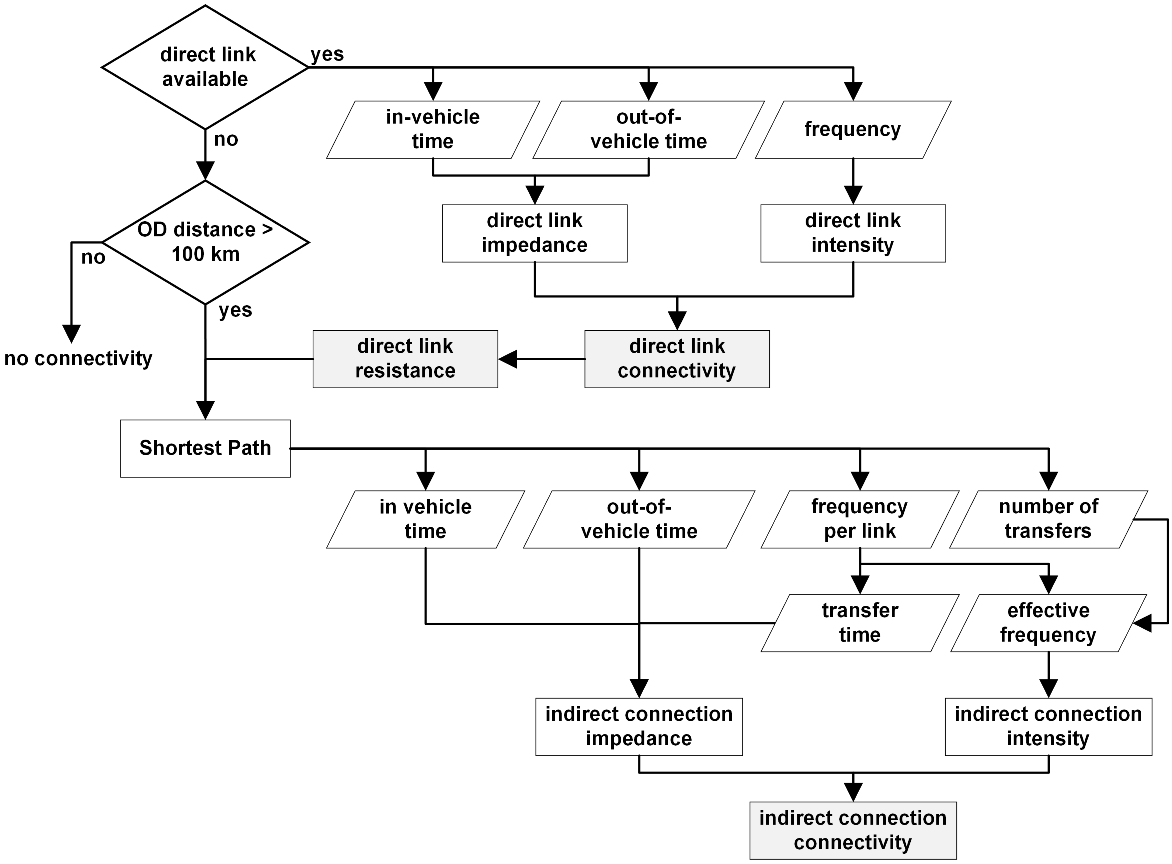


Figure 3.6: Link Connectivity Overview

Finally, it seems important to highlight that link connectivity is going to be assessed both at the local and global level, following the definitions provided in Section 2.2.5. In particular, at the local level, connectivity is measured by considering direct links only, whilst indirect connections are included at the global level. The methodology employed to compute link connectivity is summarised and schematised in Figure 3.6, highlighting the distinct procedures employed for direct and indirect connections. In order to compute the connectivity of all possible connection in the network (between all OD pairs), including all those not directly connected, the shortest paths per mode are computed using the Dijkstra method on networks with link resistance, as per formula 3.9, employed as weight. This method of selecting the shortest paths assumes that, when direct services between the desired origin and destination are not available, people will choose the route with the shortest perceived travel time, taking into account also the availability of services in terms of weekly frequencies. This follows the idea that passengers, generally, choose the route to reach the desired destination based on the total perceived travel costs of that connection (Zhu et al., 2018). This, for rail and air traffic, usually includes transport fare, perceived travel time and other parameters (e.g. transport emissions, frequency of service), which generally vary from the specific context of the study. In this case, other than the total perceived travel time, frequency is employed, as it appears to be an important parameter that passengers take into account when making decisions regarding route choice in long-distance travel. In particular the direct impact of frequency, related to the wider choice of departure/arrival services in terms of departure and arrival time is modelled through the connection intensity component. Furthermore, the indirect impact of frequencies, due to its influence on transfer times, is considered in the travel impedance component. Finally, also the number of transfers is included within the connection intensity component, whilst transport fares and emissions are excluded. The assumptions related to the topics discussed in this section are summarised in the following points:

- It is assumed that people when choosing the route to reach the desired destination, in the absence of a direct alternative, take into account total perceived travel time, number of weekly frequencies and number of transfers, disregarding other important characteristics such as travel fares or environmental impact of the trip.
- In relation to the parameter employed to weigh connectivity it is assumed that the passenger preferences do not vary depending on personal characteristics and on the specific transport mode, with the exception of in-vehicle time.
- In computing transfer times different assumptions are made in terms of the statistical distributions of arrival and departure times and their coordination, as thoroughly explained by A. K. Sen and Morlok (1976).
- It is assumed that the operational day for international transport service consists of 18 hour.

In conclusion, the outputs of the reaggregation process are employed as inputs of this process leading to the following outputs:

- Direct Link Connectivity
- Direct Link Resistance
- Indirect Connection Connectivity

3.3.2. Network Structure Topology Indicators

As already anticipated in Section 2.3, Network Science provides a wide range of indicators to assess the structure and topological properties of networks. In particular, literature widely agrees that the main network characteristics in terms of structure and topological properties can be inferred from its binary adjacency matrix, thus considering the unweighted version of the network (Paleari et al., 2010). The network structure indices employed within this research are briefly introduced in the following paragraphs.

Density

The density of an undirected network G is computed using Formula 3.13, where N is the number of nodes and L is the number of links in G. Thus, this metric represents the ratio between the number of links in a graph G and the maximum number of links that the graph can contain, ranging from a value of 0 for graphs without links to a value of 1 for complete graphs.

$$d = \frac{2L}{N(N-1)} \tag{3.13}$$

Degree

The degree of a network is defined as the number of links that connect a node to all the other nodes (Barabási & Pósfai, 2016). Within this study, considering the characteristics of the networks analysed, the degree of a node measures the number of direct connection that are available in a certain urban area. Using the a binary adjacency matrix A_{ij}, the degree can be computed as:

$$k_i = \sum_{i \in N} A_{ij} \tag{3.14}$$

Furthermore, at the network level the average degree $\langle k \rangle$ can be computed, considering in an undirected network, through the following formula:

$$\langle k \rangle = \frac{1}{N} \sum_{i \in N} k_i \tag{3.15}$$

In this study, the average degree $\langle k \rangle$ represents the average number of urban areas directly connected without requiring a transfer. Degree, however, is particularly important in network science because of its role in the calculation of network properties (Barabási & Pósfai, 2016). In fact, the scale free networks are generally identified through their degree distribution p_k, which represents the probability that a randomly selected node has degree k and is measured using the following formula:

$$p_k = \frac{N_k}{N} \tag{3.16}$$

where N_k is the number of nodes with degree k. Based on the distribution of the degree, it is possible to identify the type of network. Random networks are characterised by Poisson distributions whilst the degree of scale-free networks follows a power-law distribution. The degree distribution of networks tending to a regular configuration, on the other hand, is generally characterised by a bell curve as most nodes share the same number of links (degree).

Average path length

The average path length $\langle l \rangle$, also defined by Watts and Strogatz (1998) as characteristic path length, measures the average number of links contained in the shortest paths between all the possible node-pairs in the network, and can be computed using the following formulation:

$$\langle l \rangle = \frac{1}{N(N-1)} \sum_{i \neq j} spl_{ij} \tag{3.17}$$

where spl_{ij} represent the shortest path length between node i and j. Generally, shortest paths are computed on the binary matrix. In this case, however, the length of the shortest path computed using travel resistance weighted networks is considered instead, as these paths are deemed the most likely to be chosen by passengers in reality. Thus, the aforementioned formula is adapted to also take into account all the OD pairs that are excluded from the analysis, whose shortest path length is set to zero (i.e. OD distance < 100 km). The average shortest path length is a measure of the efficiency of the network and together with the clustering coefficient is particularly useful to detect small-world networks.

Clustering coefficient

The clustering coefficient C_i of a node i represents the share of actual links E_i connecting node i and all its neighbours (i.e. nodes directly connected by a link) out of the maximum number of possible links between them (Watts & Strogatz, 1998). Given that, to exclude self loops, the maximum number of possible links is equivalent to the product of the number of neighbours (i.e. node degree) by the number of neighbours minus one the clustering coefficient can be formulated as follows:

$$C_i = \frac{E_i}{k_i(k_i - 1)} \tag{3.18}$$

Larger clustering coefficient values imply that the node has a tighter connection systems with its direct neighbours. The values range from 0 for a node with only one connection, to 1 for a fully connected node. At the network level, the clustering coefficient is measured by the average node clustering, which is measured using the following formulation:

$$\langle C \rangle = \frac{1}{n} \sum_{i \in N} C_i \tag{3.19}$$

High values of $\langle C \rangle$ mean that node are on average closely connected to their neighbours, suggesting that the network is compact and well-connected. Moreover, implying that it is likely that nodes are well connected within a few transfers, this indicator in transport networks reflects the intensity of interconnectivity of the system (J. Wang et al., 2011). As previously mentioned, this measure is used by (Watts & Strogatz, 1998) to define small world networks. A network is, in fact, considered to be an empirical example of the small-world phenomenon when it features both a low average path length (i.e. similar to the one of a random graph) and a high average node clustering (i.e. larger than a random graph).

3.3.3. Node Connectivity

An extensive review of the literature has shown that a range of indicators has been employed to compute node connectivity. In this study, node connectivity is measured using three principle metrics, hub potential, node connectivity and connection directness. Each of these metrics aims to capture a specific dimension of node connectivity, and is computed both at the local and global level, considering only direct connections for the former and the entire network in the latter case. Furthermore, each indicator is computed and analysed for both rail and air. It is worth noting that the values from all the indices have different magnitude ranges and are consequently not directly comparable.

Hub Potential

Hub potential represents the importance of a node in terms of its role as transfer connectivity hub. To measure this a the local level the hub potential indicator proposed by Dennis (1998) for the aviation field is employed. In particular the hub potential is measured as a product of all incoming frequencies by all outgoing frequencies. Using the tools provided by Network Science, this can be translated into considering the squared node strength in the frequency weighted P-Space undirected network. Where node strength represents the node degree in weighted networks (W. Wang et al., 2020). In particular, the local hub potential hp_i^l formula is given by:

$$hp_i^l = (\sum_{j \in N} f_{ij})^2 \tag{3.20}$$

The number of total frequencies is squared to account for both inbound and outbound services, considering the use of symmetric matrices/undirected networks which impose the same number of incoming and outgoing services. This allows to capture the importance of a node in terms of direct traffic transiting it. In particular, the higher the local hub potential index the higher the number of total direct services available in a certain urban area. Furthermore, squaring the value allows to exponentially distribute the indicators, assigning considerably higher values to cities with higher numbers of frequencies. This follows the assumption that the attractiveness of a hub, relating to the total number of connection possibilities available, increases exponentially with the growth of the number of available frequencies. Mathematically, the total number of ordered combinations of incoming-outgoing services can, in fact, be measured using the combinatorics concept of permutations with repetition by multiplying all the possibilities of entering a node by all the possibilities of exiting it. It seems important to highlight that in reality only a small portion of these combinations represents viable connections due to the constraints imposed by temporal coordination and schedule synchronisation.

Global hub potential, on the other hand, aims to measure the centrality of a certain node within the wider continental context in terms of transfer attractiveness. It is measured using the betweennness centrality indicator in the P-Space undirected network with the reciprocal of frequency f_{ij}^{-1} employed as weight. In network science betweenness centrality is defined as the sum of the fraction of all shortest paths that pass through a certain node i. In particular, global hub potential hp_i^g is defined through formula 3.21.

$$hp_k^g = \sum_{i,j \in N} \frac{sp_{ij}^k}{sp_{ij}} \tag{3.21}$$

where, sp_{ij} represents the number of shortest paths between node i and j, and sp_{ij}^k the number of those shortest paths that passes through k. Betweenness centrality captures the extent to which a certain node lies in the shortest paths of the network. Thus, it allows to highlight the extent to which a certain urban area is used as a transfer hub for indirect connections. In particular, using the reverse of frequency as network weight allows to apply higher link resistances to corridors with lower frequencies. The shortest paths in such network will, thus, minimise the cumulative f_{ij}^{-1} of all travel legs rather than only minimising the number of transfers. The basic assumption that underlies this choice is that the attractiveness of indirect connections does not depend merely on the number of transfers but most importantly on the quality of these transfers. Quality, which, following the assumptions made in the previous sections, is estimated using the reciprocal of frequency. Formula 3.12, in fact, entails that there is an inversely proportional relationship between transfer time and service frequency. Thus, using this indicator implies an underlying assumption that passengers when choosing an indirect service aim to minimise their transfer time selecting the connections with the highest number of frequencies. The betweenness computed on such network reveals how many highest-frequency paths pass through a certain node, highlighting which hubs attract most transfer passengers. Higher values imply that the node is very attractive for transfers, allowing for a considerable number of connecting services, whilst lower values imply that the node is unlikely to offer a wide set of good transfer possibilities. In fact, lower In terms of the absolute value of the indicator, it is expected that the use of a weight will lead to a lower number of total shortest paths, and, consequently, to lower magnitudes of the betweenness centrality indicator compared to the case of an unweighted graph.

Node Connectivity

Node connectivity represents the node counterpart of link connectivity as defined in 3.3.1 and aims to provide a broad aggregated indicator of the connectivity of a node. Local node connectivity is computed as node strength in the P-Space undirected network with direct link connectivity as weight. This means that the connectivity of a node is equal to the sum of the connectivity of all the links adjacent to it. In other words, the connectivity of a node increases with the growth of direct connections available in terms of number and strength, expressed by the degree of link connectivity. In particular, local node connectivity nc_i^l is expressed mathematically as:

$$nc_i^l = \sum_{j \in N} LC_{ij} \tag{3.22}$$

At the global level node connectivity is measured using the matrix of connectivity between all OD-pairs, including direct links and indirect connections. In particular, global node connectivity nc_i^g represents the average connectivity for all direct and indirect available connections from node i, as per formula 3.23. It is important to highlight that for indirect connections shortest paths are computed using link resistance as weight, as explained in Section 3.3.1. Consequently, the metric proposed here appears similar to closeness centrality measured using link resistances as network weight. This measure, as defined by network science, captures the average shortest distance to any other node in the network. Which considering the specific weights would imply the average travel resistance on the shortest path to every node. As opposed to the indicator proposed, which considers the average connectivity on all these shortest paths. The decision to employ connectivity rather than resistance was mostly to be consistent with the definition of connectivity, however, it is important to highlight that a closeness centrality metric using link resistance as network weight is deemed to be equally effective in capturing node connectivity.

$$nc_i^g = \sum_{j \in N} \frac{LC_{ij}}{N_i} \tag{3.23}$$

where, N_i represents the total number of nodes connected with node i, and is computed by subtracting the number of connections with no connectivity with i (i.e. OD distance <= 100 km) to the total number of nodes in the network N. This allows to capture the connectivity degree of a node within the entire network. Higher values of global node connectivity imply that the node is well connected to a large number of other nodes.

Connection Directness

Finally, connection directness represents the availability of direct connections from and to a certain node. Local connection directness is computed as the degree of the unweighted network. Degree centrality is defined as the number of direct links incident on a node and is often described throughout the literature as the conceptually simplest metric of Network Science (Barabási & Pósfai, 2016; Z. Zhang et al., 2015). Literature widely employs the degree to measure the connectivity of a node, as higher degrees imply that more urban areas or terminals are directly accessible (W. Wang et al., 2020). In particular, local connection directness cd_i^l is expressed mathematically as:

$$cd_i^l = \sum_{j \in N} a_{ij} \tag{3.24}$$

where, a_{ij} represent the connection in the binary adjacency matrix $A(N \times N)$, assuming a value of 1 if urban areas i and j are directly connected and a value of 0 when no direct connection is available.

Global connection directness represents the relative centrality of a node or, in other words, the importance of a node in relation to the importance of its neighbours. This metric is measured computing the eigenvector centrality of the binary adjacency matrix. Eigenvector centrality has been described as a natural extension of degree centrality, where neighbour nodes do not all share the same fixed value of 1 but have a variable value that depends on the relative importance of the node in terms of the number of its direct connections (i.e. degree). In particular, the core principle of this metric is that connections to important nodes (measured in terms of degree centrality) are more important than connections to unimportant nodes. This means that higher eigenvector centrality scores are associated to nodes that are directly connected to a high number of nodes which are themselves connected to a high number of nodes (second degree nodes), which are also connected to a high number of nodes (third degree nodes) and so on. The eigenvector centrality of a node is obtained using an algorithm that iteratively computes the importance of each node in the network based on the relative importance of the adjacent nodes. In the first iteration the relative importance is set to 1, obtaining the degree centrality of all nodes. The second iteration then employs the degree centrality of adjacent nodes to compute the second degree centrality, which is then employed to compute the third degree centrality. Thus, despite the importance importance of each node starts at the same level, nodes with more direct connections start gaining importance with the progression of the iterations. In particular, at every iteration the importance of the nodes propagates to all the neighbours changing the relative importance of all the nodes until the values stabilise. Mathematically the process of the first iteration involves multiplying the adjacency matrix for an all 1 vector of size N, which implies summing over the values of each row. The vector obtained through that process, as previously mentioned, represent the degree centrality of each node. These vectors, in some cases after being normalised, is thus multiplied again for the adjacency matrix obtaining the relative importance of each node considering second degree connections and the iteration progress until an equilibrium is reached. The mathematical formulation of the eigenvector is described by equation 3.25.

$$x_i = \frac{1}{\lambda} \sum_{j \in N} A_{ij} x_j \tag{3.25}$$

where, λ represents a constant. In conclusion, eigenvector centrality allows to capture both the number and the importance of direct connections. A high value of eigenvector centrality could consequently either imply that many direct connections are available, or that the connections lead to very important (i.e. connected) nodes. Thus, considering the case of transport networks, an urban area might feature high global connection directness cd_i^g either because it is directly connected to many nodes or because it is directly connected to a few important nodes trough which all the remaining nodes are easily accessible.

3.3.4. Network-wide Components' Significance

Network-wide significant components in the network are defined as the nodes and links that play an important role within the network. They are identified using betweenness centrality on the P-Space undirected network with the link resistance LR_{ij} employed as weight. In fact, betweenness centrality, taking into account the extent to which a node/link is traversed by the shortest paths in the network could also be read as a measure of node/link significance. Link resistance is used as a weight to capture the most attractive shortest paths overall, as it is assumed that people generally try to minimise travel impedance when choosing a link. Thus, these two indicators aim to capture the nodes and links that are most likely to be traversed by connecting services considering travel time and frequencies of the links. In this regard, higher values of link betweenness centrality imply that a service is more often employed by connecting passengers, whilst higher values of node betweenness centrality imply that an urban area is more often chosen as a transfer location. Network-wide node significance is expressed by formula 3.26 whereas link significance is defined using formula 3.27.

$$c_n = \sum_{i,j \in N} \frac{\sigma_{ij}^n}{\sigma_{ij}} \tag{3.26}$$

$$c_l = \sum_{i,j \in N} \frac{\sigma_{ij}^l}{\sigma_{ij}} \tag{3.27}$$

where, σ_{ij}^n represents the number of shortest paths between node i and j that passes through node n, and σ_{ij}^l the number of those shortest paths that passes through link l. Betweenness centrality has already been used to assess global hub connectivity, however, the two indicators, despite employing the same formulation, capture two distinct characteristics. The former, in fact, identifies the theoretical hub potential merely based on the quality of the transfers, whilst network-wide significance also takes into account travel time other than frequencies, allowing to capture a more realistic scenario of the actual traffic patterns.

3.3.5. OD Relation Analysis

The methodology employed to categorise direct links is schematised in Figure 3.7. A first distinction is made between links under a monopoly regime of either rail or air services and links under a competition regime, with both modes operating services on the link. Thus, a second distinction is made within the links under competition, between competitive links and substitution ready links. The former group identifies all those links where travelling by air or rail does not imply consistent differences for passengers, whilst the latter identifies all those link where rail could possibly substitute air services. It is worth noting that these two categories are not mutually exclusive, and that all the routes are checked for competitiveness and substituitability. The fundamental condition for competition is that the OD geodesic distance d_{ij} is lower than 1500km. This threshold is widely deemed to be the longest distance over which rail can compete with air, as discussed in Section 1.2 and argued by Givoni et al. (2012). The fundamental condition for substitution, on the other hand, is that the weekly frequency of air services f_{ij}^{air} is greater than 35. Givoni et al. (2012), in fact, finds that the main determinant of the potential for mode substitution is the level of demand rather than the distance. As described in Section 3.1.2, demand is not directly estimated within this study, consequently air frequencies are employed as a measure of the level of demand. In particular the threshold value is taken from the study of Givoni et al. (2012) which considers as high demand routes all the connections with more than five daily flights. To evaluate the performances of services on each link the differences in terms of travel time and weekly frequencies between air and rail are computed as $delta_{tt}$ and $delta_f$ respectively. To compute consistent deltas it is required that values for both links are non-zero. This is ensured by the filtering undergone during the first step, through which all links under a monopoly regime are filtered out. Thus, the thresholds

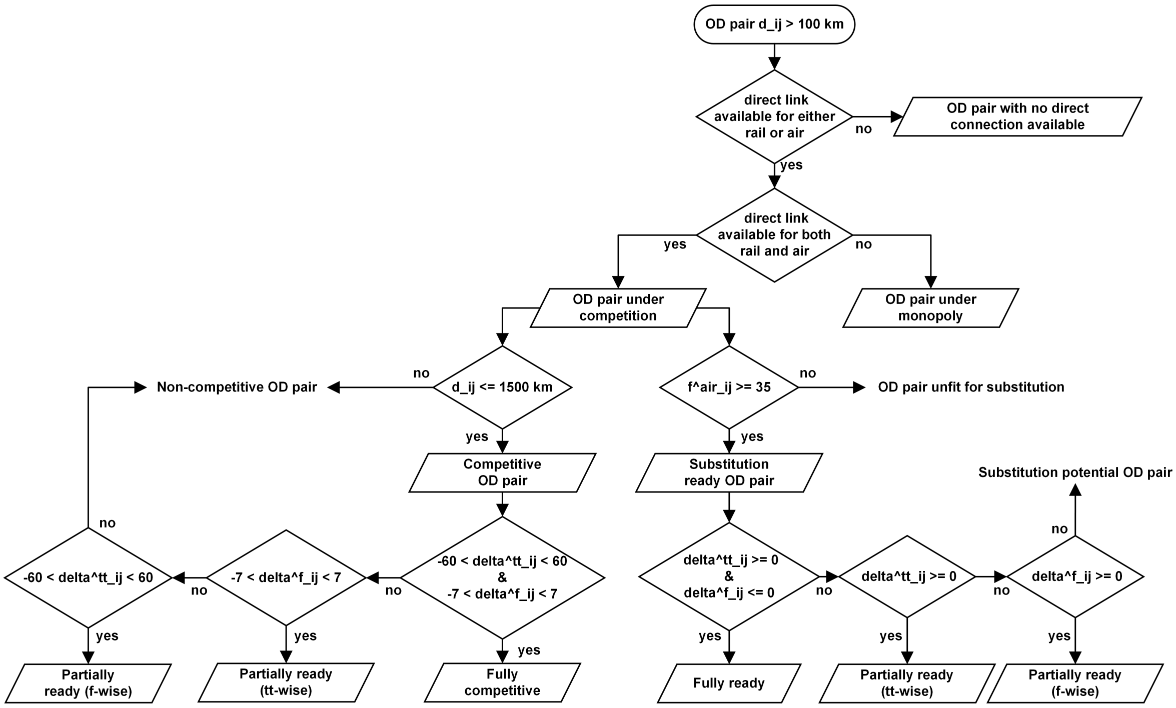


Figure 3.7: Link Categorisation Methodology Overview

for competitive links have been set to ±60 minutes in terms of travel time and to ±7 weekly frequencies (i.e. ±1 daily frequency). This implies that links are assumed to be competitive with each other when the travel time difference between the two modes is lower than an hour long and the difference in daily frequencies is 1 or less (considering a weekly schedule). These values are deemed sensible also in light of the frequency and travel time distributions provided in Section 5.2. Furthermore, to capture links that are currently ready for substitution $delta_{tt}$ is required to be greater or equal than zero (i.e. travel time for air greater or equal compared to rail) and $delta_f$ is required to be lower or equal than zero (i.e. number of frequencies for air is lower or equal compared to rail). Finally, for both competitive and substitution ready links a threefold categorisation is made, between links that fulfil both criteria, defined as fully competitive and fully substitution ready respectively, and links that fulfil only one of the criteria (i.e. either travel time or frequency), defined as partially competitive and partially substitution-ready respectively. The links under competition that do not fulfil any of these criteria are defined non-competitive links in the former case and links with substitution potential in the latter. Links with substitution potential represent all those OD pairs under a competition regime with a considerable demand (i.e. more that 35 weekly air services), where rail is currently not competitive either from a travel time or frequency perspective. Finally, all the air monopolies with frequent air service connections (i.e. more that 35 weekly air services) on the 100 - 1.500 km market segment that are not directly connected by rail services are categorised as OD pairs attractive for substitution.

4

Case Study

This chapter introduces the case study of the European long-distance rail and air transport networks, where the methodology described in Section 3 is applied. The model is practically developed based on the specific characteristics of the case study, to guarantee a realistic representation of reality. The process of network construction starts by defining the geographical scope of the problem, explained in Section 4.1. Having defined the broader geographical scope, a more specific selection of nodes to include in the network is developed. Thus, this chapter will further provide an in-depth explanation on the process that leads to the final selection of urban areas, and on the associated terminals. First of all, building on the urban areas selection criteria defined in Section 3.1.3, a fitting theoretical and statistical definition of urban areas is identified and given in Section 4.2. Secondly, the database selection process is going to be further addressed in Section 4.3. Thirdly, the actual process of selection and the final list of urban areas included in the study are thus explained in Section 4.4, and the methodology employed to define the geographical location of the nodes is explained in Section 4.5. Finally the association process between terminals and urban areas is given in Sections 4.6 and 4.7.

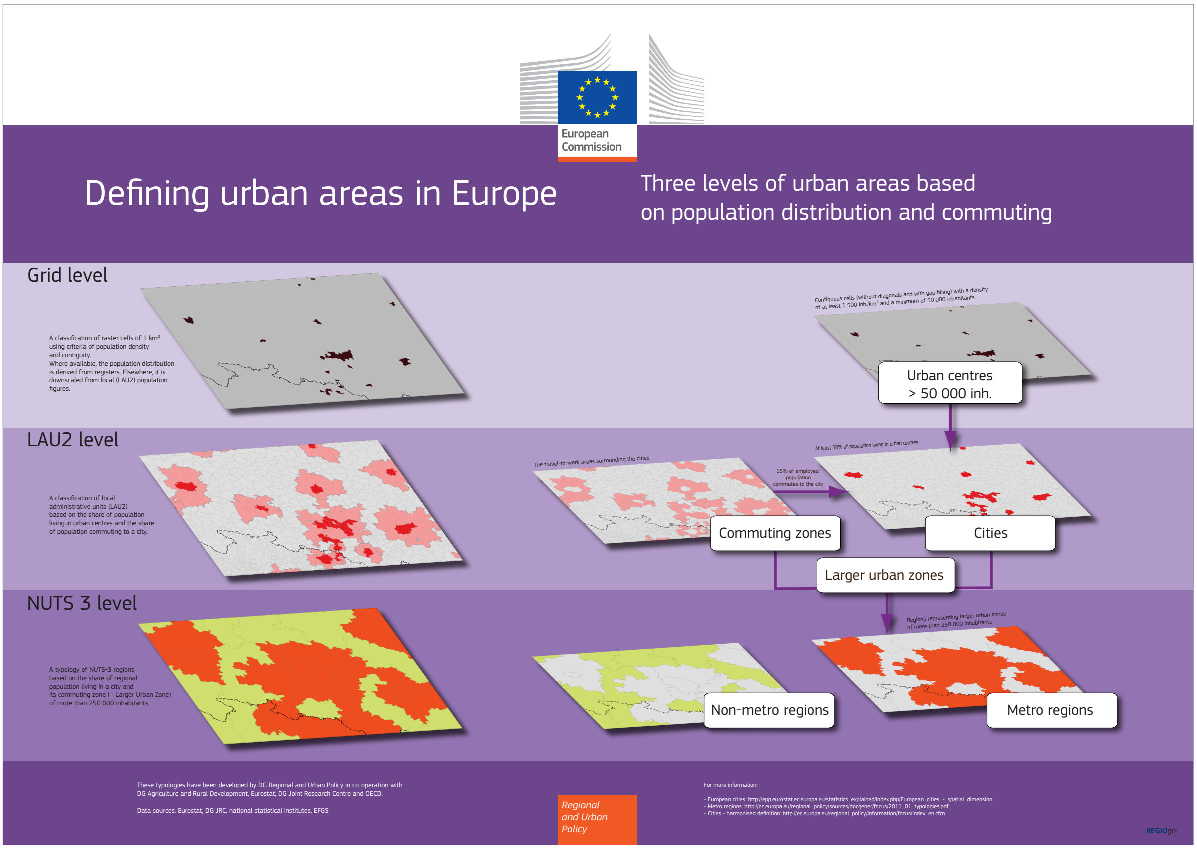
4.1. Geographical Scope

One of the practical aim of this research is to provide a disaggregated alternative to the connectivity index whose implementation feasibility is currently being assessed by the European Commission. Thus, the geographical scope of the research has been limited to the European Continent. To define it in more detail, a selection of countries has been created including the 27 countries part of the European Union. Starting from this first selection some countries have been excluded and some others have been included based on a set of criteria. A first criteria limits the scope to the European continent excluding all the overseas territories such as French Guyana (France) and La Réunion (France). A second criteria relates to the requirement that all countries considered are to be directly connected to the European international rail network. During this second step three insular countries, namely Ireland, Cyprus and Malta have been excluded. Furthermore, six islands and archipelagos, namely Madeira (Portugal), Azores (Portugal), Canaries Islands (Spain), Balearic Islands (Spain), Sardinia (Italy) and Corsica (France), other than all other minor islands, have been excluded from the scope. In fact, despite their undisputed touristic attractiveness, maintaining the comparability between the rail and air networks was deemed more important. Consequently, in a third phase, a number of European countries not part of the EU but still connected to the TEN-T European rail networks has been added, to include the UK, Ukraine, Belarus, Moldova, Serbia, Bosnia-Herzegovina, Albania, North Macedonia, Montenegro and Kosovo.

4.2. Urban Areas Definition

Once the selection criteria have been defined, in order to retrieve consistent, comprehensive and comparable set of data in terms of aggregation level, a more concrete and specific description to the "urban area" notion was researched. Dijkstra et al. (2019) highlight how comparing cities on an international level can be extremely challenging, as definitions of cities widely vary across different countries, even within the EU. Whilst it might be argued that the concept of urban centre as a node is rather clear, it is much more problematic to establish the actual extent of the urban area. In this regard, Weeks (2010) states that many definitions of Urban Areas and of what urbanness actually means and represents are to be found across the vast literature on the topic, suggesting that the matter is widely debated among researchers and that a common definition is still missing. Furthermore, Dijkstra et al. (2019) underlines that the delineation of urban areas is often based upon administrative or legal boundaries that do not necessarily reflect the functional and economic extent of the metropolitan area. This is reflected by the nomenclature of statistic territorial units for urban areas, scattered across different institutions and in some cases even within the same body. In this regard, Dijkstra et al. (2019) states that the extent of local administrative and legal units, often employed to identify city boundaries, can be utterly different across countries, urging caution when comparing city data coming from different countries. The different criteria used to define what is to be considered an urban area lead to considerable bias risks. Thus, it was decided to choose a common definition of urban areas in order to retrieve comparable datasets and avoid biases. Urban areas as specified within this research reflect the definition of metropolitan areas given by Moreno-Monroy et al. (2021), who describe them as "densely inhabited urban centres and their surrounding and interconnected lower-density areas".

Given the pan-european international scope the two main institutions taken into account are the European Union (Eurostat) and the OECD (OECD.Stat). These two institutions have jointly developed a methodology to consistently define Functional Urban Areas (FUAs) across different countries aiming to create an harmonised definition of cities and of their influence areas. However, whilst the OECD focuses mostly on FUAs, the European Union, features a wider range of definitions for cities and urban areas based on population grids, Local Administrative Units (LAUs) and NUTS regions. In the Urban Audit, Eurostat identifies three main levels of territorial units for urban areas: Cities, Functional Urban Areas and Metropolitan Regions. First of all, cities are defined as local administrative units (LAUs) where "a majority of the population lives in an urban centre of at least 50.000 inhabitants" (European Commission & Eurostat, 2019). Secondly, Larger Urban Zones (LUZs) or FUAs consists of a densely inhabited city and of its commuting zone, with the latter being defined as the surrounding areas where at least 15% of the employed residents works in the main city (European Commission & Eurostat, 2019). Finally, Metropolitan regions are defined as agglomerations of at least 250.000 inhabitants based on NUTS 3 regions or a combination of NUTS 3 regions identified using Functional Urban Areas (FUAs) (Eurostat, 2022b). This differs from the OECD nomenclature which defines metropolitan areas as estimated Functional Urban Areas (eFUAs), thus aggregations of grid cells not adapted to local administrative units or statistical territorial units (OECD, 2022). An overview of the different definitions from the EU



Defining urban areas in Europe

Three levels of urban areas based on population distribution and commuting

A classification of raster cells of 1 km² using criteria of population density and contiguity. Where available, the population distribution is derived from registers. Elsewhere, it is downscaled from local (LAU2) population figures.

Contiguous cells (without diagonals and with gap filling) with a density of at least 1 500 inh./km² and a minimum of 50 000 inhabitants

A classification of local administrative units (LAU2) based on the share of population living in urban centres and the share of population commuting to a city.

The travel-to-work areas surrounding the cities

15% of employed population commutes to the city

At least 50% of population living in urban centres

A typology of NUTS-3 regions based on the share of regional population living in a city and its commuting zone (= Larger Urban Zone) of more than 250 000 inhabitants.

Regions representing larger urban zones of more than 250 000 inhabitants

These typologies have been developed by DG Regional and Urban Policy in co-operation with DG Agriculture and Rural Development, Eurostat, DG Joint Research Centre and OECD.

Data sources: Eurostat, DG JRC, national statistical institutes, EFGS

For more information:

- European cities: http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/European_cities_-_spatial_dimension
- Metro regions: http://ec.europa.eu/regional_policy/sources/docgener/focus/2011_01_typologies.pdf
- Cities - harmonised definition: http://ec.europa.eu/regional_policy/information/focus/index_en.cfm

Figure 4.1: Urban Areas Definition in Europe (Source: European Commission and Eurostat, 2012)

and of the process employed to derive them is provided in Figure 4.1.

The statistical territorial units that most reflected the definition of urban areas given in the following paragraphs are the EU's LUZs/FUAs and the OECD's metropolitan areas/eFUAs. Furthermore, this aggregation level was deemed the most fit to capture the size of the problem analysed in this research. On one hand, cities were excluded for two main reasons. First and foremost, because they are related to administrative boundaries (LUAs) leading to the risk of important biases in terms of comparability across different countries. Secondly, because cities were considered to be excessively disaggregated. The limited extent in many cases (especially for larger metropolitan areas) is not representative of the actual catchment areas of international transport hubs such as stations and airports. In fact, with longer travel times the access and egress times that passengers are willing to accept also increase. Furthermore, the numerous number of entries makes the data-handling process more complex and probably leads to the need of a considerable number of corrections related to the aggregation of cities located in close proximity. On the other hand, metropolitan areas, as defined by the EU, were excluded because as for cities they are believed to have excessively tight bonds with administrative boundaries (NUTS 3 regions). Consequently their boundaries might include other than urban areas also rural areas located in the proximity. This might result in skewed and incomparable values especially in terms of indicators such as GDP and population. In fact, both greatly vary between rural and urban areas, meaning that depending on the percentage of urban and rural area included withing the borders of the urban areas the results may greatly differ.

4.3. Database Selection

Having defined the statistical territorial units to be employed, as illustrated in Section 4.2, the Eurostat and OECD.stats FUAs databases have been researched to operationalise the criteria defined in Section 3.1.3. The OECD.Stat database has been preferred to the Eurostat one because, even though the number of datasets are more limited, it is deemed more extensive in terms of reliability, completeness and number of entries. Especially considering that it includes also countries outside of the EU, such as Switzerland, the UK and Norway, which are not always portrayed in Eurostat datasets. The data related to population and GDP was, thus, retrieved on this level. However, the OECD.Stat database lacked elements to cover the other criteria and, consequently, the Eurostat one has been employed in this regard. From this database the total number of jobs (labour market attractiveness) and of EU foreigners (intra-EU migrations) have been retrieved. Finally, the number of bed-places in tourist accommodation establishments (touristic attractiveness) has been retrieved from an Eurostat dataset at the Cities level. The data was cleaned, ordered and adapted on the geographical scope defined in Section 4.1. Thus, all the urban areas located on islands were excluded, with the exception of Sicily and Great Britain as they currently feature stable train connections to the continent.

4.4. Urban Areas Selection

Given that the data is retrieved from different databases and at different aggregation levels it is not deemed to be directly comparable. Thus, two primary demographic and economic indicators are employed to define a first selection of cities. The union of the first 100 cities in terms of population and GDP (both from OECD.Stat) is selected as the basic input, accounting for a total of 111 cities. Then, the first selection of cities is projected on a European map and compared with a map of the European airports with more than 1 million passengers, and with a map of the rail TEN-T network as proposed by the European Commission. Consequently, the urban areas are aggregated based on proximity (Rotterdam-The Hague, Liège-Maastricht-Aachen), redundant cities are removed (close to major cities, without an airport in the catchment area and outside of the TEN-T railway corridors) and a few other cities were added based on the other three indicators (touristic attractiveness, labour market attractiveness and intra-EU migrations). A further check for under-representation and over-representation within each country is done, leading to the inclusion of some capital cities and other important urban areas of countries not featured in the OECD databases, such as Ukraine, Belarus and the Balkan countries. Finally, cities not connected by rail services to any other major urban centres are excluded as isolated nodes would cause issues in analysing the network. Through this process a first selection of 137 urban areas is defined. This is further refined based on the availability of data for both rail and air, as the same set of nodes is considered in each network to ensure comparability. Due to the war in Ukraine during the summer of 2022 no flights are available between Belarus, Ukraine, Russia and Europe. This leads to the exclusion of the urban areas of Minsk in Belarus, Saint Petersburg in Russia and Kiev, Kharkiv, Odessa and Lviv in Ukraine. Consequently, also the urban area of Helsinki in Finland is excluded as, being connected to Europe by rail only through Russia, it would become an isolated node. Furthermore, the UIC merits database used to retrieve rail data does not include the stations of Riga (Lithuania), Tallin (Estonia) and Chisinau (Moldova), which are consequently excluded. The stations of Pristina (Kosovo) and Skopje (North Macedonia), despite being included in the in the list of UIC stations do not feature any available service for any of the other urban areas and are consequently excluded. Thus, the final selection of 125 urban areas is illustrated in Figure 4.2. The countries included in the geographical scope that were excluded due to the unavailability of data are highlighted in orange.

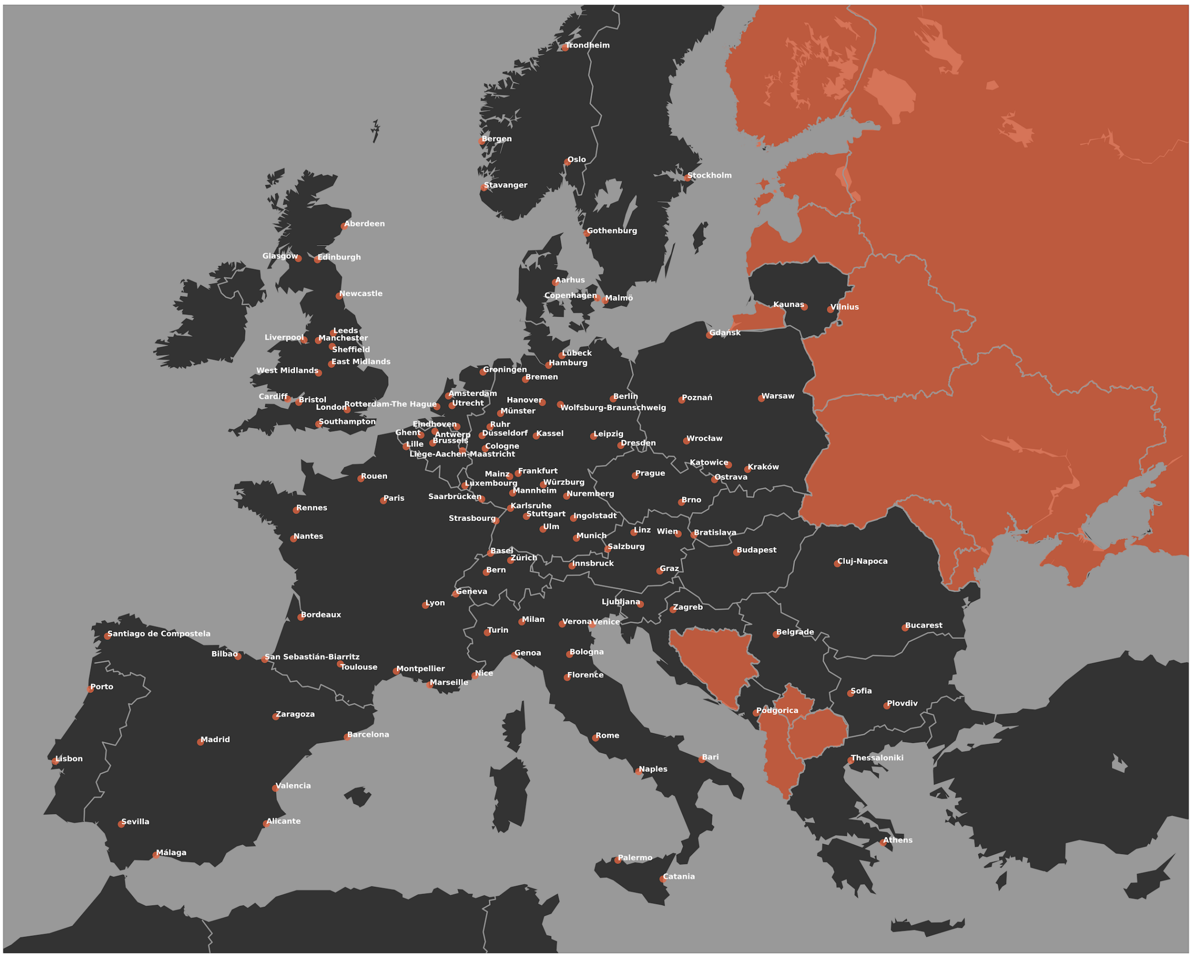


Figure 4.2: Selected Urban Areas and Excluded Countries

This methodology has been preferred to the use of linear aggregation methods, such as the summation of weighted and normalised individual indicators, because the latter would have required the use of weights. A thorough literature review has shown that there are not valid and commonly agreed methods to assign weights to the different indicators making this decision rather arbitrary.

4.5. Geographical location of the nodes

The geographical location of the selected urban areas is required in order to map them, to compute the distance between each OD pair, and to calculate catchment areas. To do so, the latitude λ and longitude ϕ are retrieved using the python client GeoPy and the Nominatim API geocoder for OpenStreetMap. A sample of the results has been checked manually to ensure that the correctness of the data, and an overall control is undergone by plotting all the geographical coordinates on a European map.

4.6. Airports Selection

First, to reduce computational efforts, the relevant commercial airports are filtered out based on their annual traffic, setting a minimum threshold of 1 million passengers per year. The traffic data employed is retrieved from the Eurostat database on air passenger transport measurement (Eurostat, 2022a). In particular, the data is taken from the year 2019 due to the impact of the Covid-19 pandemic on air traffic during 2020 and 2021. The filtered airports are then associated to each of the urban areas selected in the previous chapter making use of the catchment areas defined in Section 3.1.4. Thus, 90-min isochrones are defined using the TravelTime API, which allows to easily compute accurate catchment areas (Travel Time, 2022). In a few cases, no major airport appeared to be available within the city's catchment area (e.g. Zaragoza). In that case, commercial airports with passenger traffic under the threshold were included (e.g. Zaragoza Airport). Finally, for each airport-urban area pair the specific access/egress time are computed manually using Google Maps to take into account of the average traffic conditions. During this phase, driving times are compared to public transport accessibility, and access/egress time are set based on the fastest mode between the two.

4.7. Rail Stations Selection

Rail stations associated to each urban area have been manually selected based on the author's knowledge. First of all, for smaller urban areas only the major station, providing intercity and long-distance services, was selected. In larger urban areas (e.g. Wien) and in aggregated urban areas (e.g. Rotterdam - The Hague) multiple stations have been included to make sure to capture all the services available. In particular, in some urban areas (e.g. London and Paris) head stations are associated to specific lines, each handling the traffic going in a different direction. In that case all the major terminal stations are included. Finally, all the stations are manually matched to their UIC code, in order to allow the collection of the data, as described in Section 3.2.2.

5

Results and Discussion

In the following chapter the results obtained by applying the methodology described in Chapter 3 are provided and discussed. First of all, in Section 5.1 the network structure is analysed, the topological properties are identified and the results are compared across the two modes, providing an overview on the main patterns found at the network level. Then, the analysis dives into the analysis of the OD pair relationships, first analysing Link Connectivity in Section 5.2, and then in Section 5.3 categorising OD pairs into direct connections under monopoly and competition regime and further subdividing the latter into competitive and substitution ready. Section 5.4 adds to the analysis focusing on node connectivity, expressed in terms of hub potential, node directness and an aggregated measure of the aforementioned link connectivity metric. Next, the most significant components in the network are identified and analysed in Section 5.5. Finally, Section 5.6 concludes the chapter summarising the key findings.

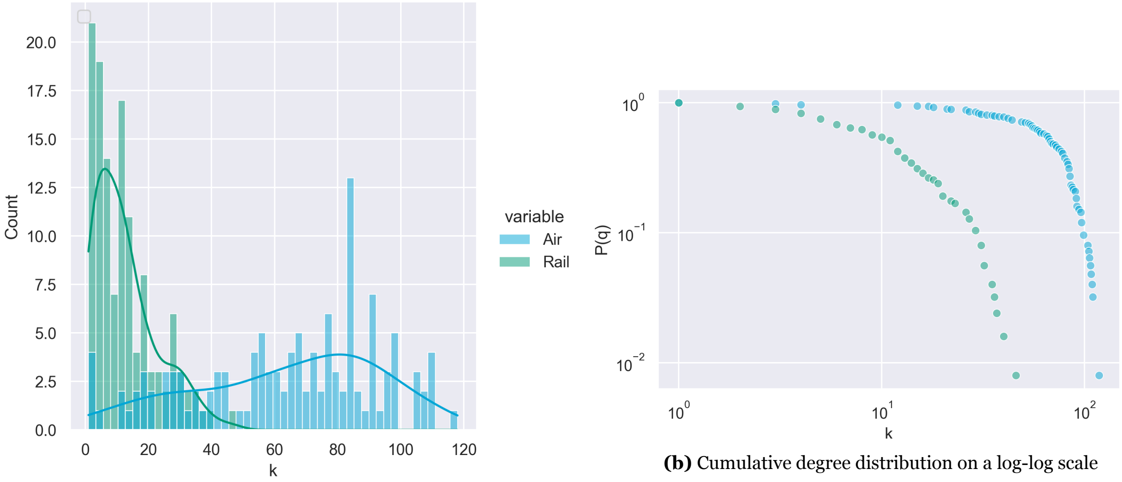
5.1. Network Structure and Topological Properties

First, a brief topological analysis of the rail and air unweighted networks (with the exception of shortest paths, which are computed on a link connectivity weighted network) is developed using the indicators defined in Section 3.3.2. To do so, some basic descriptive statistics are provided in Table 5.1. Comparing the rail and the air network it is possible to notice how the latter features a considerably higher number of links. Given that both networks share the same number of nodes, the density of the air network is necessarily higher than the density of the rail network. This implies that air currently dominates the market, at least in terms of the number of direct connections available. This aspect, is further underlined by a comparison of the average degrees, which represents the average number of neighbours per node, or the average number of urban areas reachable directly without transfer. In particular, the ratio between the indicators computed on the air and rail networks remains consistent at around 5 across the number of links l, the density d, and the average degree $\langle k \rangle$. Thus, in terms of the supply of unique routes between major urban areas across Europe air currently outperforms rail at around five to one. Furthermore, the average path length in the rail network is twice as long as the average shortest path in the air network. This implies that on average reaching a destination through rail services requires double the amount of transfers in comparison to flying. In terms of average node clustering the values are quite similar, with the air network being slightly more clustered. However, it must be considered that clustering depends on the density of the network. Consequently, in relation to that, the rail network appears to be extremely clustered. This is in line with the expectations, as generally neighbours in the rail network are geographically close and tend to have stable relationships, whilst this is less the case for the air network, which is less reliant on the geographical dimension. Finally, the network degree distribution of the rail and air networks is compared using a histogram with a kernel density estimation (kde) curve, as illustrated by Figure 5.1a. It is possible to notice how in the rail network most nodes have a rather low degree, with only a few high connected nodes. However, considering that the maximum degree is around 50 it also appears that no crucial hubs are present in the network, especially compared to the air network. This reflect the characteristics of the rail network and particularly its spatial constraints, which do not allow for high centralisation. On the other hand, the air network appears to be quite evenly distributed, with a peak around 80. This is rather surprising, as the air networks generally tend to allow for high degrees of centralisation, with many smaller airports and a few important hubs. However, in this case, it is important to remember the characteristics of the network as defined in Chapter 4. In fact, only 125 major urban areas are included, implying that the aforementioned smaller nodes might be present in the real network but have been excluded by the selection criteria. Furthermore, the traffic from all airports within the catchment area has been aggregated, further flattening the degree distribution curve.

Table 5.1: Basic descriptive statistics of the network structure and topology indicators of the European Rail and Air Networks

Network	Nodes N	Links L	Density d (%)	Average Degree $\langle k \rangle$	Average Path Length $\langle l \rangle$	Average Node Clustering $\langle C \rangle$
Air	125	3994	51.53	63.904	1.488	0.739
Rail	125	778	10.04	12.448	3.085	0.628
Random Air	125	4013	51.78	64.208	1.482	0.518
Random Rail	125	767	9.90	12.272	2.173	0.100

In order to evaluate whether the two networks feature small-world characteristics, two random graphs are generated using the Erdős-Rényi (ER) model, following the methodology employed by Humphries et al. (2006). In particular, to allow the comparability of the random networks $G(n, p)$ the number of nodes n is set to 125 and the probability p that two nodes are connected by a link is set to the density d of each specific network. Comparing each random network to the respective real network it is possible to notice that how air is generally more performing in terms of average path length, whilst rail appears to be more performing in terms of average node clustering. This represents the core characteristics of the two modes, air allowing quick connections on long distances, whilst rail creating compact clusters of directly connected urban areas. With regards to average path length, the air network appears to behave as a comparable random network, whilst the rail network features slightly higher shortest path on average. This is probably due to the spatio-temporal constraints of the rail system. The average node clustering of real networks is for both rail and air considerably higher than a comparable random network. The average node clustering of the rail network is particularly interesting featuring a sixfold figure compared to its random counterpart. After a comparison with the random network, the characteristics of the real networks appear to be in line with the expectations. In terms of topological properties, the three ratios $\gamma = \langle l \rangle / \langle l \rangle_{random}$, $\lambda = \langle C \rangle / \langle C \rangle_{random}$ and $S = \gamma / \lambda$ have been computed for both rail and air. In fact, Humphries et al. (2006) argues that, in order to meet the small-world criteria set by Watts and Strogatz (1998), the network should respect two criteria: $\lambda > 1$ and $S > 1$. Both the air and rail network respect the conditions set on the λ (λ^{air} = 1.43 and λ^{rail} = 6.28) and on S (S^{air} = 1.42 and S^{rail} = 4.42). Thus, it is possible to conclude that both the rail and air networks might be considered as small-world networks. However, it is recommended caution in using such definitions, as many argue that the number of real networks with small-world characteristics is largely overestimated (Telesford et al., 2011).



(b) Cumulative degree distribution on a log-log scale

(a) Degree distribution on a linear scale

Figure 5.1: Network degree distributions of the air and rail networks

As mentioned in Section 3.3.2 and previously in Section 2.3 the scale-free is evidenced by a power law distribution of the node degree. Power law distributions can be easily identified by plotting the data points on a log-log scale. In fact, if the data is power law distributed ($P(k) \sim k^{-\gamma}$) the resulting data points will follow a straight line with slope γ, as $log(P(k)) \sim -\gamma log(k)$. In particular, Barabási and Pósfai (2016) highlight that real networks rarely observe a degree distribution that follows a pure power law, often featuring low-degree saturation and high-degree cutoff. To best capture the power law distribution of the plotted data and to allow for a more accurate estimate of the degree exponent Barabási and Pósfai (2016) suggest the use of complementary cumulative degree distribution $P(q) = \sum_{k=q+1}^{\infty} P(k)$. The plotted cumulative degree distribution, illustrated on a log-log scale in Figure 5.1b, clearly show that both the air and rail network are unlikely to feature scale-free characteristics, as the distributions do not follow a straight line. This is particularly accentuated for the air network due to the flattened degree distribution, as shown in figure 5.1. Given the scope of this research the analysis won't dive deeper into the small-world and scale-free topological properties of the two networks. However, more thorough statistical tests are deemed necessary to assess the actual small-world properties of the networks beyond reasonable doubt. On the other hand, following the arguments made in this paragraph, it appears highly unlikely that a more in-depth analysis would prove the networks to be scale-free.

Table 5.2: Comparison of average path length and average node clustering across different studies

Case Study	European Rail Network		Chinese Rail Network		European Air Network		Chinese Air Network		
Study	Calzada-Infante et al., 2020	This Study	W. Li and Cai, 2007	W. Wang et al., 2020	Paleari et al., 2010	This Study	Paleari et al., 2010	J. Wang et al., 2011	Cai et al., 2012
$\langle l \rangle$	2.89	3.085	3.5	2.66	2.80	1.488	2.34	2.23	2.20
$\langle C \rangle$	0.77	0.628	0.835	0.49	0.38	0.739	0.49	0.69	0.79
SW/SF	SW	SW	SW & SF	SW	SW	SW	SW	SW	SW

To contextualise the results of the network analysis, the figures are compared with the outcomes of other studies in Table 5.2. The results of this study concerning the European Rail Network appear to be in line with the results from Calzada-Infante et al. (2020). Although, the higher average path length and lower clustering coefficient imply that the rail network analysed in this study is slightly less compact and connected. These patterns, however, are probably also influenced by the considerably higher number of nodes (412) and links (7732) analysed by Calzada-Infante et al. (2020). In comparison to the Chinese Rail Network, the rail network analysed in this study appears to lie between the results of W. Li and Cai (2007) and W. Wang et al. (2020). Also in this case the important differences in values between the two cases are probably imputable to the much larger amount of nodes and links analysed by the former authors. Taking into accounts the results from W. Wang et al. (2020), and considering the shorter average path length and the lower clustering coefficient, it is possible to argue that the Chinese rail network appears to offer more long-distance direct connections, whilst the European network seems to focus more on the medium range. This is in line with the actual supply of rail services across the two continents, and highlight the current fragmentation of the European rail market, as opposed to the Chinese one. The average path length of the air network computed within this study appears to be considerably shorter than the figures of comparable studies, whilst the clustering coefficient appears to be higher. Again, this is probably a consequence of the exclusion of smaller airports and the use of urban areas rather than terminals as nodes of the network. At the same time, it is possible to notice that across all the papers analysed there are important differences in the results, even within the same case study. Comparing rail with air networks, it is possible to notice that the former, in general, feature longer average shortest path and slightly higher clustering coefficient. This means that rail network tend to require more transfers on average to reach a destination, but provide a more compact network on the local level with more neighbour nodes being directly interconnected by a link. Finally, analysing the small-world and scale-free properties across the studies, it is interesting to notice how each network features the former topological characteristics but only the rail network analysed by W. Li and Cai (2007) shows scale-free behaviours. However, as previously mentioned, caution is suggested when talking about real scale-free and small-world networks without the support of proper statistical tests and thorough analyses. Many papers have, in fact, showed how real networks appear to be overly identified with these two concepts, as the network properties tend to be often simplistically analysed (Broido & Clauset, 2019; X. F. Wang & Chen, 2003).

Table 5.3: Distribution of air shortest paths by number of transfers

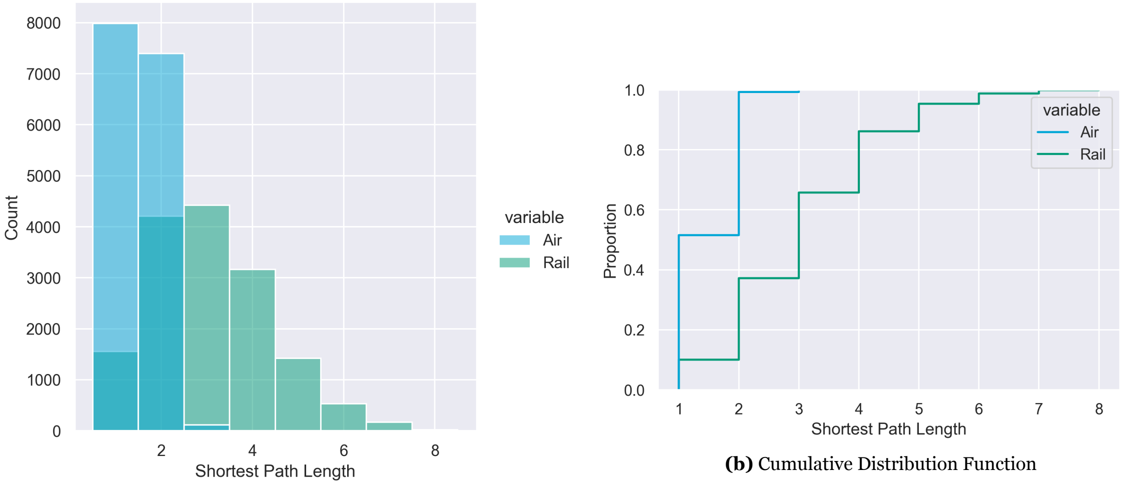
Shortest Path	Number of Paths	Percentage of total paths	Cumulative percentage of total paths	Number of Transfers
1	7988	51,92%	51,92%	0
2	7280	47,32%	99,25%	1
3	116	0,75%	100,00%	2

Interesting insight into the structure of the network is also provided by the distribution of the shortest path lengths. Table 5.3 and 5.4 illustrate the distribution and cumulative distribution of shortest paths in the air and rail network respectively. Furthermore, Figures 5.2a and 5.2b illustrating the distribution of the shortest path length and the cumulative distribution function respectively provide an overview on how the figures compare across the two modes. The network diameter is a measure often used to describe networks, representing the longest shortest path in the network. In this case, the diameter could also be seen as a measure of the maximum number of transfers needed to get from any node to any other node in the network. The diameter of the air network is 3, implying that every node in the

Table 5.4: Distribution of rail shortest paths by number of transfers

Shortest Path	Number of Paths	Percentage of total paths	Cumulative percentage of total paths	Number of Transfers
1	1454,00	9,45%	9,45%	0
2	4196,00	27,28%	36,73%	1
3	4422,00	28,74%	65,47%	2
4	3164,00	20,57%	86,04%	3
5	1422,00	9,24%	95,28%	4
6	530,00	3,45%	98,73%	5
7	172,00	1,12%	99,84%	6
8	24,00	0,16%	100,00%	7

network is accessible within a maximum of two transfers. Furthermore, an overwhelming majority of the OD pairs can be reached within one transfer, with less than one in a hundred connections requiring two transfers. In the case of rail, the diameter is 8 meaning that the maximum number of transfers required to move between any two nodes is 7. This value is considerably higher than in the case of air due to the characteristics of rail, which are spatially and temporally bounded by infrastructural and speed-related constraints. However, it is possible to notice that more than 95% of the OD pairs can be reached within 4 transfers. This suggests that there are some minor areas which are particularly poorly served. Considering that the rail system is bound to the geographical localisation of the nodes, it is expected that these areas are located in the external peripheries of the continent. On the other hand, more information is required to identify the location of the few poorly connected nodes within the air network. It is particularly interesting to notice that direct air services cover over half of the possible connections. This, highlighting the scale of the availability of direct connections in the network, shows the crucial role of air transport in the European long-distance market. Rail, on the other hand, features much inferior figures, offering direct connections only between less than 10% of the possible connections. However, it is worth noting that rail covers a considerable share of OD pairs within two transfers, serving almost 65% of the all the possible routes. This underlines the crucial role of transfers and schedule coordination in the rail network if this is ever to compete with air on a network-wide level. In this regard, Figures A.1 and A.2 portray the average number of transfers required to reach from each



(a) Distribution

(b) Cumulative Distribution Function

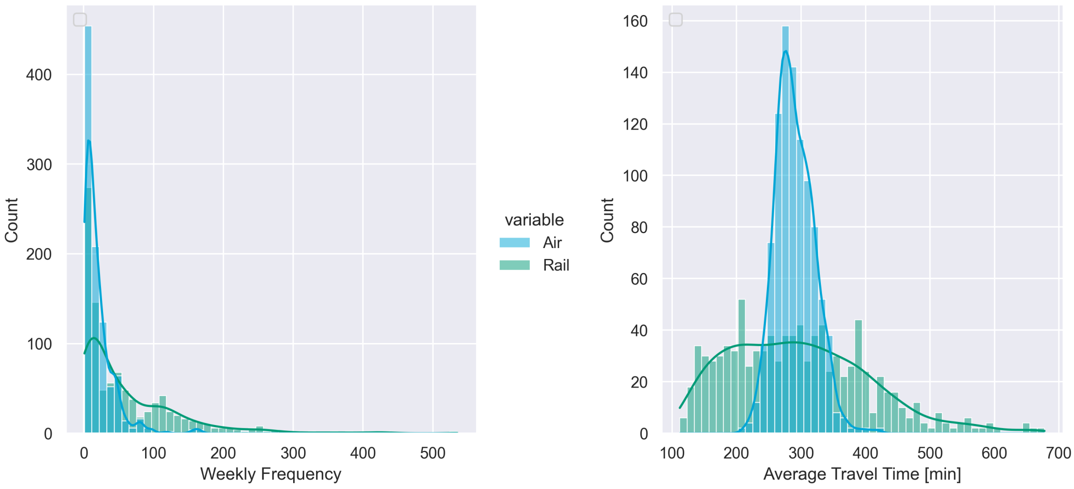
Figure 5.2: Distribution of Shortest Path Length within the rail and air networks

node any other node in the network, for the rail and air case respectively. It is worth noting that in both cases the average number of transfers is inversely related to the number of available direct connections, as illustrated by Figures 5.24 and 5.25. In particular, the average number of transfers within the rail network appears to be related to the geographical location of the nodes and especially to their centrality. In fact, central European nodes generally feature considerably lower average shortest paths compared to peripheral areas, such as the Iberian peninsula or Greece. This matter is discussed in more details in Section 5.4.2.

It is possible to conclude that constraining the number of nodes in a network and excluding smaller cities has an impact on the node structure. Furthermore, aggregating the data to urban areas rather than terminals also influences the results in terms of network structure. It is particularly interesting to note that the effects are more noticeable in the air network than in the rail network, probably because of the use of catchment areas. Thus, it seems important to note that when developing models employing urban areas rather than terminals as nodes it is crucial to make considerations based on these two elements, being aware on the impact that the design choices might have on the results. Furthermore, when analysing their outcomes it is fundamental to consider the influence of the aforementioned choices on the results.

5.2. Link Connectivity

This section dives into the link connectivity, measured using the methodology described in Section 3.3.1, analysing the characteristics of the data retrieved for direct connections and on the connectivity of all the links across the network. In particular, link connectivity consists of two main components, travel impedance, measured using travel time and connection intensity, measured using weekly frequencies. To explore and compare the characteristics of the two networks in terms of travel time and frequency, the distributions of direct travel time and direct frequency are plotted in Figure 5.3. In terms of frequency, the air network shows a considerably higher concentration of low frequency connections, with a maximum of 165.5 weekly frequency equivalents between London and Glasgow. Rail, on the other hand, has a much wider distribution of frequencies which peaks at 535 on the corridor between Wien



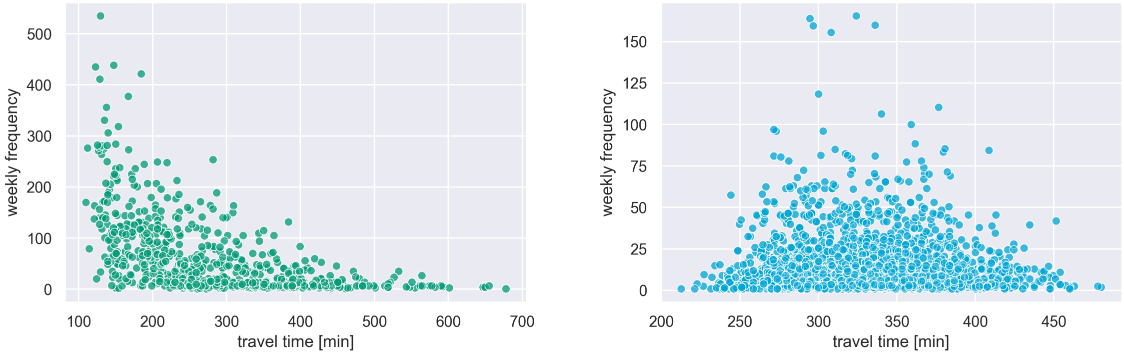
(a) Frequency

(b) Travel time

Figure 5.3: Network components distribution for direct links

and Linz. It is interesting to notice how for both the air and rail network the strongest connection in terms of frequency is a national route. The distributions of travel times follow a similar pattern. Air travel times are concentrated at around 300 min, with a minimum of 212,5 min between Geneva and Lyon and a maximum travel time of 480 min on the Lubeck-San Sebastián, Barcelona-Kaunas, Glasgow-Thessaloniki and Porto-Bucarest routes. On the other hand, rail travel times are more evenly distributed ranging from a minimum of 112 min between Brussels and Liège to a maximum of 677,5 min (more than 11 hours) between Hamburg and Salzburg. It is important to note that rail travel times are perceived door-to-door travel times, which are travel times weighted using the modal time sensitivities defined in Section 3.1.6. Thus, they are shorter than actual travel times, especially in the case of night trains. The trends and patterns identified within this paragraph are strongly influenced by the travel time composition of the two modes, as defined in section 3.1.5. Air total door-to-door travel time is mostly influenced by the fixed out-of vehicle times, as the high speeds allow for a low variability of travel time across different distances. On the other hand, rail service total travel time mostly depends on the in-vehicle time, which is highly variable considering its dependence on spatial distance due to the lower average speeds.

To explore the relationship between travel time and frequency, the two components of each direct link are plotted in Figures 5.4a and 5.4b. Some interesting patterns are noticeable comparing the case of the rail and air network. The former features high weekly frequencies (over 300) only for short travel times (under 200) min. The latter, on the other hand, features the highest frequencies for medium travel times, ranging between 275 and 375 minutes. This suggests that the current rail service supply tends to focus on OD pairs reachable within short travel times. In fact, with increasing travel times rail frequencies generally tend to decrease, in some cases linearly and in others exponentially. The overwhelming superiority of rail on these shorter routes is possibly due to inability of air to offer competitive travel times on such connections (i.e. travel time < 200 min), due to the considerable access/egress and waiting time that are required to use the mode. On the medium travel time range the frequencies of rail and air appear to be similar. Thus, the market of trips with perceived door-to-door travel time between 250 and 400 minutes (i.e. 4 and 6,5 hours circa respectively) appears to have comparable size for both rail and air. In this regard, it is worth noting that comparable frequencies within the European market generally imply that rail features higher service capacity in terms of offered seat compared to air, as discussed in Section 3.1.2. Looking at the distribution of travel times in the air network, it could



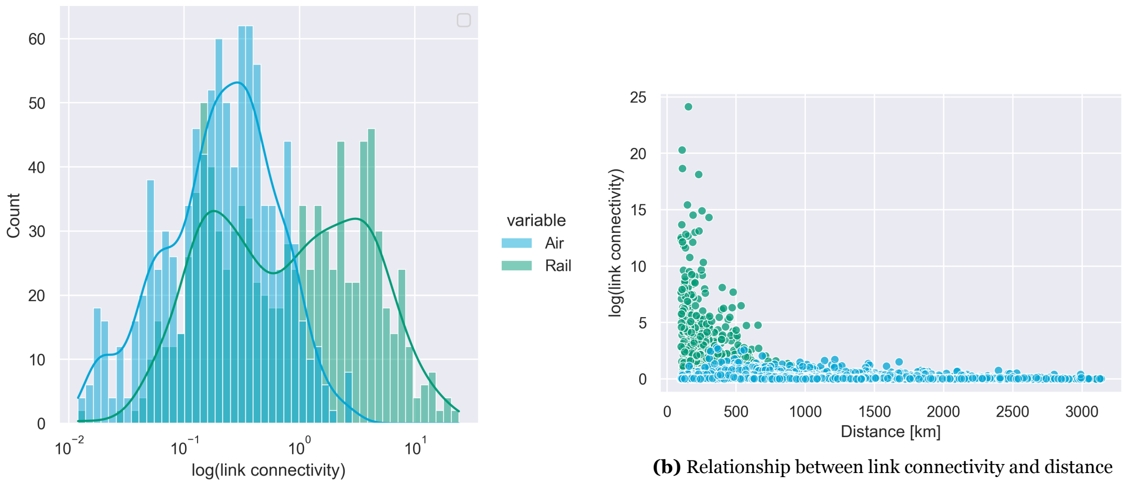
(a) Rail Network **(b)** Air Network

Figure 5.4: Relationship between travel time and frequency in direct links

be argued that in order to be really attractive long-distance rail transport within Europe should aim to offer perceived door-to-door travel times lower than 450 minutes (7,5 hours). Furthermore, it is worth noting that air does not provide any service with travel time lower than 200 minutes, being hardly able to offer any travel times lower than 250 minutes. Thus, using 250 minutes as a time threshold and considering the rail access/egress times and waiting time of 70 minutes and an in-vehicle time sensitivity of 0,75 (compared to 1 for air), as defined in Sections 3.1.5 and 3.1.6 respectively, it is possible to estimate the distance range where rail can compete with air. Offering comparable or better performances, rail appears to be able to compete with air, and possibly even substitute air routes on direct connections between OD pairs with physical distance until:

- 1200km with an average speed of 300km/h
- 1000km with an average speed of 250km/h
- 800km with an average speed of 200km/h

Finally, the link connectivity of all the possible connections (i.e. OD pairs with distance greater than 100km) is analysed. Link connectivity in the air network ranges from a minimum of 0,00037 between Ostrava and Southampton to a maximum of 2,76 between Amsterdam and London. On the other hand, in the rail network link connectivity varies from a minimum of 0,00014 on the Dresden-Athens corridor to a maximum of 24,13 on the Wien-Linz corridor. Thus, also in this case, the distribution within the rail network appear to be more widely distributed. In particular, looking at the histogram of link connectivity, provided in figure 5.5a, it is possible to notice how rail generally appears to have an edge over air in terms of connectivity. This, is due to the formulation of link connectivity, illustrated by formula 3.8, which places particular importance on high frequencies and short travel times. These, however, are both extremes where air is not particularly present. Figure 5.4, in fact, shows that the magnitude of the highest rail frequencies figures is more than three times over the one of the air counterparts. At the same time, travel times ranging between 100 and 200 are only available for rail. Thus, the important advantage of rail in terms of connectivity can be explained highlighting the large number of high-frequency low-travel time corridors, which thanks to these combination of characteristics greatly outperform all other links. In particular, it should be highlighted the prominent importance of extremely high frequencies in establishing connectivity, as the Wien-Linz corridor, featuring the highest



(a) Link connectivity distribution

(b) Relationship between link connectivity and distance

Figure 5.5: Link Connectivity Distribution and Relationship with OD distance

figure for frequency, is also the most connected link across both networks. To better understand the dominance of rail over air connectivity, the relationship between link connectivity and spatial distance is analysed in figure 5.5b. Link connectivity is generally evenly distributed for values between 0 and 2,5 (on a logarithmic scale) with the exclusion of an important cluster of outliers composed of short distance (i.e. <500km) rail connections. Whilst link connectivity within the air network appears to be independent of distance, connectivity in the rail network plummets for connections over 500 km. In fact, high values of link connectivity in the rail network always relate to the aforementioned short distances. This is probably due to the structural characteristics of rail networks, and to its infrastructural dimension. Direct services between OD pairs located further apart generally stop in urban areas on the path providing additional frequencies to the surrounding cities. Thus, it is reasonably foreseeable that the amount of frequencies for rail services exponentially grows for nodes with strategic infrastructural roles between major cities and in areas with higher densities of urban areas. It is possible to conclude that, despite the link connectivity index defined within this study provides an effective tool to measure and compare the degree of connectivity between specific links, there are some important factors to take into account when considering the entire network. This measure, in fact, appears to be indirectly dependant on spatial distances between OD pairs for rail. Thus, when comparing aggregated values and figures for the entire network across rail and air it is important to take into account the spatial dimension. For instance, including connections between 100 and 500 km appears to strongly skew the overall link connectivity distributions in favour of rail.

5.3. OD Relation Analysis

Having defined the connectivity of direct and indirect connections in the network, OD pairs directly connected by either rail or air (or both) are categorised using the definitions and the methodology described in Section 3.3.5. First of all, the influence of spatial distance on direct connections within the air and rail network is assessed in Table 5.5 using four distance thresholds. Whilst the previous chapter analysed and compared the quality of the connections across the two networks, this chapter aims to capture their quantity, building over the information provided in Section 5.1. The aforementioned table, in fact, highlights the number of supplied air and rail direct connections within each specific threshold, relating these figures to the total number of possible OD pair connections in terms of percentages. As already mentioned in the previous sections, due to the absence of spatio-temporal constraints the air network features a considerably higher number of direct connections compared to the rail network, and thus is a more densely connected network. However, when considering only OD pairs with spatial distances within a threshold of 500km this difference almost disappears, with the two networks featuring a similar number of routes. This implies that for geodesic distances up to 500 km the air-rail route supply shares are rather balanced. The 100-500 km market, in fact, appears to be, by far, the most important market for rail service supply, accounting for almost 75% of all the direct connections

Table 5.5: Influence of distance on the total number of direct connections

Distance Threshold (km)	**Possible OD pairs**		**Supplied Direct Connections Air**			**Supplied Direct Connections Rail**		
	Link number	Cumulative	Link number	Cumulative	% possible OD pairs	Link number	Cumulative	% possible OD pairs
500	2.636	2.636	1.328	1.328	50,38	1.090	1.090	41,35
1.000	5.256	7.892	2.854	4.182	54,3	360	1.450	0,68
1.500	4.128	12.020	2.246	6.428	54,4	4	1.454	0,01
∞	3.364	15.384	1560	7.988	46,37	0	1.454	0,00

provided. In particular, the remaining share of direct connections is almost exclusively made up of links between 500 and 1.000 km long, with the exclusion of 4 links, which exceed in route length the upper threshold. In contrast, the availability of direct air connections tend to be the highest between 500 and 1.500 km, serving more than 50% of all the possible OD pair connections. Thus, rail despite supplying a competitive number of direct routes under 500km, appears to be overwhelmingly behind on longer ranges. This is particularly interesting in light of the widely supported argument that rail can offer competitive services on routes up to 1.000km, with some suggesting that could take some market shares from air even on longer distances. In fact, despite most of the possible routes (34,16%) in the network have a distance between 500 and 1000 km, rail currently supplies an extremely limited number of services in this segment of the market. This, together with the fact that air serves more than half of those connections, indicates the importance of this market and suggest that there is an unexplored potential for rail to capture this demand. Some plausible causes of this are the fragmentation of the European rail market and the infrastructural and capacity constraints. The former, in fact, hinders the supply of cross-border international trains, whilst the latter limits the supply of long-distance services in denser areas (i.e. the Netherlands), where large shares of capacity are consumed by local and regional services. Furthermore, the structure of the rail service supply, with the considerable focus on shorter routes, suggests that rail might rely on transfer services rather than on direct connections to cover longer-distance markets. This, has a twofold set of implications. On the one hand, the considerable number of transfers might negatively impact the attractiveness of the mode. In fact, as discussed in Section 3.8 when defining the theta parameters, transfer time is often perceived in a particularly negative way by passengers. Consequently, imposing transfer services on passengers might greatly reduce the appeal of the mode, leading to them choosing direct air alternatives. On the other hand, relying on transfers imposes strict and ambitious requirements in terms of schedule coordination and disruption management, to allow the feasibility of transfers, to ensure their attractiveness in terms of waiting time and to avoid the propagation of delays along the network. Failing in that could further damage the attractiveness of the mode, leading travellers to prefer comparable air services. These matters are certainly problematic also in the aviation field, but the specific structure of the European rail supply network appears to create the conditions to make the issue even more crucial.

Table 5.6: Influence of distance on the number of direct connections under competition and monopoly

Distance Threshold (km)	**Unique Supplied OD pairs**		**Inter-modal Competition Connections**		**Rail Monopoly Connections**		**Air Monopoly Connections**	
	Link number	% possible OD pairs	Link number	% unique connections	Link number	% unique connections	Link number	% unique connections
500	1.774	67,30	644	36,30	446	25,14	684	38,56
1.000	2.908	55,33	306	10,52	54	1,86	2.548	87,62
1.500	2.246	54,41	4	0,18	0	0	2.242	99,82
∞	1.560	46,37	0	0	0	0	1560	100

The first distinction is made between direct link under a regime of monopoly and inter-modal competition. The figures are provided in Table 5.6, which highlights the total number of unique supplied OD pairs (i.e. OD pairs which have at least a direct connection either in the rail or air network), of links with inter-modal competition (i.e. both air and rail provide direct service on the route), and of links with a monopoly of either rail or air. It is possible to notice how the number of unique supplied OD pairs, despite decreasing with the growth of OD distances, tends to be rather balanced especially between 500 and 1500 km. This, however is allowed by the supply of air services that account for the quasi-totality of unique supplied OD pairs on routes longer than 500 km. In particular, under 500 km the shares of links with inter-modal competition, rail monopoly and air monopoly appears to be quite evenly distributed, with a slight majority of links being supplied by only air services. Over the 500 km the market appears to be almost exclusively served by air services, with rail contributing only in a minimum part. Finally, the distribution of direct links under a competition regime between competitive and substitution-ready connections is provided in Table 5.7. It is possible to notice that there is a limited but non-negligible number of routes where currently rail has the potential to substitute air services. Furthermore, some others could become part of this selection by either reducing the travel time thresholds or by increasing the frequency thresholds employed for the selection of the links. The number of competitive links is considerably higher, indicating that rail can effectively compete with air on a wide set of routes, even on links without a sufficient demand to justify substitution. A complete overview of the distribution of the OD pairs across the categories identified within this study is provided in Figure 5.6. It is possible to notice how more than 50% of all the possible OD pairs (with distance greater than 100km) are connected either by air or rail. In particular, the large majority (82,87%) is connected only by air, with rail monopolies accounting for a mere 5,89% and the remaining 11,24% of OD pairs being served by both modes. Furthermore, over more than half (i.e. 63,31%) of the OD pairs under competition rail and air are non-competitive, so one has an edge over the other in terms of both travel time and frequency. This finding is particularly interesting, as it suggests that other important factors might influence passengers' modal choice, as highlighted in Section 3.1.2. Finally, the limited number of substitution ready OD pairs shows that in many cases demand is not large enough to allow for considerable investments in rail infrastructure and services required to guarantee the modal competitiveness of rail with air. This calls for more specific demand estimation models to forecast the evolution of the market and the economic feasibility of future investments.

Table 5.7: Relationship between distance and distribution of OD pairs by category

Distance Threshold (km)	Competitive			Substitution-ready			Substitution Potential	Substitution Attractive
	Fully	Part. (tt-wise)	Part. (f-wise)	Fully	Part. (tt-wise)	Part. (f-wise)		
500	27 (58,70%)	147 (71,36%)	50 (51,02%)	13 (68,42%)	15 (93,75%)	8 (61,54%)	11 (35,48%)	34 (23,13%)
1000	19 (41,30%)	59 (28,64%)	48 (48,98%)	6 (31,58%)	1 (6,25%)	5 (38,46%)	19 (61,29%)	69 (46,94%)
1500	0	0	0	0	0	0	1 (3,23%)	44 (29,93%)
Total	46	206	98	19	16	13	31	147

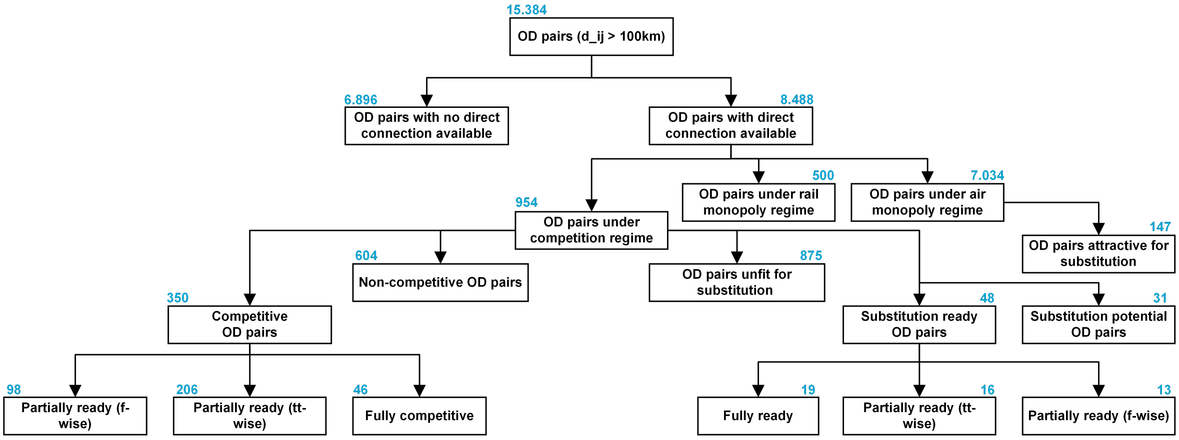


Figure 5.6: Distribution of the OD pairs across the categories identified within this study

5.3.1. Spatial Distribution of OD Pair Connections under a Monopoly Regime

In order to highlight the spatial distribution of the aforementioned corridors, they are plotted on maps. In particular, to provide some additional insights into the quality of the links the difference in connectivity (air - rail) is projected using a scale of colour ranging from green, to white and blue. Darker shades of green indicates negative difference values, representing the links where rail most comprehensively dominates air in terms of connectivity. Lighter colours represent links with connectivity differences around nil, where the degree of connectivity is similar across both modes. Finally darker shades of blue represent increasingly air-dominated links in terms of connectivity. This twofold visualisations aims to provide an overall overview of the quantity, quality and spatial locations of links. In particular, including the difference in connectivity allows to understand how the connectivity on each link compares across modes, providing interesting insights for policy-making.

Figure 5.7: Spatial Distribution of the Links under a Rail Monopoly Regime

Figure 5.7 portrays all the routes under a rail monopoly regime. A majority of these links consists of routes either national or short, whilst long-distance cross-border connections represent only a minority. Some notable exceptions are the connections between Northern Italy and Switzerland (e.g. Geneva-Venice, Zürich-Genova) and France (e.g. Lyon-Milan), the routes connecting Berlin to Poland (e.g. Berlin-Gdańsk) and Czech Republic (i.e. Berlin-Ostrava), Paris and western Germany (e.g. Paris-Cologne) and Rome to Salzburg. Furthermore, it is interesting to notice that the largest share of cross-border rail monopolies concentrates between Austria, Hungary and Germany, suggesting that the national systems of these countries are integrated, at least up to a certain degree. In particular, using the link connectivity difference it is possible to identify two main link categories within the map. On the one hand, those links where rail heavily outperforms air which can be identified by darker shades of green. This characteristic implies that the good performances of rail services, largely outweighing the ones of the air network, are probably keeping air outside of the market. Generally these links cover short-distance routes where rail necessarily performs better than rail. However, some notable exceptions with longer distances are the Naples- Florence and Naples-Bologna corridors, both served by an efficient high-speed rail service. On the other hand, the white links highlight routes where the differences between the two modes are less pronounced. A notable example is the case of Zaragoza, which features a considerable share of rail monopolies. Some possible causes for this are the central location of the city within the Spanish rail infrastructural system, the small size of its airport and consequently the absence of a strong demand for international flights. In particular, the latter relates the performing high speed rail connections that connect Zaragoza to both Madrid and Barcelona, allowing passenger to effectively employ inter-modal service compositions to reach further destinations. It is also interesting to notice that air does not provide a better alternative in terms of connectivity on any of the routes under rail monopoly. It is also interesting to notice that in central Europe a large share short-connections are rail monopolies despite not rail is not overwhelmingly more performing compared to air. This suggests that despite the rail system is well connected and offers smooth cross-border connections, its performances have still space for improvement.

Figure 5.8 illustrates the spatial distribution of all the air monopoly routes on distances between 100 and 500 km. It is interesting to notice how also in this case the connectivity on the majority of links is rather balanced between the two modes, even though there are some air monopoly links where connectivity is higher for rail. This implies that despite imposing one or more transfers on such connections, rail might still be more convenient for passengers than taking a direct flight. In this regard, it is particularly interesting the case of the Southampton-Newcastle corridor, which features an especially high connectivity for rail, probably due to the high frequencies of trains transiting within London. However, it is important to consider that London doesn't really have a centralised rail system that allow for easy and quick transfers, due to the considerable number of head stations. Thus, providing a direct express service between Newcastle and Southampton could actually allow to foster competition on the corridor, maybe even leading to the substitution of the air connection. In contrast, corridors like Madrid-Porto, Utrecht-London and Copenhagen-Oslo appear to be better connected by air. In these cases, the transfers are clearly penalising, and it is important to provide direct services in order to increase the connectivity of rail and make the mode more attractive. Analysing the map, it appears that there are multiple reasons why air monopolies are still present on short-routes where rail is supposed to be able to provide more performing services. First of all, the detours that the geographical conformation of the territory imposes to rail. This is for instance the case of the Rome-Nice and Barcelona-Nice connections, where the land imposes considerable detours compared to a direct flight connection that cuts across the sea. Similarly, in the UK only London is served by a direct rail connection to the European continent, due to its central position that attracts and redirects nationally and internationally all the traffic coming from the country. Secondly, the limitations and constraints imposed by the available rail infrastructure. For instance between the Netherlands and Northern Germany (Bremen and Hamburg), it is clearly noticeable the gap in terms of infrastructure. Another interesting case regards Spain. Its radial system centred in Madrid in fact leaves out most connections connections between the urban

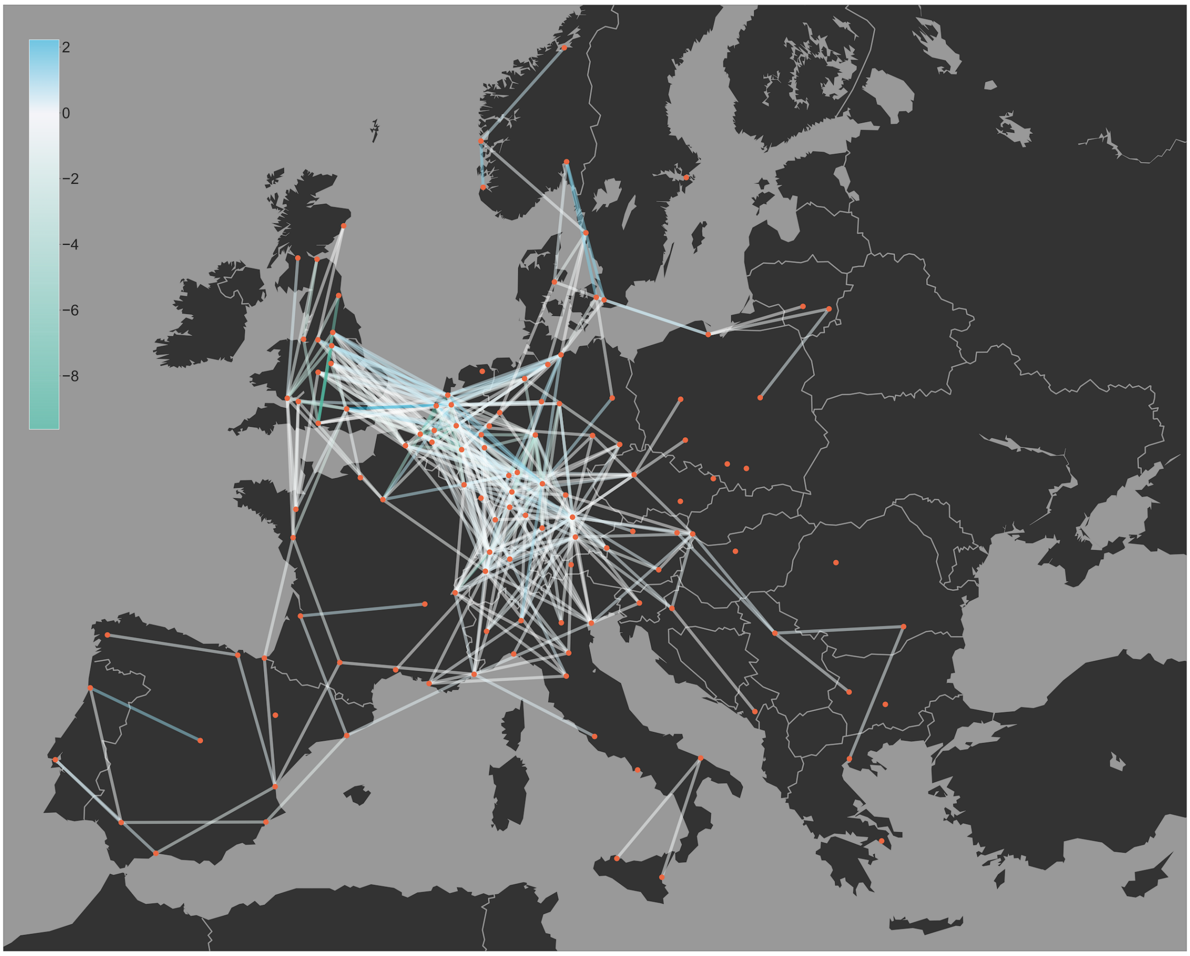


Figure 5.8: Spatial Distribution of the Connections within a 500 km distance threshold under an Air Monopoly Regime

areas on the coasts, which in fact appear to be air monopolies. Similarly, in Norway all rail services run directly to the capital city, Oslo, leaving the other cities disconnected from each other. Furthermore, in the south of Italy the relatively close cities of Bari and Catania are not connected by a direct rail service. This is due to the poor rail infrastructure in the southern regions of Italy, and to the structure of the Italian rail network that mostly runs on parallel lines on the two coasts. In both the Norwegian and Italian case, it is important to highlight the rather complex topology and morphology of the terrain, which makes it very expensive to build infrastructure. Finally, another possible reason is the high density of rail services, which implies high frequencies and high capacity consumption. This, is for instance the case of central Europe, where surprisingly a considerable number of urban areas is not directly connected by rail. However, in this case the limitations in terms of capacity might mean that it is not possible to offer more direct services, or that the high service frequencies allow for short transfers which are make indirect services attractive enough to avoid the necessity for a direct service.

5.3.2. Spatial Distribution of Competitive OD Pair Connections

The spatial distribution of competitive links is shown is Figures 5.9, 5.10 and 5.11. The first map allows to identify the links where rail is effectively competing with air in terms of both frequency and travel times. Furthermore, the colour scale allows to understand whether either rail (i.e. green links) or air

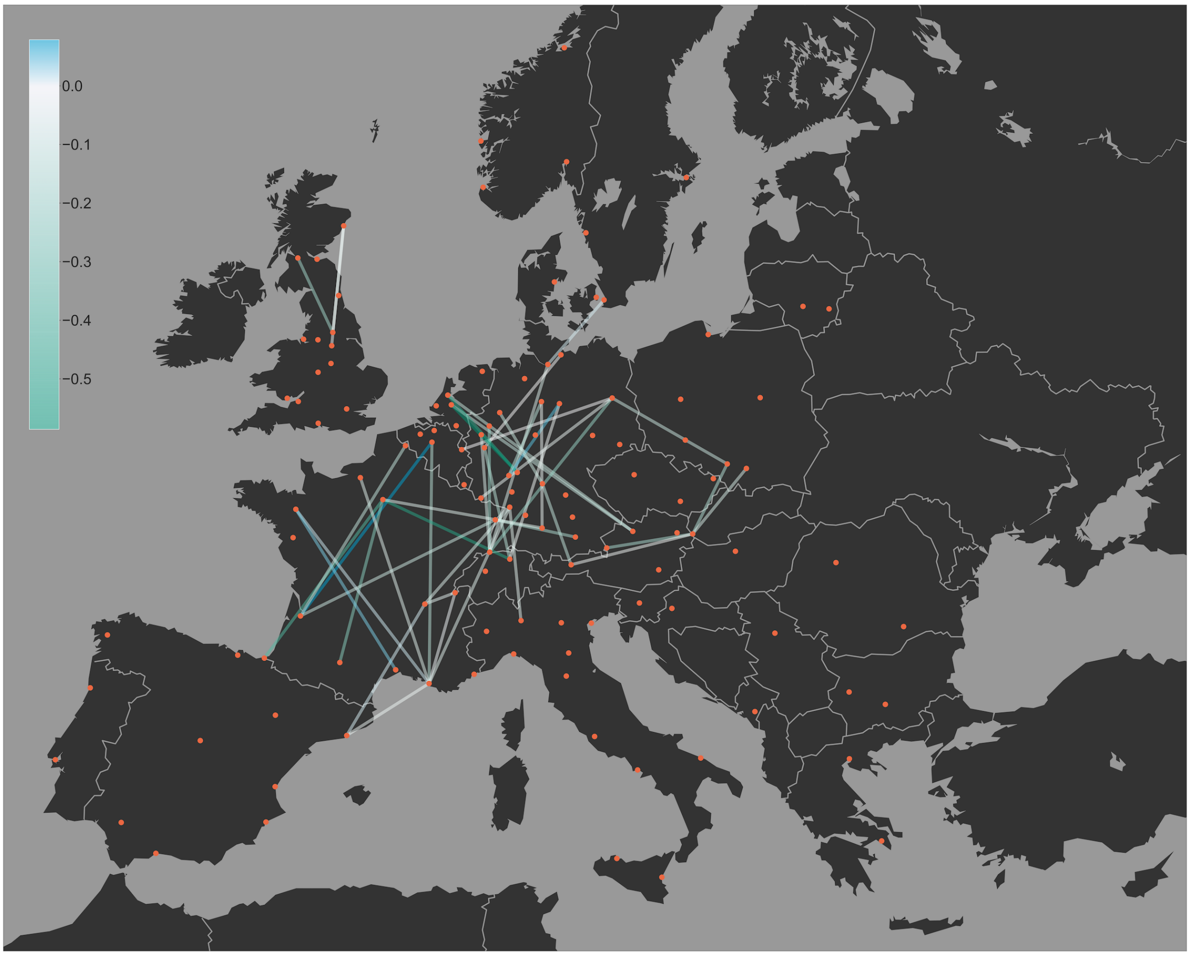


Figure 5.9: Spatial Distribution of the Fully Competitive Links

(i.e. blue links) have an edge in terms of the aforementioned connectivity or whether the two modes are quite evenly competing on the route (i.e. white links). A considerable share of the fully competitive corridors appear to depart and arrive either within France or Germany. In particular, there is a range of cross-border corridors that appear to be competitive, including the Barcelona-Marseille/Lyon, the Milan-Karlsruhe, the Amsterdam-Linz, the Paris-Zürich and the Frankfurt-Amsterdam corridors. Furthermore, the connections between Bratislava and both Poland (i.e. Katowice and Krakow) and Austria (i.e. Salzburg and Linz) also appear to be competitive, featuring similar degrees of connectivity across both modes.

The map, shown in Figure 5.10, highlights illustrates all those links where rail travel times are competitive with rail. Links coloured in blue represent connections where air frequency is higher, whilst green links portray routes with more rail frequencies. It is possible to notice some patterns in terms of the spatial distribution of such links. In most countries (i.e. Germany, Italy, UK, Poland, Czech Republic, Austria, Switzerland and Denmark) national routes with comparable travel times feature considerably higher rail frequencies compared to air. Two exceptions are Sweden and Spain, which appear to feature more limited differences between the two modes in terms of frequencies. Furthermore, in France, some domestic routes, notably Paris-Nice, appear to be better connected by air. Considering these spatial distribution patterns it could be noticed a correlation between country size and population density

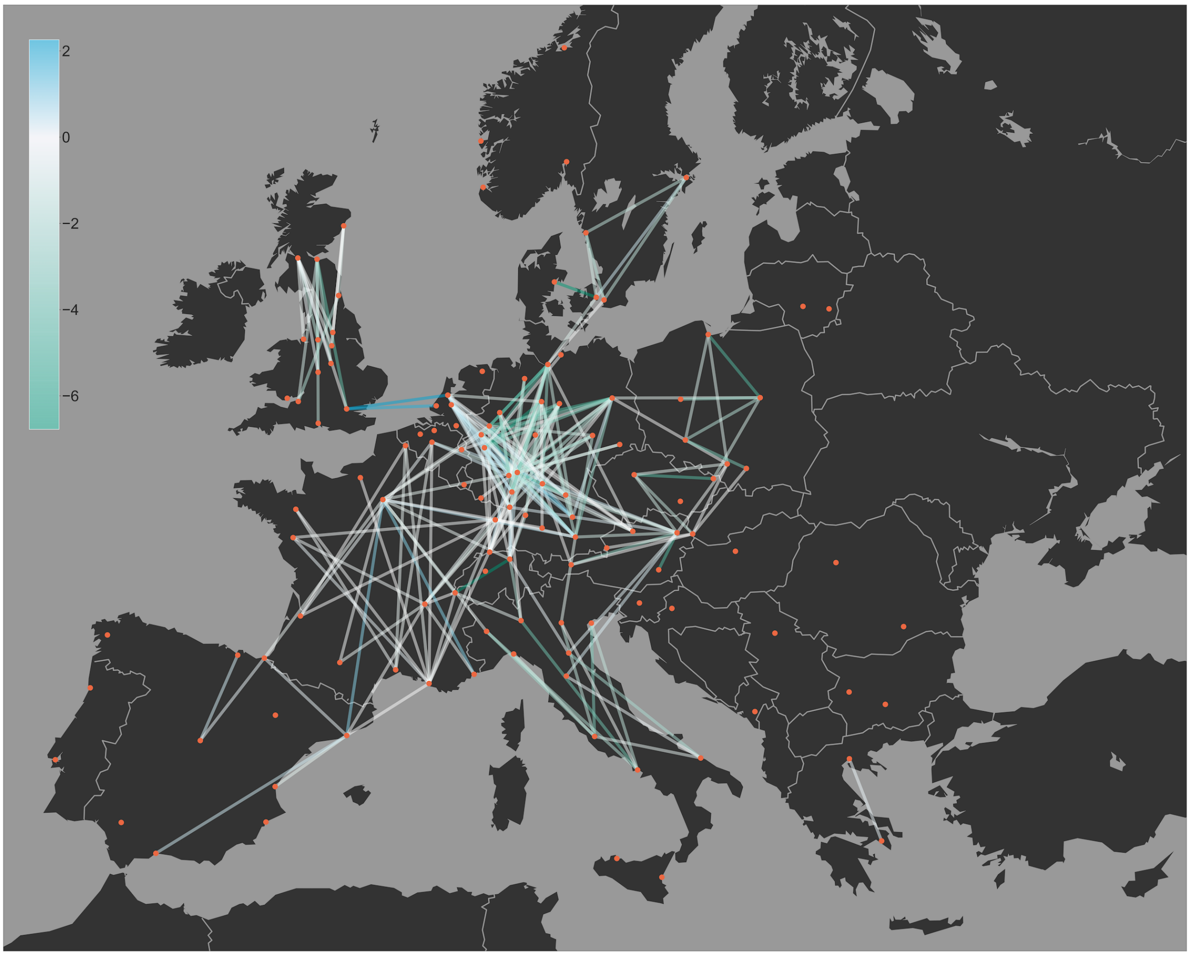


Figure 5.10: Spatial Distribution of the Partially Competitive Links (travel time wise)

and rail frequencies, as France, Spain and Sweden are the three largest countries in terms of area. On cross-border connections rail appears to generally have a more limited superiority in terms of performances over air compared to domestic routes. In particular, the Paris-Barcelona and London-Amsterdam connections, despite featuring similar travel times appear to have considerably higher frequencies for air. In this regard, it is important to highlight that increasing the frequencies on such links could allow to create a state of full inter-modal competition between the two modes, stimulating further the growth of the rail market.

The map illustrated by Figure 5.11 provides insights into the connections that feature competitive frequencies across the two modes. In this case the green colour represents links where rail travel times dominates and blue links represent links where air travel times are more performing. In contrast to the previous case a wider share of links with similar frequencies feature higher connectivity for air. An interesting finding relates to the contrasting patterns found in France and Germany. Whilst in the former country, routes with competing travel times generally feature higher air frequencies, in the latter rail frequencies appear to considerably exceed the ones of air. At the same time, on routes with similar frequencies rail appears to be more competitive in terms of travel times in France whereas in Germany air appears to have an edge. These differences can be explained by the type of rail services that the two country offer. Whilst in Germany, rail is focused on offering wide-spread coverage and high frequen-

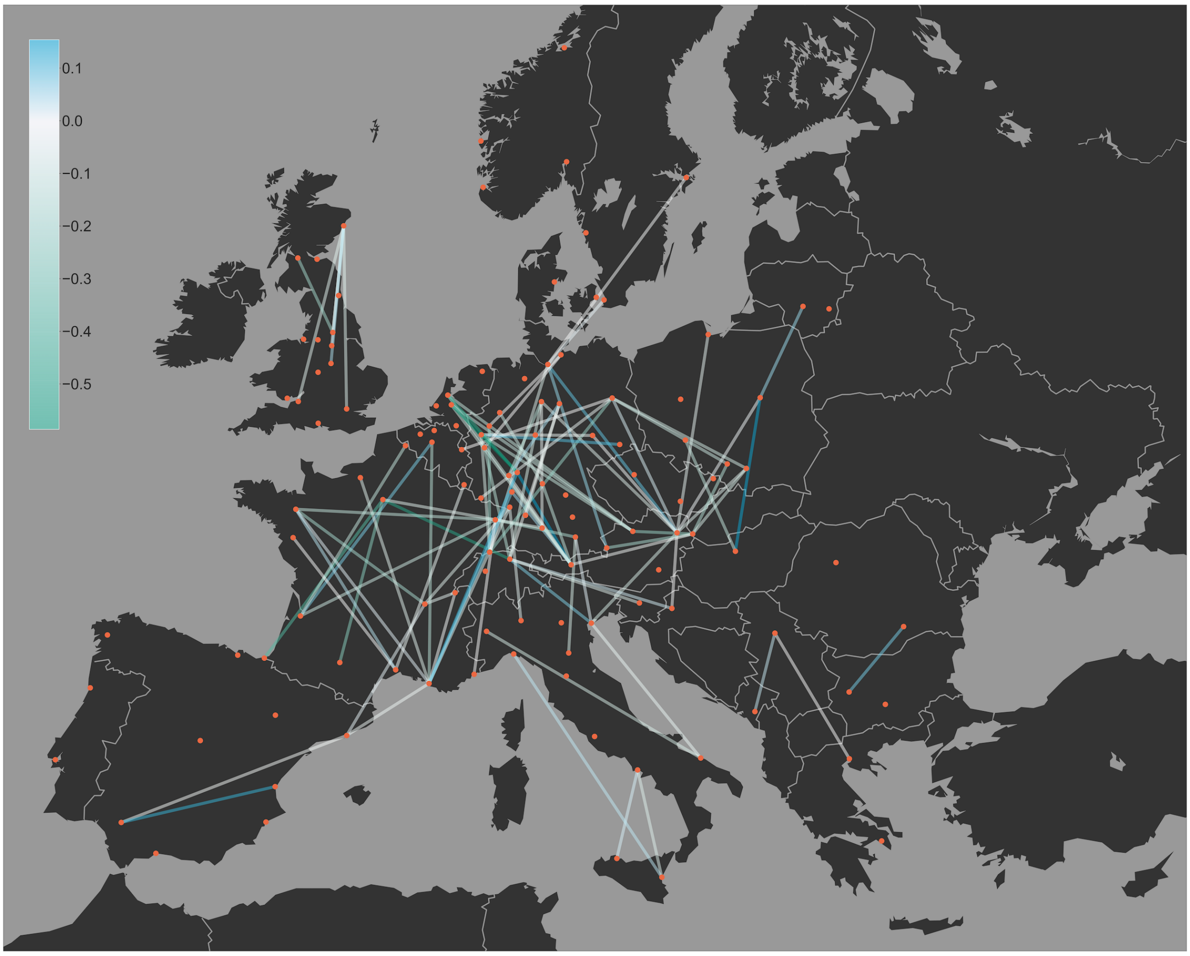


Figure 5.11: Spatial Distribution of the Partially Competitive Links (frequency wise)

cies, in France rail aims to offer competitive travel times using high-speed rail. Overall, adding to the insights provided by Figure 5.4, it could be concluded that on links with similar travel times generally rail provides more frequencies, whilst on links with similar frequency air features lower travel times. Finally, in terms of the spatial distribution of lines, it is interesting to notice the Amsterdam-Frankfurt and Paris-Zürich corridors, where frequencies are similar across the two modes, despite rail travel times are shorter. The presence of high air frequencies despite the less competitive travel times is probably related to the provision of feeder services towards the main European air hubs to serve intercontinental destination. Air frequencies on these routes could be thus decreased fostering air-rail inter-modality.

5.3.3. Spatial Distribution of Substitution-ready OD Pair Connections

The routes that, at present, offer the right conditions for the substitution of air services with rail are shown in Figure 5.12. These represent the connections where there is a considerable air demand (i.e. at least 5 daily flights) that could be captured by rail providing similar or better services. Across all these routes rail is deemed able to effectively replace air connections without reducing the level of supply quality. Also in this case, the colour scale highlights the modal difference in the connectivity degree of links, where darker greens represent an increasing dominance of rail over air. The route with the greatest potential for substitution is the Milan-Rome corridor. The high substitution potential of

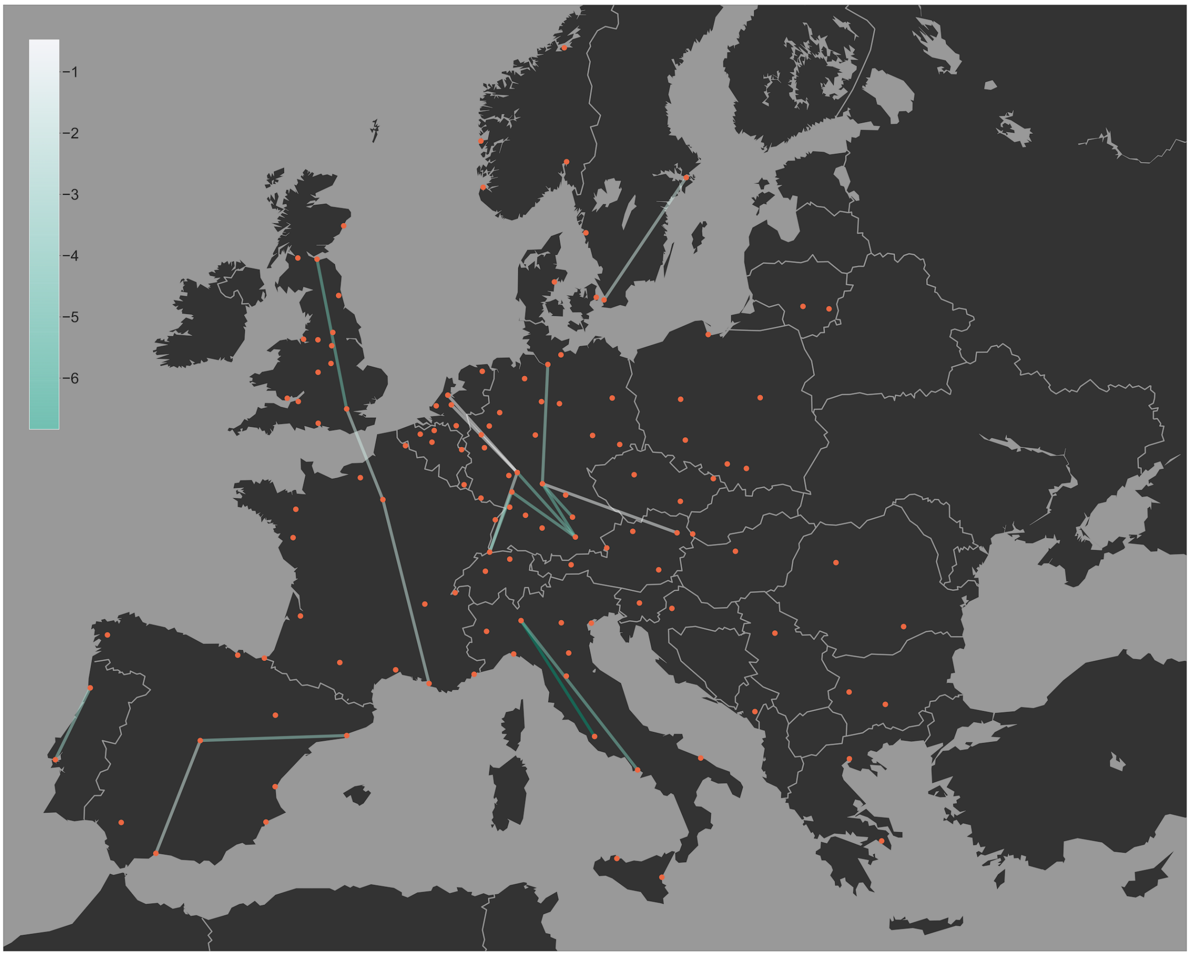


Figure 5.12: Spatial Distribution of the Fully Substitution-ready Links

this route is probably due to the very effective high-speed connection, which offers short travel times and high frequencies. Given the high quality of the supplied rail service, the air connections available between this OD pair are probably mostly employed to offer connections for international/intercontinental transfer passengers. If that is the case, air services could be substituted offering alternative rail services between airport terminals with inter-modal fare integration. Probably, similar solutions could also be useful for the Paris-Marseille, the London-Edinburgh, the Porto-Lisbon and the Stockholm-Malmö routes, which have very similar characteristics, all connecting two of the major cities within the country, despite the absence of a high-speed service in the latter three cases. Similarly, this could also be argued for routes connecting major European airport hubs, such as Madrid-Barcelona, Frankfurt-Munich, Amsterdam/Utrecht-Frankfurt and London-Paris, given that air services on such connections, as previously mentioned, are most probably in place to offer inter-hub connectivity. In this regard, it is worth noting that the travel times used in this analysis capture the total door-to-door extent of the trip. However, when considering these routes as legs of longer indirect connections rather than direct connections the travel time should be computed as in-vehicle time and transfer connection only. For this reason, in many cases the advantage of rail over air in terms of connectivity might be overestimated for such services. It is, thus, deemed crucial to analyse each route to clearly understand the substitution requirements of the specific connection. However, it is envisaged that in most cases frequent rail services directly connecting the airports are required to effectively substitute the current air transfer services.

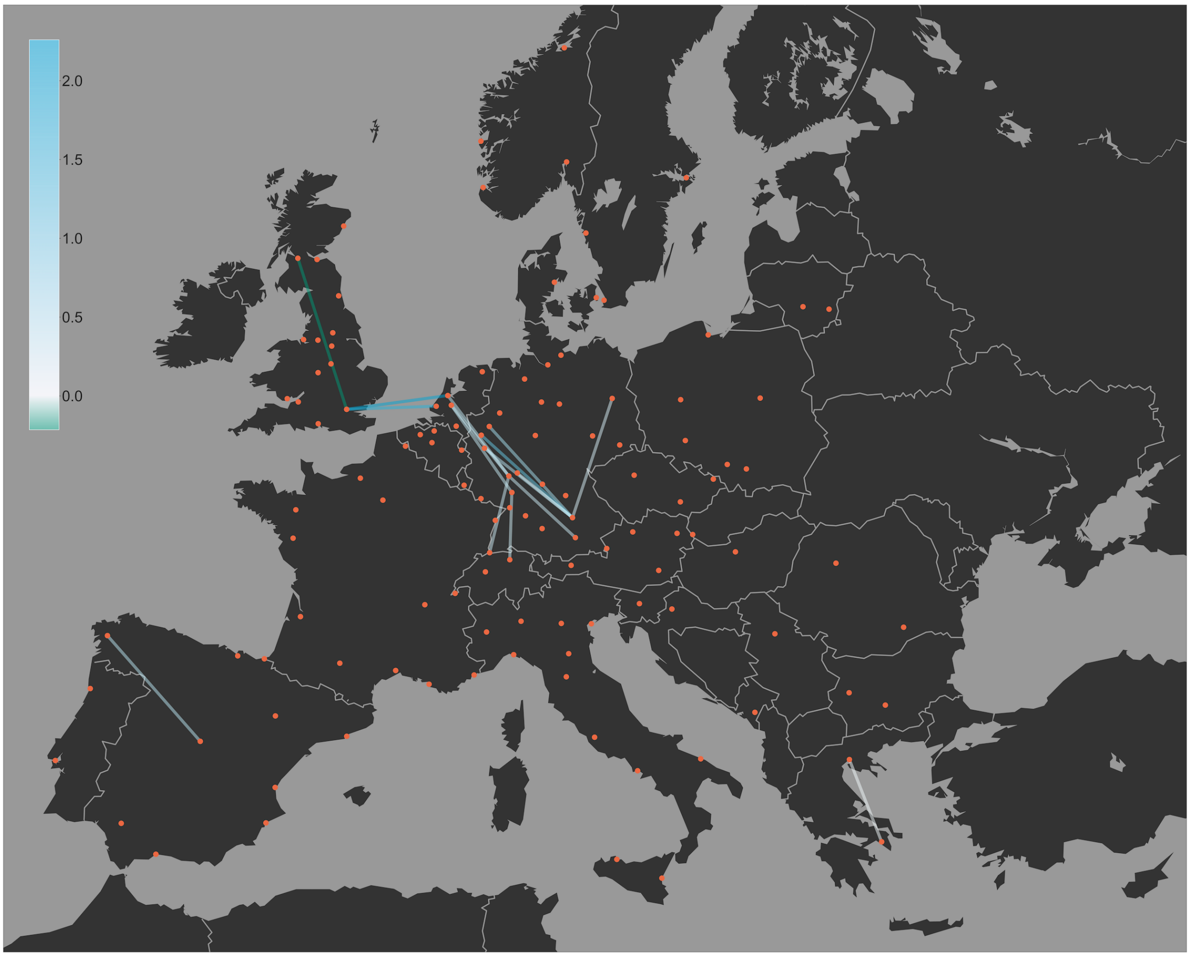


Figure 5.13: Spatial Distribution of the Partially Substitution-ready Links (travel time wise)

Figure 5.13 shows all the routes where an increase in frequencies is required in order to allow for the complete substitution of air services. Thus, on these routes, despite rail provides competitive travel times, frequencies are not as attractive. In particular, it is worth noticing how connectivity for these routes is generally lower for air than for rail counterparts, in contrast to the routes identified by Figures 5.12 and 5.14. This is probably due to the important role of frequency in computing connectivity. In this case, an interesting connection is the Glasgow-London route, where rail connectivity most clearly outperforms air connectivity. On the other hand, the London-Rotterdam/Amsterdam route, already mentioned in the description of the partially competitive links, appears to be the route where air connectivity, and consequently air frequencies, have the greatest advantage over rail counterparts. The remaining links are mostly concentrated between the Netherlands, Germany and Switzerland, with the exception of the Madrid-Santiago de Compostela route in Spain and all feature rather balanced values for the connectivity of the two modes with air being generally more performing.

Figure 5.14 captures the routes where a decrease in travel times is needed to allow for the complete substitution of air services. On these routes rail provides competitive frequencies but fails to provide travel times lower than rail alternatives. It is interesting to notice that the connectivity of these links is

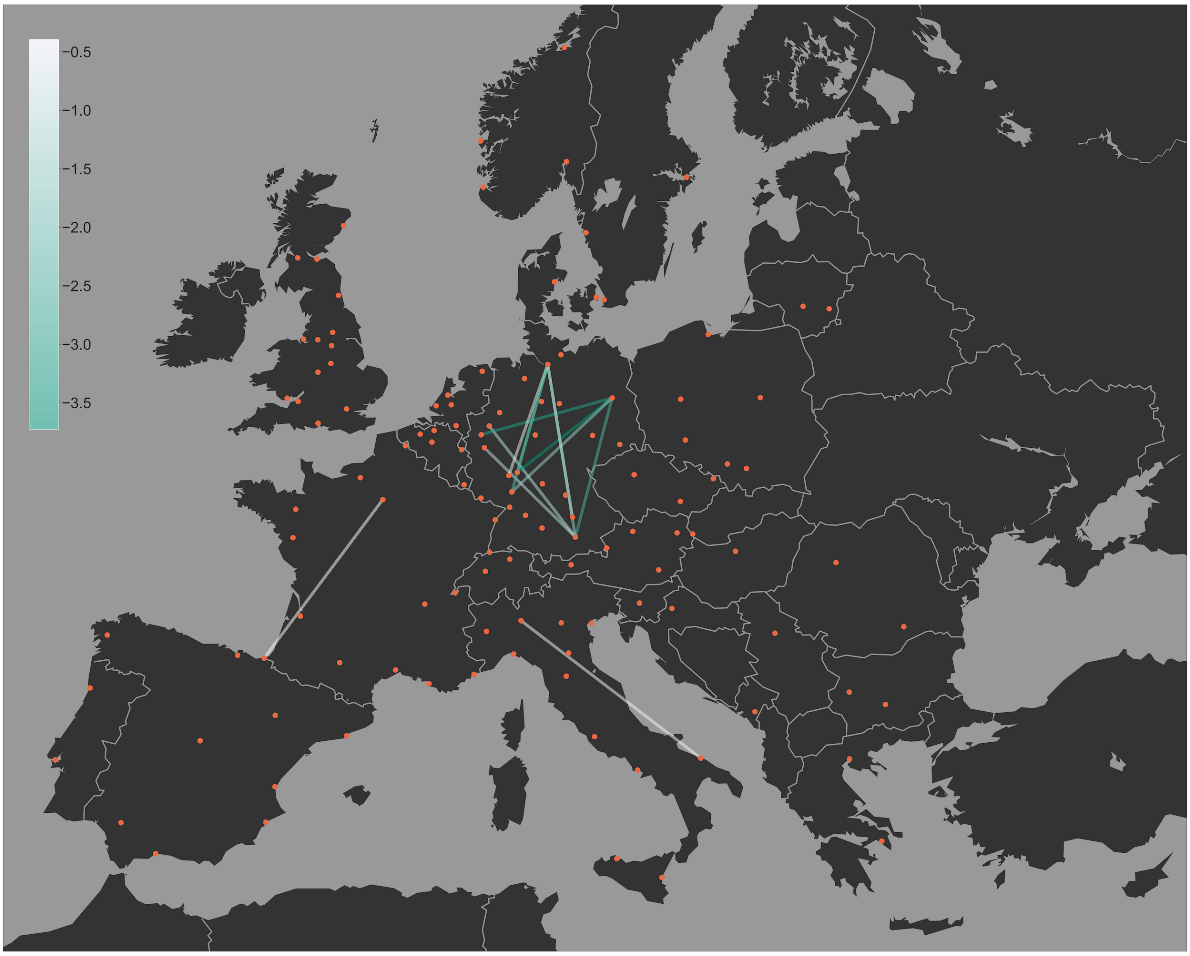


Figure 5.14: Spatial Distribution of the Partially Substitution-ready Links (frequency wise)

always higher for rail. Most routes are located within Germany, notably the corridors connecting Berlin to Düsseldorf, Frankfurt to Munich, Hamburg to Frankfurt and Munich, and Munich to the Ruhr and Cologne. Only two links are located outside of Germany, one being the Bari-Milan and the other one being the San Sebastián-Paris. Over all these routes a decrease in travel times would allow rail to outperform air services. Thus, improving the infrastructure or providing faster services would lead to increased likelihood of substitution. In conclusion, Germany appears to be the country which currently features the highest share of substitution-ready routes. Considering the current rail supply scenario in the country it could be argued that the eventual implementation of high-speed rail services, providing shorter travel times, might lead to an increased attractiveness of rail which could in turn contribute to the substitution of a consistent portion of the national flights currently available in the market. However, more specific studies are required to assess the actual costs and benefits of such scenario, and its implications for the entire European long-distance transport market.

The routes with substitution potential are highlighted in Figure 5.15. These are routes under a competition regime with a considerable demand (i.e. more that 35 weekly air services), where rail is currently not competitive either from a travel time or frequency perspective. These corridors have a considerable demand that could be captured by rail services if their performances were to be improved, without the need to create new services. Notable examples are the Lisbon - Madrid, and the routes linking Paris to

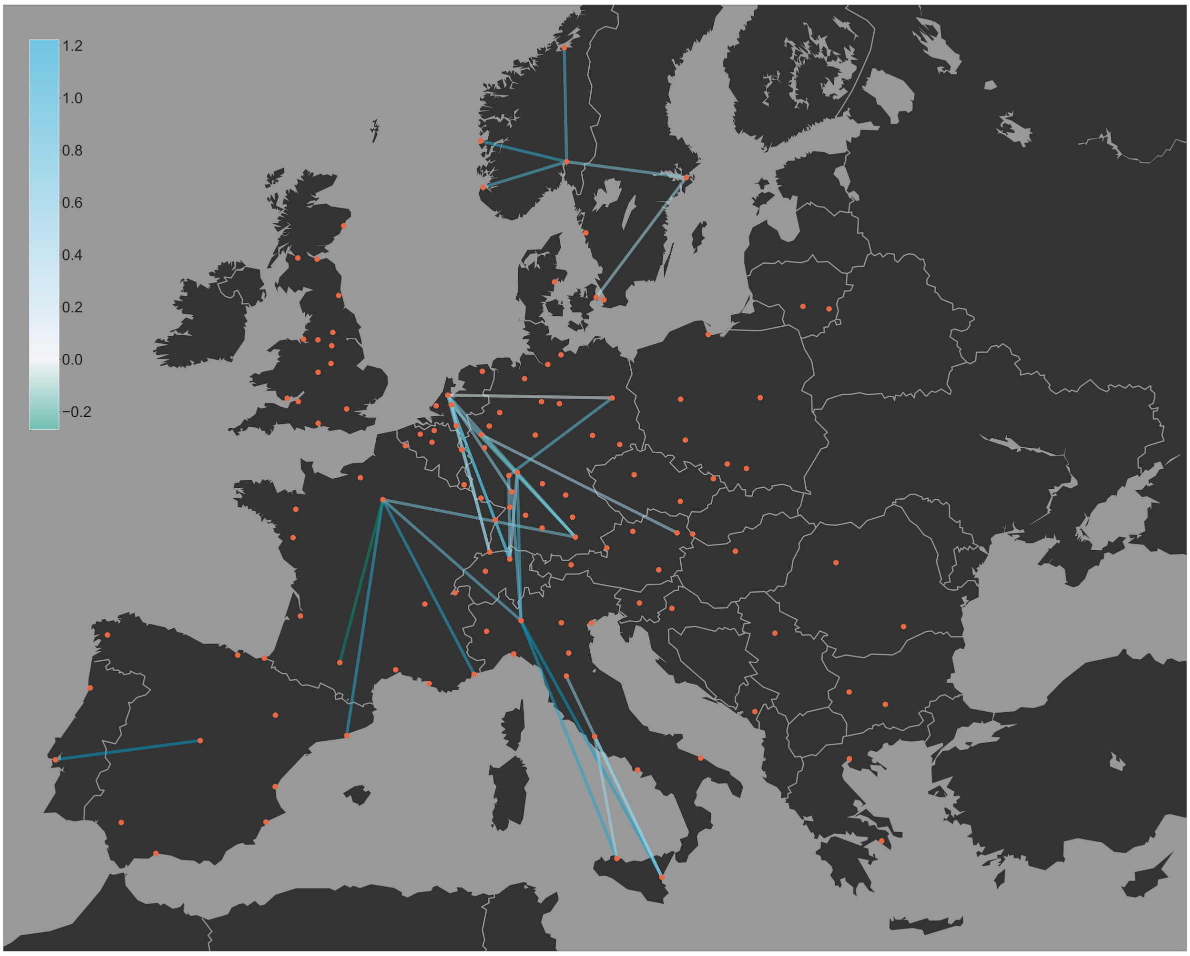


Figure 5.15: Spatial Distribution of the Routes with Substitution Potential

Barcelona, Toulose, Nice, Milan and Munich, the Zürich to Amsterdam and Frankfurt, and the Wien - Düsseldorf. Furthermore, Figures C.1, C.2 and C.3 highlight the OD pairs attractive for substitution, on the 100 - 500, 500 - 1.000 and 1.000 - 1.500 km market segment respectively. These are links with frequent air service connections (i.e. more that 35 weekly air services) that are not directly connected by rail services. Excluding a few links where the geographical morphology of the terrain prevents rail to provide competitive services (e.g. the Bergen - Stavanger and the Manchester - Amsterdam on the 100 - 500 km segment, the Copenhagen-London and the Oslo-Amsterdam on the 500 - 1.000 km segment and the Rome-Madrid the 1.000 - 1.500 km segment), the remaining links represent corridors that are attractive for substitution. In particular, the high air demand on these lines suggest that implementing new rail services on these OD pairs would probably be economically attractive in the long run, even if some investments are required in the short term. Notable examples in the 100 - 500 km market segment are the Oslo-Malmö/Copenhagen, the Rotterdam - Frankfurt, the Utrecht - London and the Düsseldorf - London. In the 500 - 1.000 km it is worth noting the Milan - Barcelona, the Amsterdam - Copenhagen, the Franfurt - London, the Amsterdam - Geneva, the Ruhr/Düsseldorf - London and the Zürich - London. Finally, notable examples in the 1.000 - 1.500 km market segment are the Barcelona - Lisbon, the Paris - Porto/Lisbon, the Copenhagen - Paris and the Florence - Amsterdam. These are all corridors where new rail service connections could be launched to attract part of the passengers that currently use air services.

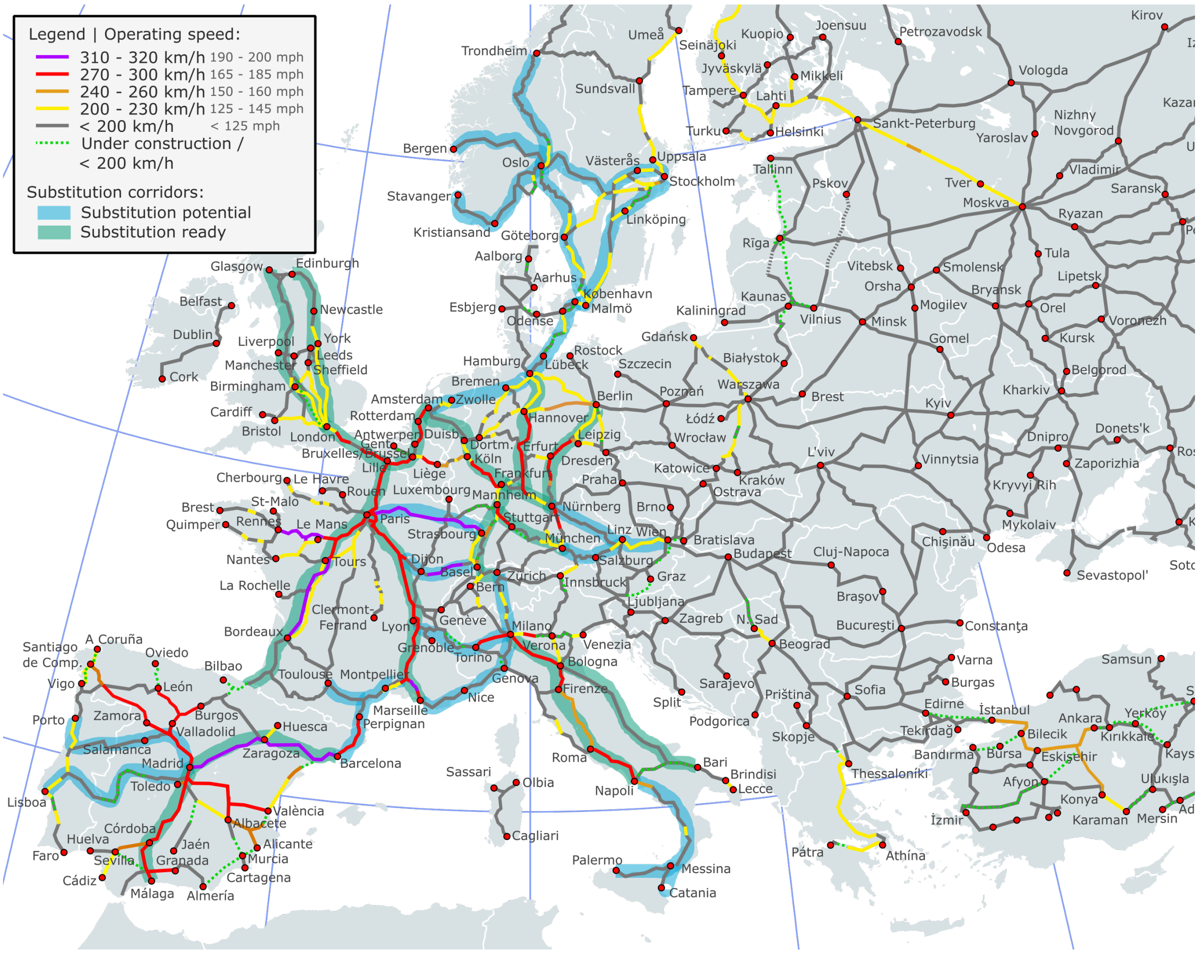


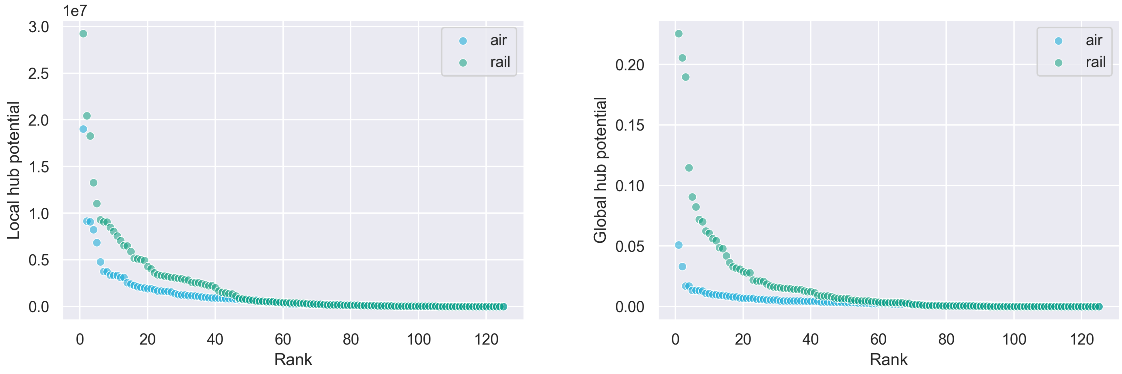
Figure 5.16: Substitution Potential and Substitution-ready Routes at the Infrastructure Level (Source: adaptation from Wikimedia Commons, 2022)

Finally, the substitution-ready links and the routes with substitution potential/attractiveness are projected over a map of the rail infrastructure. Figure 5.16 show the European rail infrastructure including the maximum speed on each section of the tracks. All the infrastructure sections employed by the substitution ready links highlighted in Figures 5.12, 5.13 and 5.14 are highlighted using the green colour. These represent track sections traversed by rail services that directly compete with air on routes where the latter manages to capture a considerable share of the demand, despite the comparable or worse performances. Thus, investing in improving the maximum speed and/or capacity on these links might allow rail to capture the aforementioned demand over these routes. In contrast, the infrastructure sections required to make rail competitive on routes up to 1.000 km with a considerable air demand are highlighted in blue. Investing in these routes could also prove beneficial, however, considering the current poor performances of rail more efforts are probably required. Furthermore, this analysis provides further insights on the causes of the failed competitiveness of rail with air. For instance, between Spain and Portugal an important lack of rail infrastructure is evident, which is most probably hindering the capacity to offer competing services on the Lisbon-Madrid corridor. Moreover, between Barcelona and Paris, there appears to be a gap in the high-speed infrastructure, which also relates to the connectivity of the Toulouse-Paris corridor. Another considerable infrastructure gap is the one between Scandinavia/Northern Germany and the Netherlands/Belgium, as no rail infrastructure currently connects

Bremen and Groningen. Other noticeable gaps are the two alp transit corridors, the Frejus, connecting Turin to Lyon and the Gotthard, connecting Milan to Zürich, the connection between Lyon and Nice which requires a detour to go around the alps and the standard rail lines with lower speeds between Germany and Austria and between Sicily and Naples. It is worth noticing that two high-speed corridors, the Paris-Basel and Paris-Strasbourg lines are currently not employed by substitution-ready links. However, due to the state of the infrastructure it is possible to argue that the reasons for this are most probably related to gaps in the service supply rather than to infrastructural deficiencies. In that regard, offering high-speed cross-border services from Paris to a wider set of destinations within Switzerland and Germany would probably allow to reduce the total number of flights currently connecting the capital of France to central Europe.

5.4. Node Connectivity

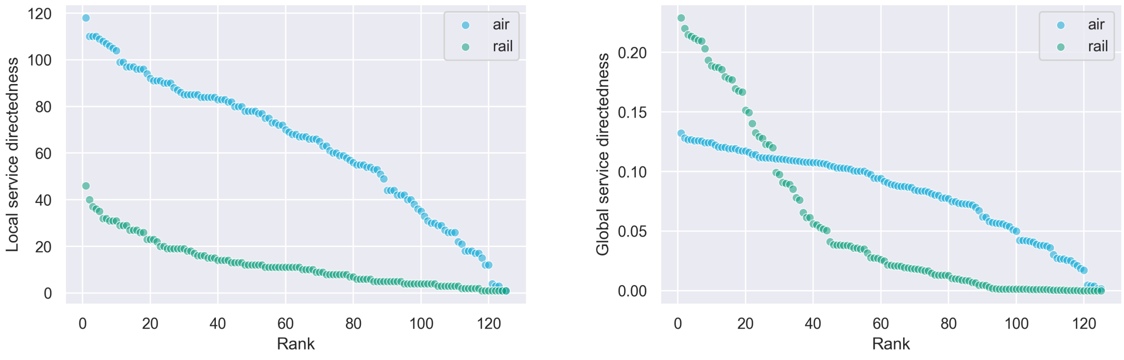
Node connectivity is measured using three different indicators, each computed both at a local and global level, using the methodology described in Section 3.3.3. The distributions of the indicators across the 125 urban areas analysed with the air and rail networks are compared in Figures 5.17, 5.19 and 5.18. At the local level hub potential measures the squared number of frequencies of all direct connections, as described in Formula 3.20. At the global level, this metric aims to capture the role of an urban area within the entire network, highlighting the nodes that are most traversed considering passengers that only aim to minimise transfer time. In particular, it is measured using the betweenness centrality on the networks weighted with the reciprocal of direct link frequencies, as described by Formula 3.21. The distribution show that rail hub potential follows for both rail and air an exponential function, which is in line with the expectations of the indicator, which aims to highlight the major hub. Rail has an edge over air in terms of the magnitudes of the indicator in the major hubs, due to the overall higher frequencies of the rail mode. This could be corrected for by taking into account only high speed services, and excluding slower but more frequent intercity services. In terms of distributions, the rail hub potential follows similar patterns both at the local and global level. In contrast, air hub potential appears to feature a steeper curve with proportionally lower figures in relation to rail counterparts at the global level. This is probably due to the higher density of the air network, which implies a low magnitude of betweenness centrality metrics.



(a) Local **(b)** Global

Figure 5.17: Hub Potential Distribution

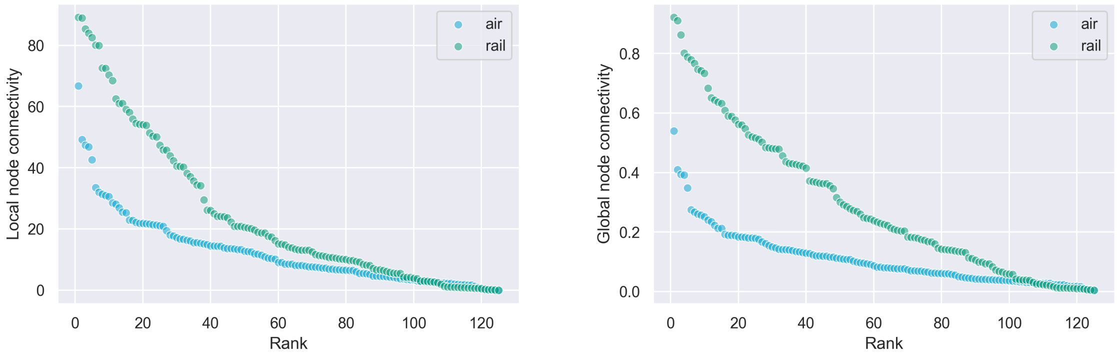
Node directness captures the ease of reaching a destination in terms of directness, expressed by the



(a) Local **(b)** Global

Figure 5.18: Node Directness Distribution

number of direct connections available at the local network and by the relative centrality of the node at the global level. The latter is measured using the eigenvector centrality, which takes into account of the number of connections reachable with increasing transfers, assigning more importance to nodes from which a larger number of nodes can be reached with a low number of transfers. Node directness shows rather different patterns at the local and global level. In the former case, the number of direct connections is consistently higher for nodes within the air network, whereas at the global level rail has higher node directness for the top 25 urban areas. Furthermore, local node directness values follow similar distribution patterns for both rail and air. On the other hand, at the global level the slope of the rail distribution is considerably steeper compared to air. This suggests that whilst the air networks allows to reach any destination in the network with a limited number of transfers from almost every origin, this is not true for the rail network. The considerable heterogeneity in the global node directness of the rail network, in fact, highlights that a few central nodes are well-connected (short amount of transfers required) to all other points, whilst from a large portion of urban areas more transfers on average are required to reach all the rest of the nodes.



(a) Local **(b)** Global

Figure 5.19: Node Connectivity Distribution

The third indicator employed is node connectivity, which represents an aggregate of the link connectivity presented in Section 5.2. At the local level this is computed by summing over the connectivity of all links incident on a node, whilst at the global level is measured as the average connectivity for all direct and indirect connections. This indicator aims to provide a broad overview on the overall connectivity patterns across the nodes of the network. Node connectivity is generally higher for rail compared to air, as the indicator is based on a ratio, which gives more importance to high frequencies (i.e. numerator of the ratio), compared to low travel times (i.e. denominator of the ratio). It is worth noting that to correct for this the logarithm of the Link Connectivity might be employed. Furthermore, it is interesting to notice how the gap in connectivity widens when considering node connectivity at the global level, thus taking into account indirect connections. A reason for this might be that transfer times being related to frequencies, are considerably higher within the rail network. This could be adjusted by taking into account schedule coordination and including actual transfer times. The lower tail of the distributions show that rail features poorer connectivity for the least connected nodes in the network. Thus, the rail network is characterised by higher heterogeneity in connectivity, compared to the air network, where connectivity is more evenly distributed across the nodes. Compared to the hub potential, the distributions are less steep, being almost linear in the case of rail at the global level.

5.4.1. Hub Potential

The first indicator employed to measure node connectivity is hub potential. The spatial distribution of the local hub potential within the air network, as illustrated by Figure 5.20, shows that nodes in central Europe are highly connected in terms of rail frequencies. In particular, the top ten is monopolised by

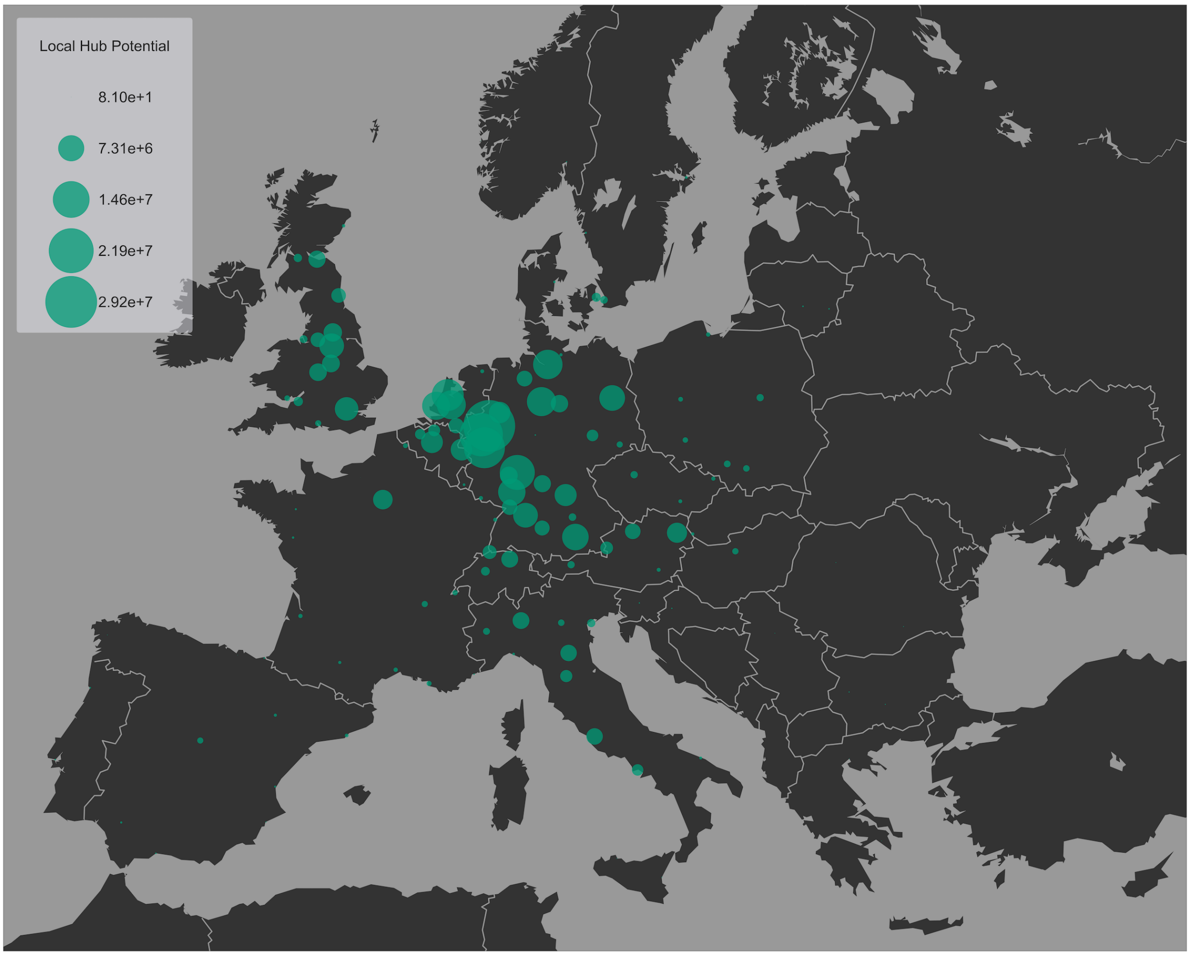


Figure 5.20: Spatial Distribution of the Local Hub Potential within the Rail Network

German and Dutch urban areas. The four most connected urban areas are the closely linked cities of the Ruhr, Düsseldorf, Cologne and Frankfurt. Then, the fifth ranked, Amsterdam is followed by Hamburg and Hannover. Finally the top ten is closed by Utrecht, Rotterdam and Mannheim. From the spatial distribution it appears rather clear that geographical centrality of nodes is quite important for this indicator. The nodes with the lowest degree of local hub potential are Rouen, Podgorica, Stavanger, Bilbao and Athens. Unsurprisingly, all urban areas located in the peripheries of the European rail network. This is due to the fact that rail frequencies are related to the density of urban areas. In particular, the German rail system, as previously highlighted appears to favour high frequencies over short travel times.

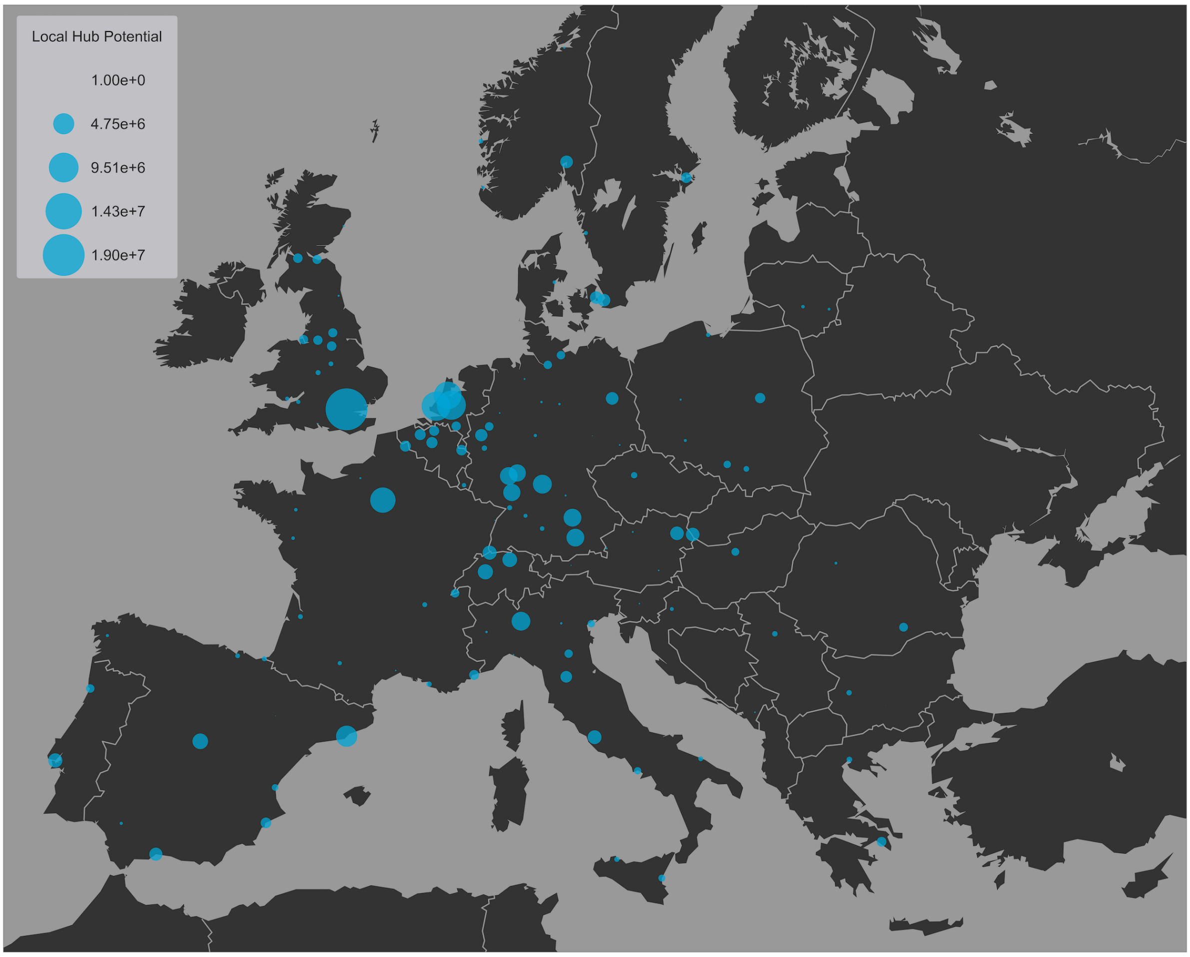


Figure 5.21: Spatial Distribution of the Local Hub Potential within the Air Network

In the case of air, frequencies are independent from urban areas density. As previously explained, air is in fact not as bound to spatio-temporal constraints. However the spatial distribution of the nodes hints that geographical centrality is important also in this case. Most nodes with a considerable node potential are also in this case concentrated in the central areas of the continent. In particular, the most connected urban area is London, followed by the Randstad area (Amsterdam, Rotterdam and Utrecht), Paris and Barcelona. The area of Frankfurt closes the top ten with Milan, Munich, Madrid Zürich and Lisbon. The nodes with the lowest degree are Groningen, Brno, Plovdiv, Ostrava and Saarbrücken. In this case, the poorly served nodes do not seem to be related in any way. It is interesting to note how local hub potential both within the air and rail network is particularly low in Eastern Europe. In regards to national systems, French networks feature similar performances and structure, in both cases centred on Paris, which is the only hub. In contrast, Italy and Germany are characterised by a polycentric network structure. Finally, the UK despite having a rather polycentric structure for rail features a very centralised air system.

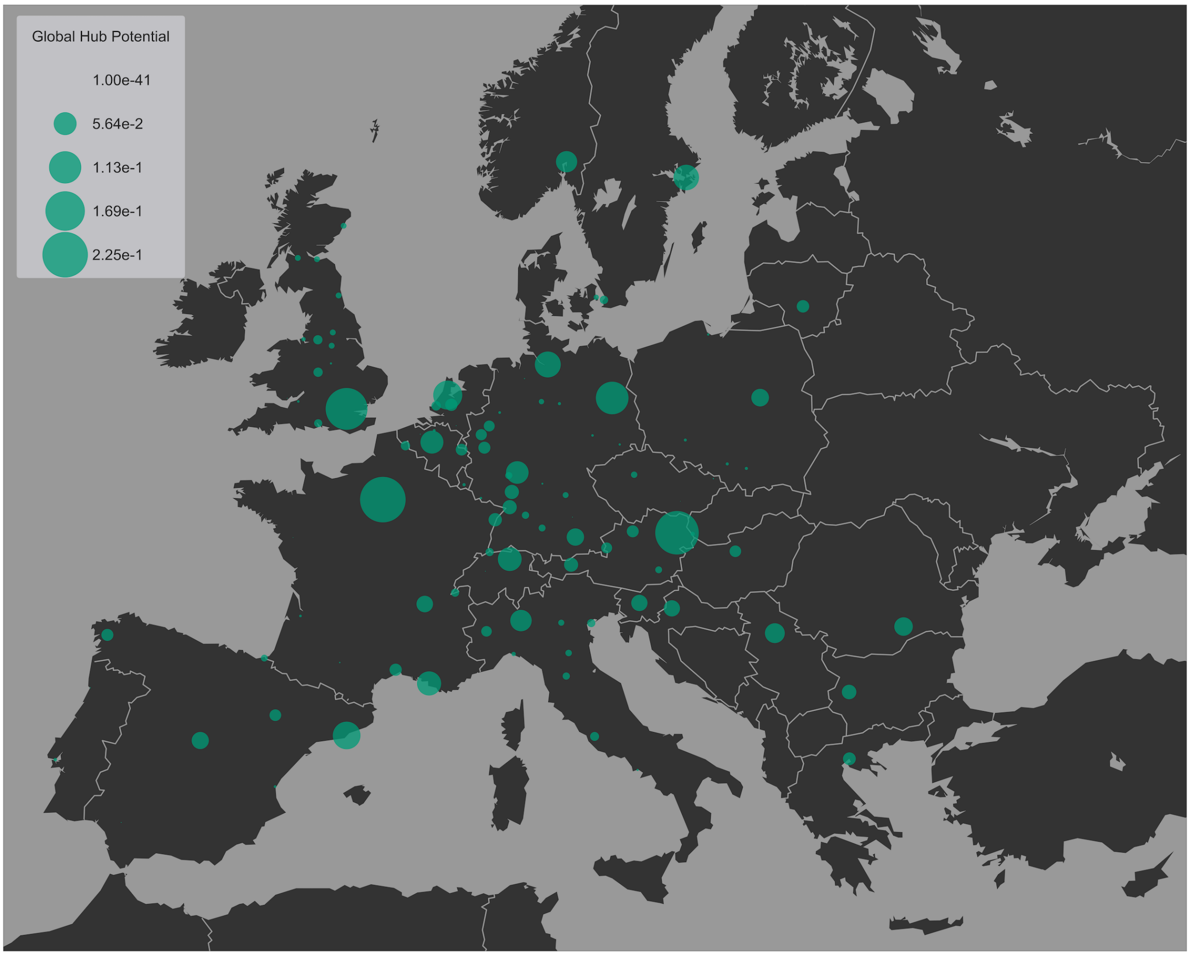


Figure 5.22: Spatial Distribution of the Global Hub Potential within the Rail Network

The spatial distributions show that the global hub connection is more evenly distributed across the continent compared to the previous indicator. In particular, figure 5.22 highlight the main hubs in the rail network, including Paris, Wien, London and Berlin, followed by Amsterdam, Barcelona and Hamburg. It is worth noting how these nodes represent cross-border gateways that connect different European markets. Barcelona is the gateway between Spain and France, Wien and Berlin are the gateways that connect central Europe to Eastern European countries and Hamburg is the gateway that connects the Scandinavian market to the rest of Europe. Finally, the Amsterdam-Paris-London triangle represent a crucial junction of the system receiving the traffic from Northern Europe, Southern Europe and the UK and redistributing it across the entire network. Also in this case, the urban areas with the lowest global hub potential are concentrated in the continental peripheries, including Palermo, Malaga, Vilnius, Poznan, Bergen, Plovdiv and Athens. It is worth noting that local hub potential within the rail network successfully highlights the busiest stations, with most traffic reaching direct connections. However, due to the characteristics of the rail system, more transfers are often required for long-distance transport. In this regard the global hub potential highlight that stations with a high local hub role are not necessarily the ones with a more significant hub potential on long-distance travel.

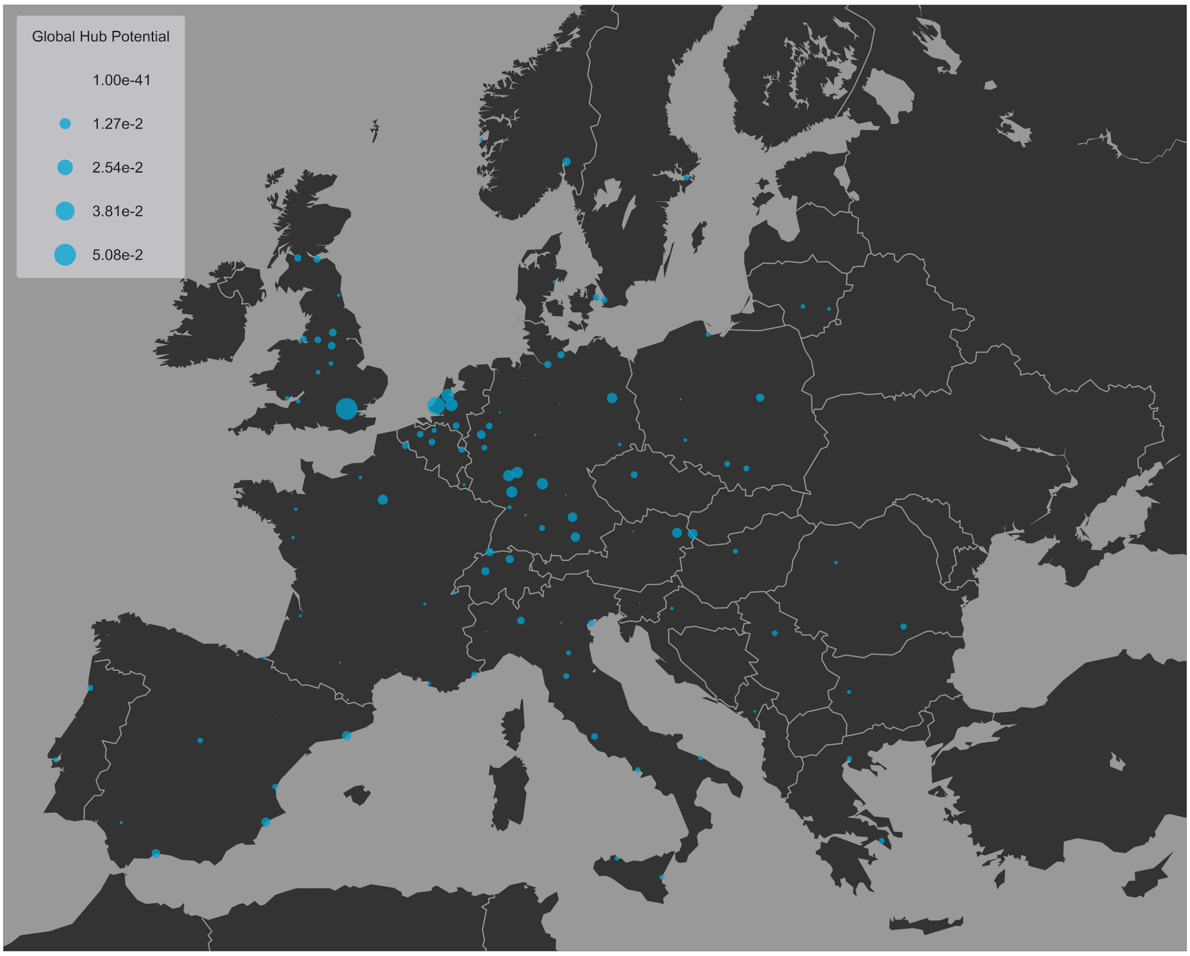


Figure 5.23: Spatial Distribution of the Global Hub Potential within the Air Network

The global hub potential in the air market, as illustrated by Figure 5.23, features considerably lower values compared to the case of rail. This is due to the higher amount of direct connections found in the air network, and consequently to the reduced amount of transfers required to connect all OD pairs. As for the local hub potential, London is the best scoring urban area, followed by the Randstad, the Frankfurt area, Berlin, Paris, Wien/Bratislava and Munich. It seem interesting to highlight that Paris and Barcelona have a more important role at the local level, whilst Frankfurt and Wien have a more significant role considering the global indicator. This hints that the former two might represent crucial hubs in relating neighbouring nodes, whilst the latter might be more attractive transfer hubs for less connected cities. Considering the assumptions made in the creation of the network, especially in terms of terminal aggregation, it is important to highlight that the results should be read considering the structure of the airport systems within each urban area. In particular, it is worth noting that London and Paris feature a polycentric airport system, as opposed to the centralised airport system that characterise the Randstad and Frankfurt. Thus, it could be argued that the hub potential is overestimated for the former two and underestimated for the latter two. In fact, transfers with terminal change for European trips are deemed to be highly unattractive for passengers. The results in terms of urban areas with the lowest global hub potential reflect the results the the local level.

In conclusion, this analysis suggests that Amsterdam and Frankfurt are the two main European airport hubs, closely followed by London and Paris. Within the rail network, Wien, Paris, London, Amsterdam and Berlin appear to be the crucial nodes in terms of long-distance hub potential. Finally, taking into account the limited average number of transfers in the air network the local hub potential is deemed to be a more accurate indicator to estimate the hub potential in the air network. In contrast, due to the higher average path length of the rail network, the global hub potential indicator is considered to better capture the long-distance hub potential of the rail network nodes.

5.4.2. Connection Directness

Local connection directness represent the number of available direct connections. It is possible to notice how in the rail network this number is heavily dependent on the urban area density. Thus, it is worth noting that the node selection process has a crucial impact on this indicator. The urban area with the highest number of rail direct connections is Wien with 45, followed by Berlin (40), Paris (36), Frankfurt (34), Munich (33) and Zürich (30). It is particularly interesting to note the high number of destinations offered in Austrian cities, especially in proportion to their limited city size. For instance

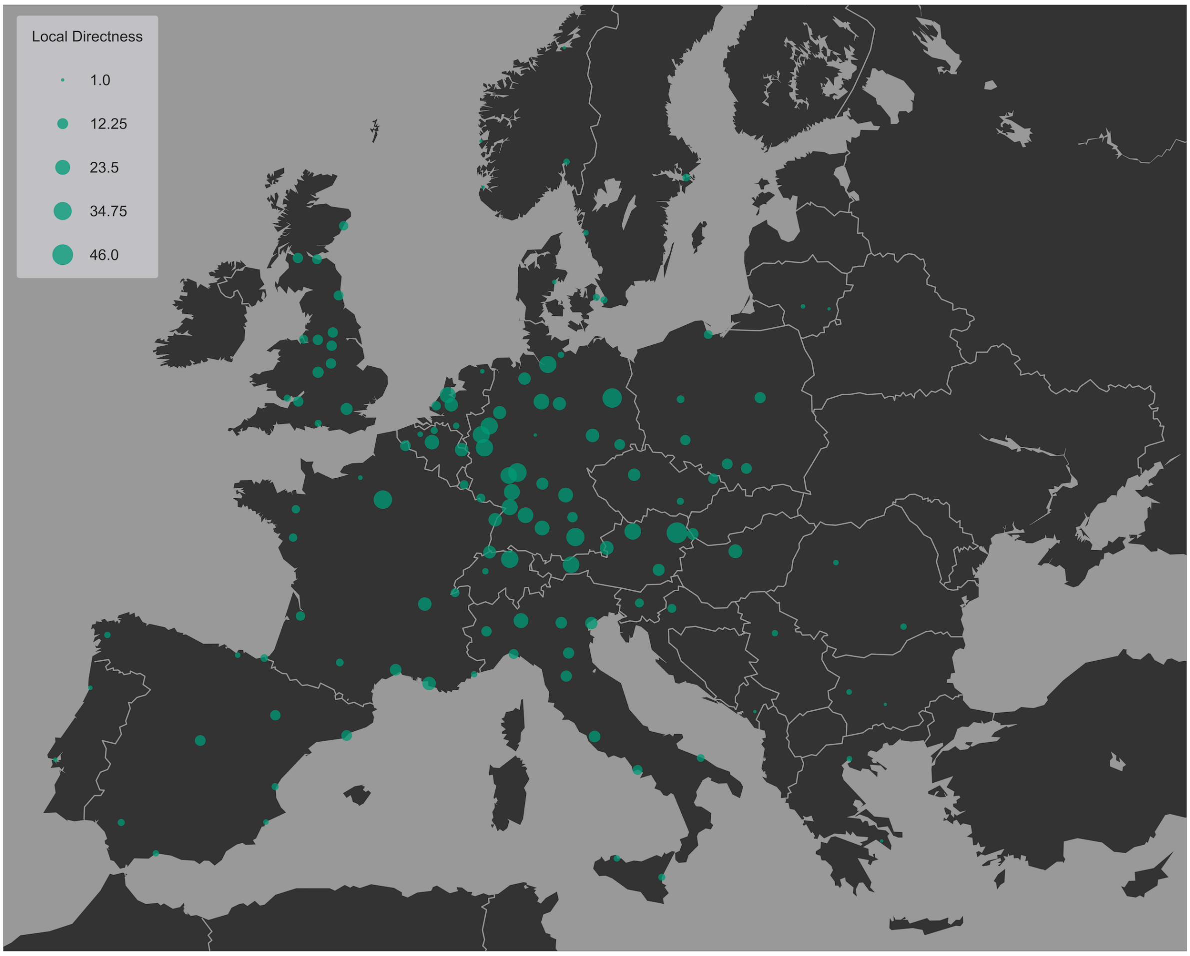


Figure 5.24: Spatial Distribution of the Local Connection Directness within the Rail Network

Linz offers 29 direct connections, Innsbruck 28 and Salzburg 20. This suggest that the Austrian rail system favours national and international coverage, even more so than the German one. In contrast, France and Spain feature very centralised systems where traffic tends to concentrate in the centres (Paris and Madrid respectively) from where it is redistributed along the different radial lines. Finally, eastern European countries appear to have a heavily underdeveloped rail service supply, with only a few destinations available at each urban area. The only exceptions are Budapest, Bratislava and Czech and Polish urban areas. Similarly, also Lithuania is almost completely disconnected from the European rail network. In this regard, it seems important to highlight the development of the Rail Baltica infrastructural project, which is going to connect the Baltic countries to Poland and Europe in the foreseeable future, reshaping the rail supply in the easternmost regions of Europe.

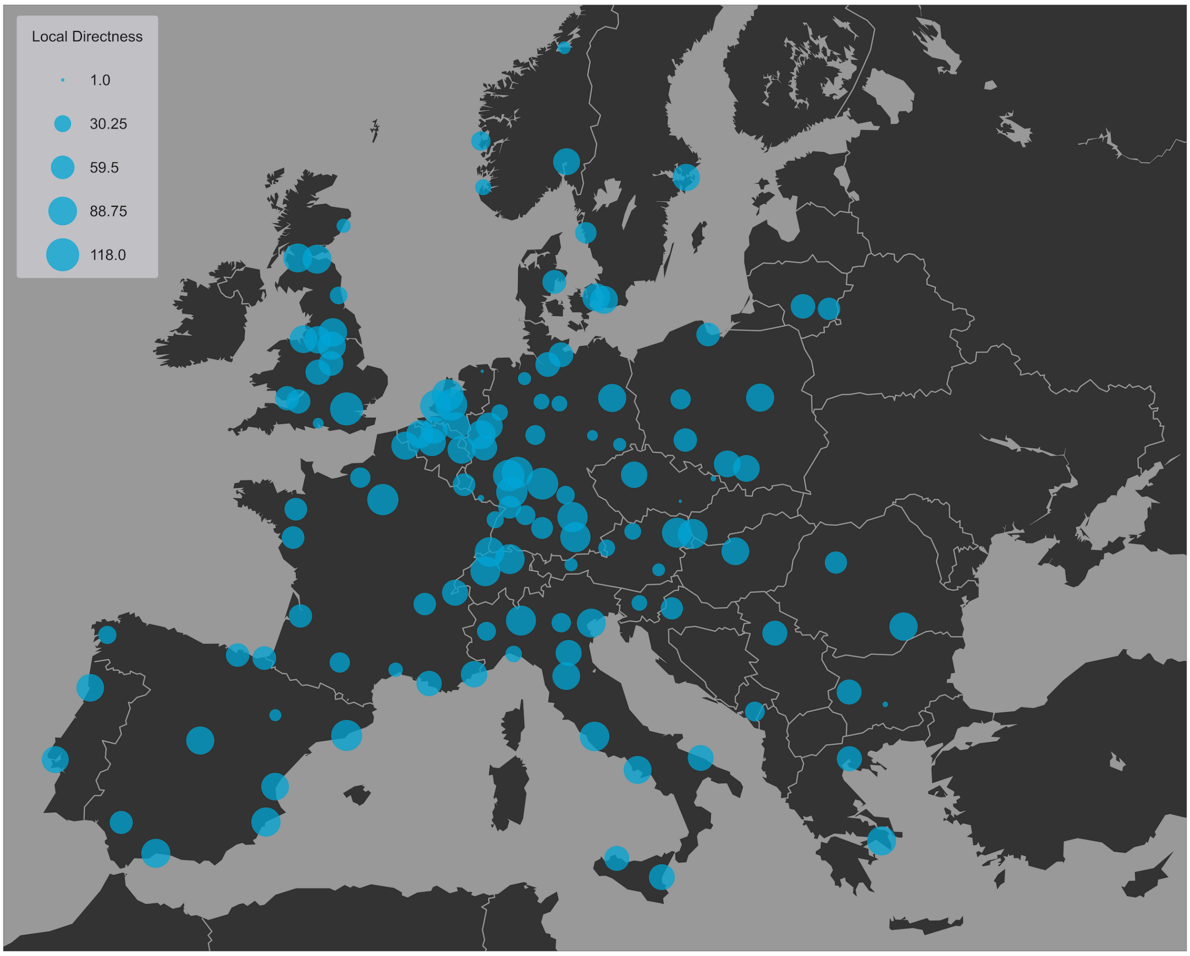


Figure 5.25: Spatial Distribution of the Local Connection Directness within the Air Network

The local connection directness within the air system is much higher, due to the differences in network density. London is the urban area with the largest amount of direct connections (118), followed by the Randstad (110), the Frankfurt area (107), Paris (105), Barcelona (104), Munich (99), Milan (97), Wien/Bratislava (97), Zürich (96) and Rome (94). The number of direct connections is quite evenly distributed across all the nodes, and the hubs, despite topping the chart, are barely noticeable. In this regard, it is worth noting that air allows to offer point-to-point routes with low frequencies on long-distances over which rail cannot compete. This represents a core reason for the flattened distribution.

Furthermore, it is worth noting that only a very limited number of nodes offer less than 10 direct connections, namely Saarbrücken (4), Plovdiv (3), Ostrava (3), Brno (1) and Groningen (1). The considerable number of direct connections present across most Eastern European urban areas leads to conclude that there is enough demand to justify the presence of more long-distance rail services in the area. In particular, this scenario suggests that the main factors hindering the presence of rail services in the area most likely reside in the low infrastructural speeds and in the poor quality of service supply.

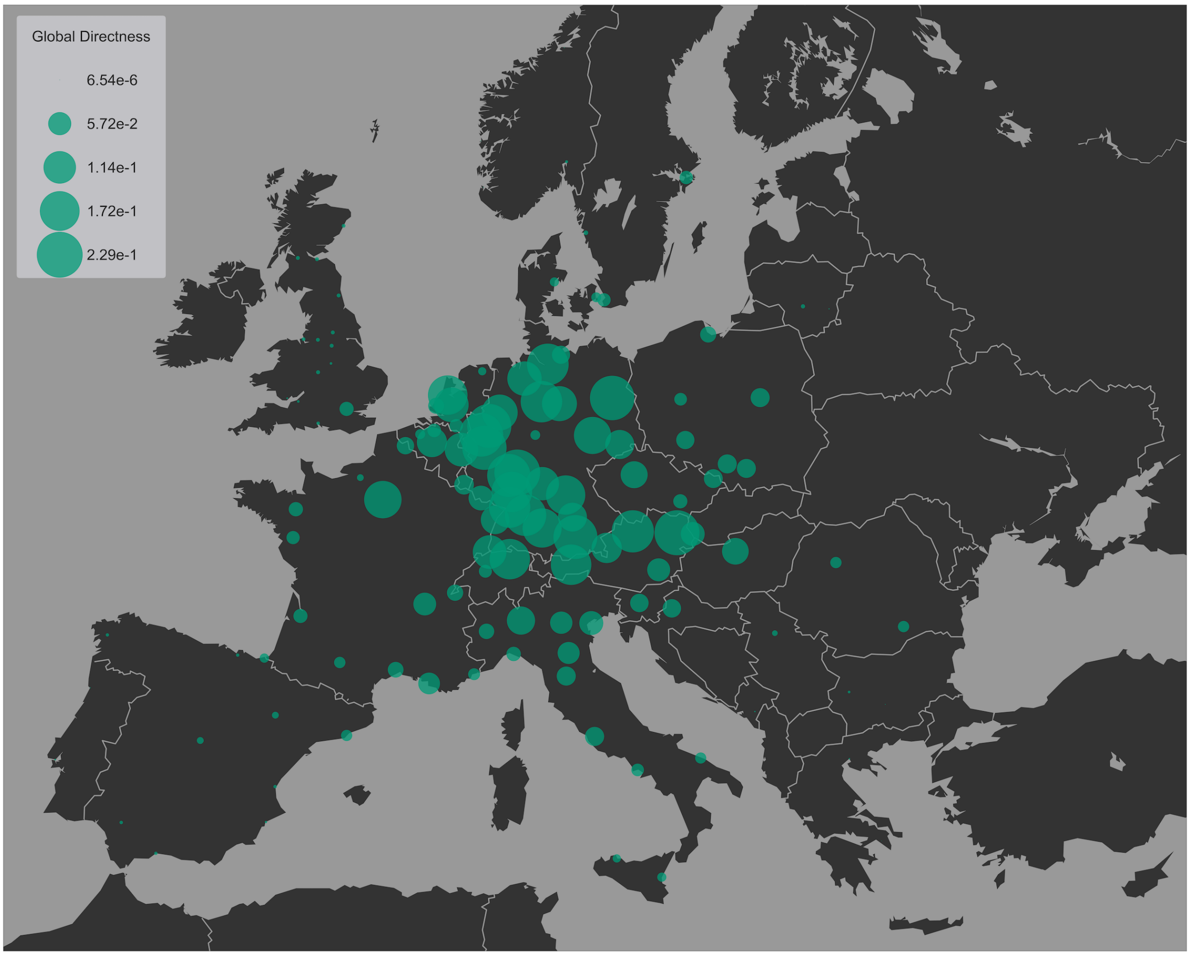


Figure 5.26: Spatial Distribution of the Global Connection Directness within the Rail Network

At the global level, rail connection directness features very heterogeneous patterns. In particular, the spatial distribution highlights the considerable geographical constraints that affect the rail service supply. Geographical centrality is, in fact, directly related to higher connection directness scores. The entire top ten is made up of German and Austrian urban areas, notably Wien, Berlin, Frankfurt, Munich, Linz, Cologne, the Ruhr, Düsseldorf, Mainz and Hamburg compose the top 10. On the other hand, peripheral countries, like Portugal, Spain, Norway, the UK and Eastern Europe consistently rank as the urban areas with lowest degree of connection directness. In particular, Vilnius, Athens, Plovdiv, Trondheim, Bergen, Stavanger and Porto for the bottom of the list. This implies that there is an important imbalance in the share of connections where rail can substitute air services between central European countries and more peripheral states. In fact, whilst in central areas rail allows to reach a large number of destination with a few transfers, this is not true for more external territories.

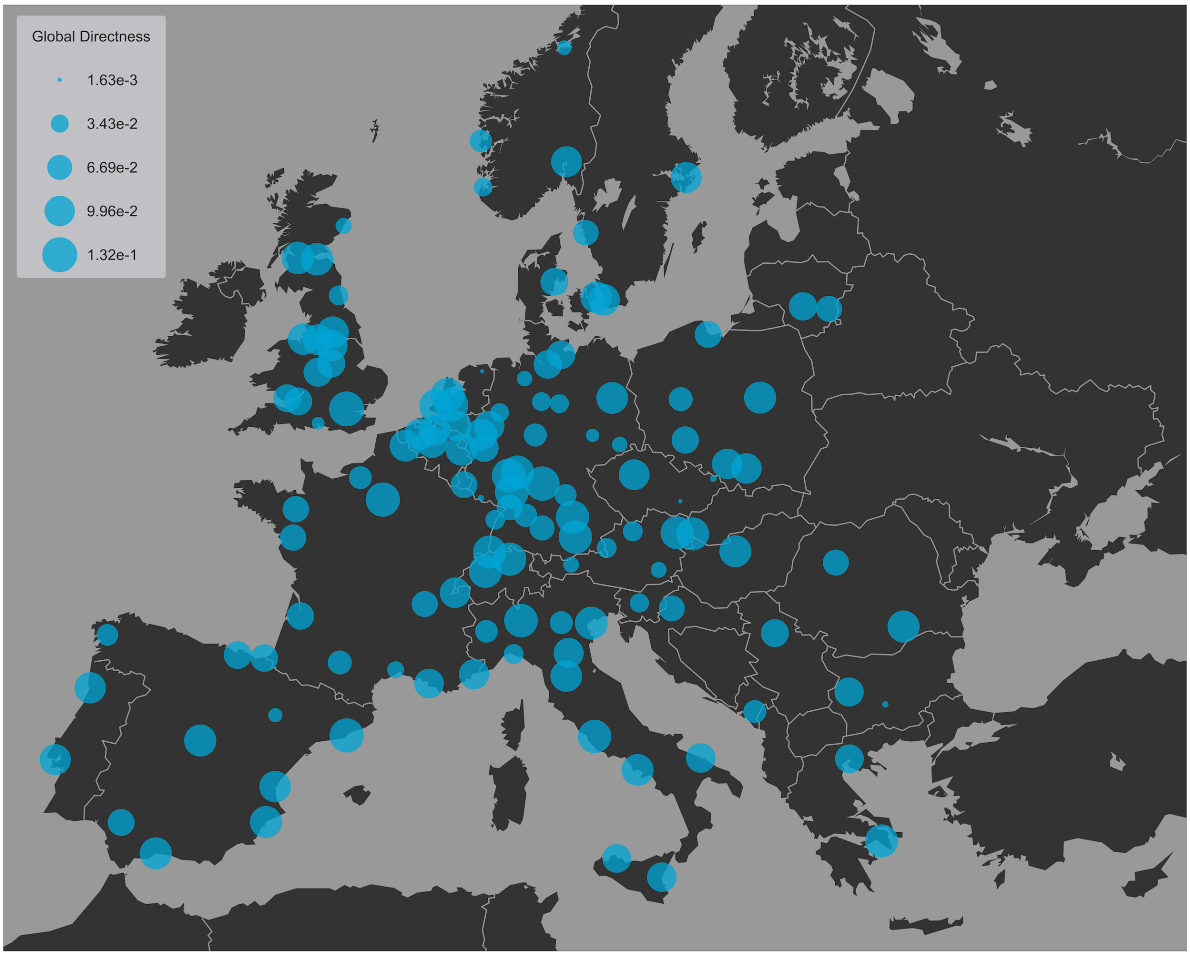


Figure 5.27: Spatial Distribution of the Global Connection Directness within the Air Network

Finally, London, Frankfurt, Barcelona, Paris, the Randstad, Milan, Rome, Munich and Zürich are the top scoring urban areas for global connection directness within the air network. Despite these reflect the results of the indicator at the local level, it is possible to notice how Frankfurt, Barcelona, Milan and Rome have gained importance, whilst the Randstad and Munich have dropped in the rankings. This suggest that the former group is connected on average to more important nodes (i.e. nodes with more direct connections) compared to the latter group. It is, thus, plausible to imagine that Amsterdam and Munich airport provide access to some smaller airports in the network, that are connected to few other airports. In summary, the patterns found for this indicator closely reflect the values of the local connection directness. This, once more, highlights that first degree connections and local indicators in general are already a pretty accurate measure of the connectivity for the air network, with global metrics generally reflecting such values. On the other hand, rail networks require to be analysed both at a local and global network to have a complete overview on the characterising patterns.

5.4.3. Node Connectivity

This third indicator aims to provide a broad overview on the node connectivity, merging together all the different elements including frequencies (as employed within the hub potential indicator), the number of connections (as employed within the connection directness indicator), and travel times. In both the air and rail networks, rather similar patterns can be noticed among the local and the global level. The spatial distributions of local node connectivity are provided by Figures B.1 and B.2, whilst the spatial distributions of global node connectivity are illustrated by Figures 5.28 and 5.29. The main differences are found in the rail network, and regard border transit cities, such as Cologne and Liège, which feature relatively higher rankings at the global level. These cities, in fact, compensate the lower amount of direct connections compared to the most connected urban areas (e.g. Frankfurt) with the ease of accessing neighbouring national systems. Similar patterns are noticeable across the spatial distribution of both modes. In particular, node connectivity within the rail network appears to be influenced by urban area density as areas with higher concentration of urban areas show higher values for node connectivity. In contrast, it is apparent that airport size replaces urban area density as the main influence factor in determining the degree of node connectivity. Despite also in this case high figures are clustered in central Europe, important differences are noticeable across individual nodes within each country, notably in Germany and Austria. In this regard, it is important to note than in the air network multiple cities access the same airports, so the data needs to be analysed taking into account the unique terminals. Some notable examples are the Randstad formed by Amsterdam, Rotterdam and Utrecht served by Amsterdam Schiphol Airport, the Swiss conurbation formed by Bern, Basel and Zürich and served by Zürich Airport and the Frankfurt area, formed by Frankfurt, Mainz, Würzburg and Mannheim served by Frankfurt Airport. Another interesting finding relates to the generalised poor connectivity of eastern

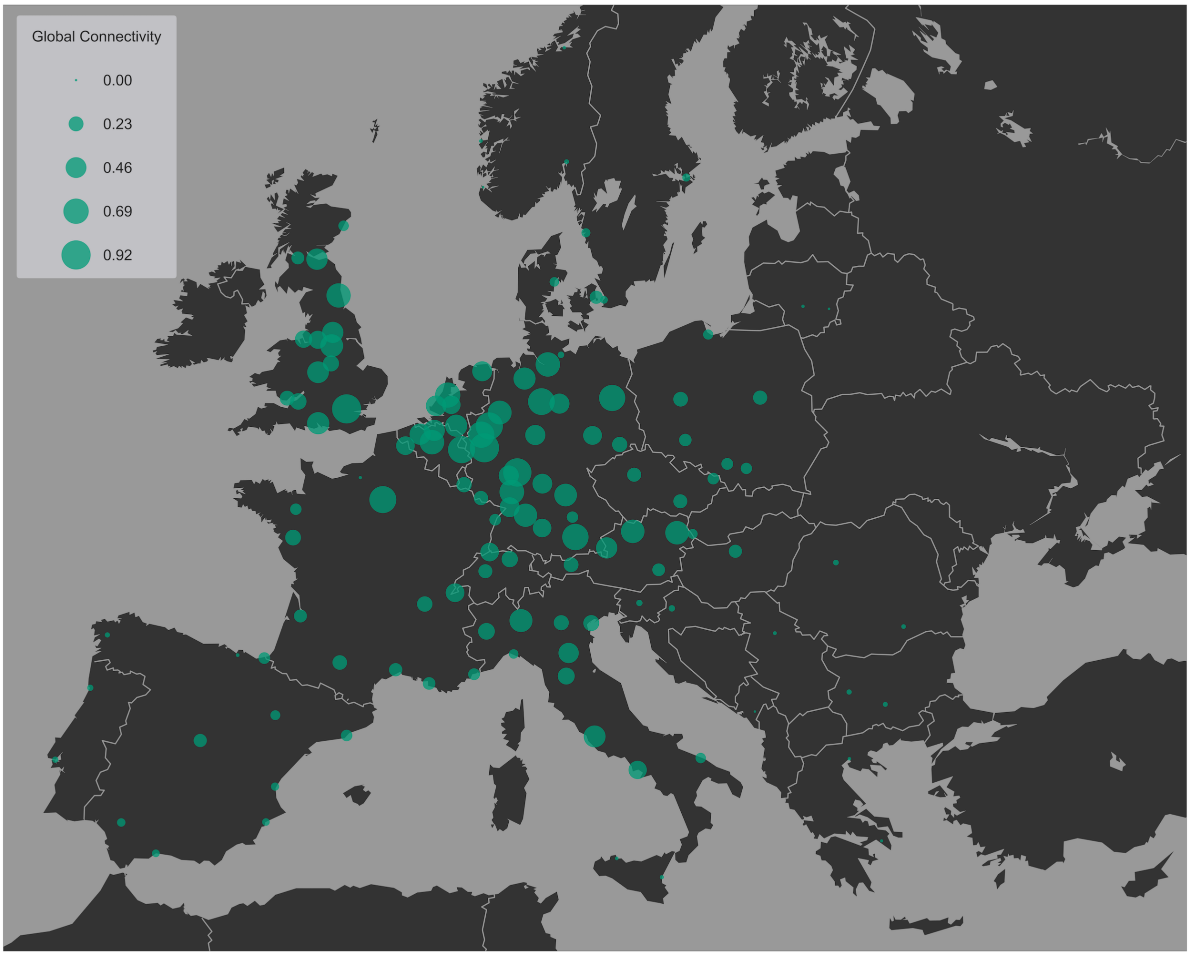


Figure 5.28: Spatial Distribution of the Global Node Connectivity within the Rail Network

Europe. An important distinction must be made between Poland, Czech Republic and Hungary and the remaining eastern European countries, as the former group of countries features considerably higher degrees of node connectivity in the rail network. What is particularly surprising is that this is not only limited to the rail sector but refers also to the air network. In fact, eastern European countries despite featuring numbers of direct connections similar to the figures of western Europe, as highlighted by Figure 5.25, show consistently low degrees of node connectivity. Despite the main reason for this probably relates to the more limited demand for long-distance transport and the absence of important airport hubs, which also suggests that geographical location might still play a role also for air connectivity.

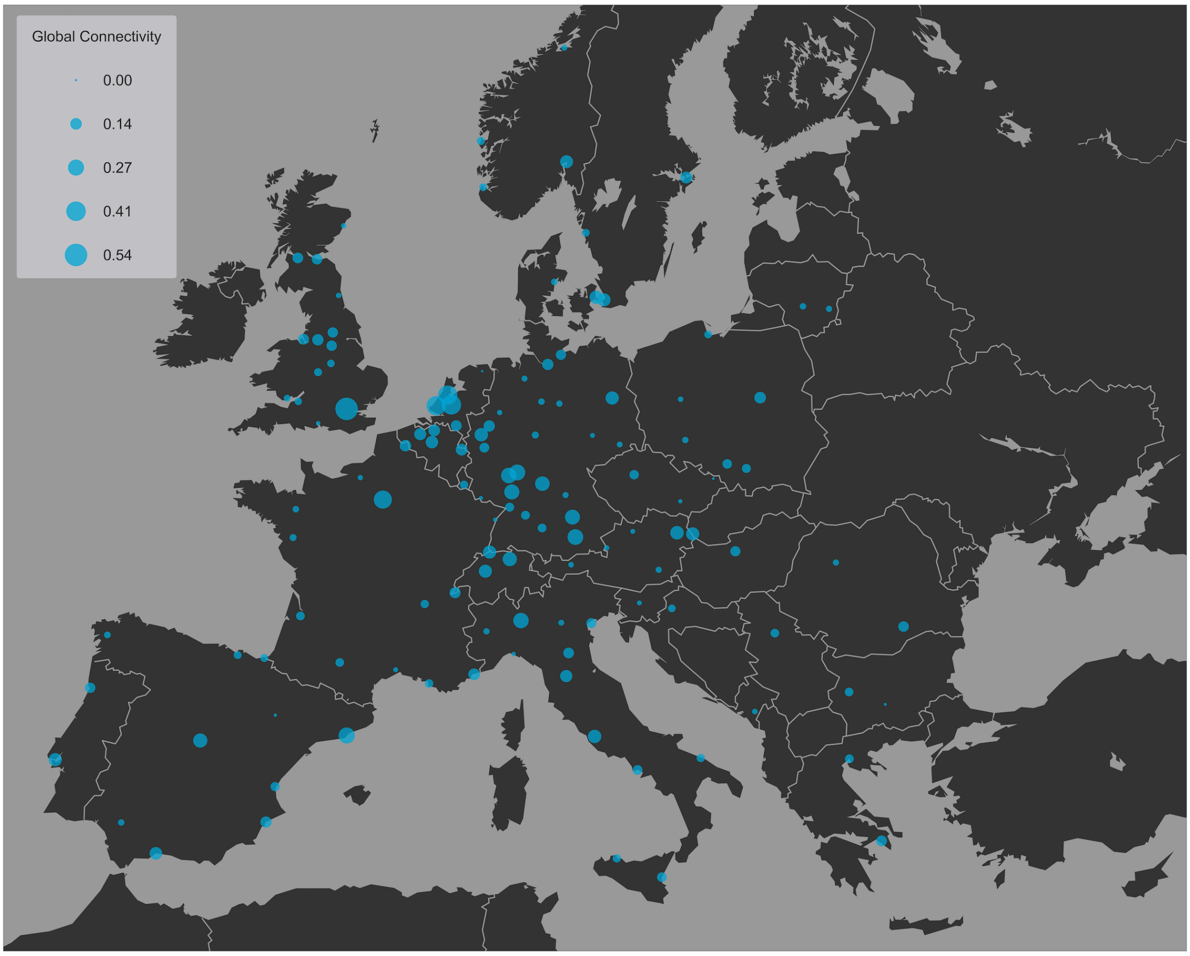


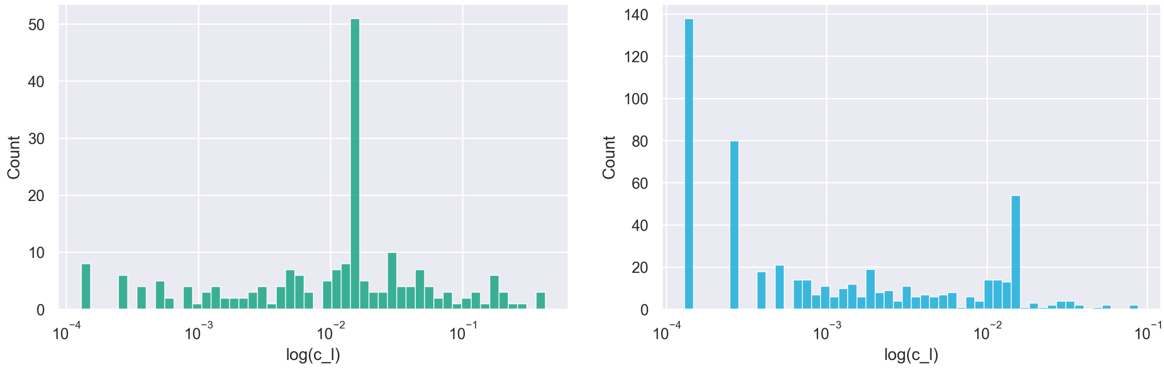
Figure 5.29: Spatial Distribution of the Global Node Connectivity within the Air Network

The node connectivity spatial distribution, however, is particularly useful to identify the specific transport network characteristics of each national system. In regards to rail, most countries appear to have a decentralised and polycentric structure, with low heterogeneity in connectivity degrees across most nodes. In particular, Germany, the Netherlands, Belgium, Austria, Poland, Czech Republic and the UK feature these characteristics. It is worth noting that all these countries feature generally high urban areas densities and have an important focus on standard train services. In contrast, France and Spain are characterised by a centralised and monocentric structure, with a radial infrastructure topology. Furthermore, both countries have a focus on high-speed and have lower densities of urban areas. An interesting case is Italy. In fact, whilst sharing the decentralised characteristics of the former group of countries, its focus is mostly on high-speed rather than standard rail services. However, the linear shape of the country allows to equally serve all urban areas located on the north-south axis. In terms of the air network national air network systems generally appear to be very centralised across all Europe with each country featuring a few hubs. The most polycentric countries are Germany and Italy, whilst France, Spain, Austria are very centralised systems.

5.5. Network-wide Components' Significance

Throughout this section the most significant components in the network are identified. In particular, the significance is assessed using betweenness centrality indicators, following the methodology delineated in Section 3.3.4.

Thus, significant components are defined within this study as the links and nodes that are most often traversed by shortest path in the network. They are defined as elements with network-wide significance, because they are crucial components that allow the seamless and smooth connectivity of the



(a) Air Network

(b) Rail Network

Figure 5.30: Network-wide Connections' Significance Distribution

entire network. Furthermore, the consistent pressure of indirect traffic transiting through these components makes them particularly vulnerable. Thus, making sure that these components are always on point is particularly important to ensure the reliability of the entire network. These are also nodes that are particularly sensitive to disruptions and attacks, and that, as such, should be preserved as far as

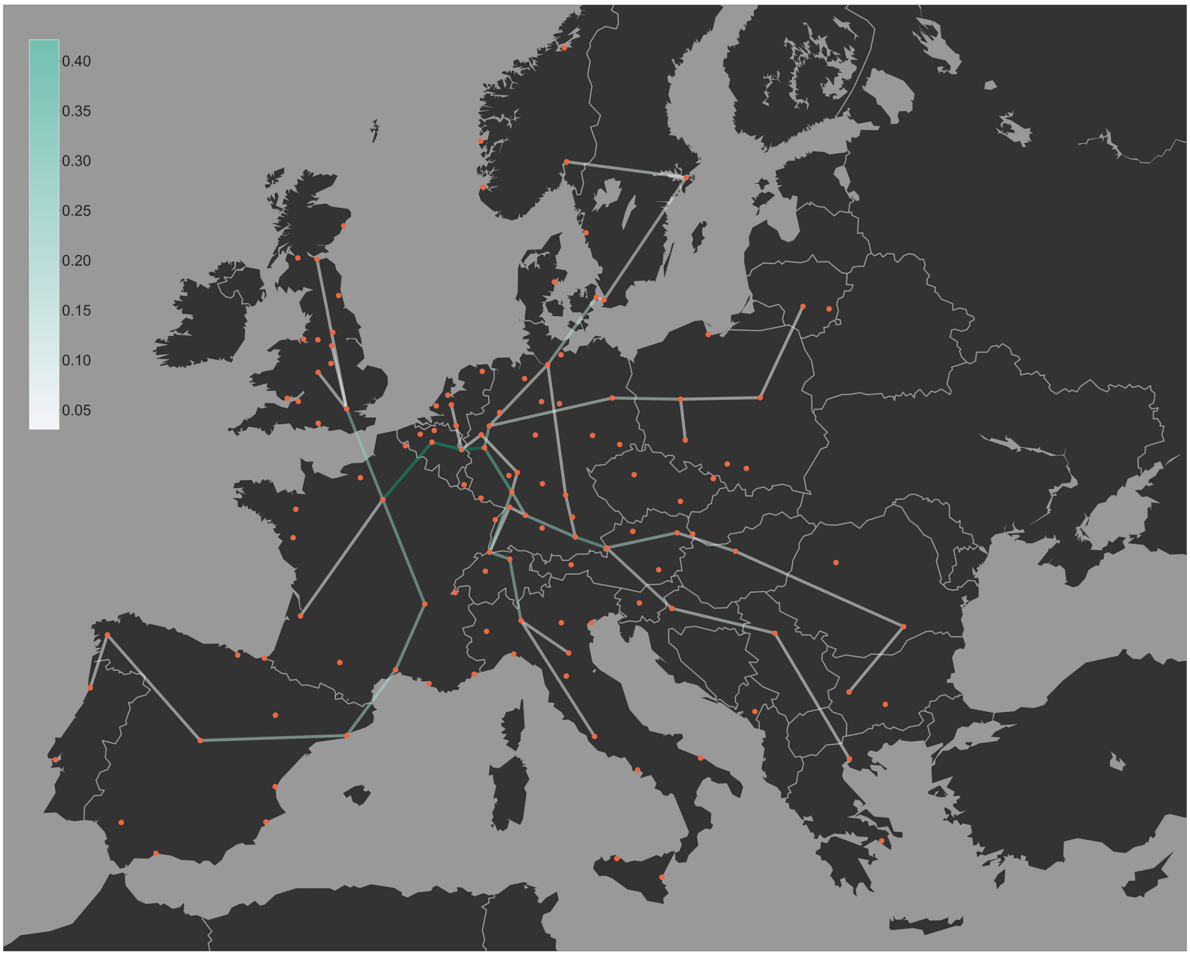


Figure 5.31: Spatial Distribution of the 50 Most Significant Connections within the Rail Network

possible.

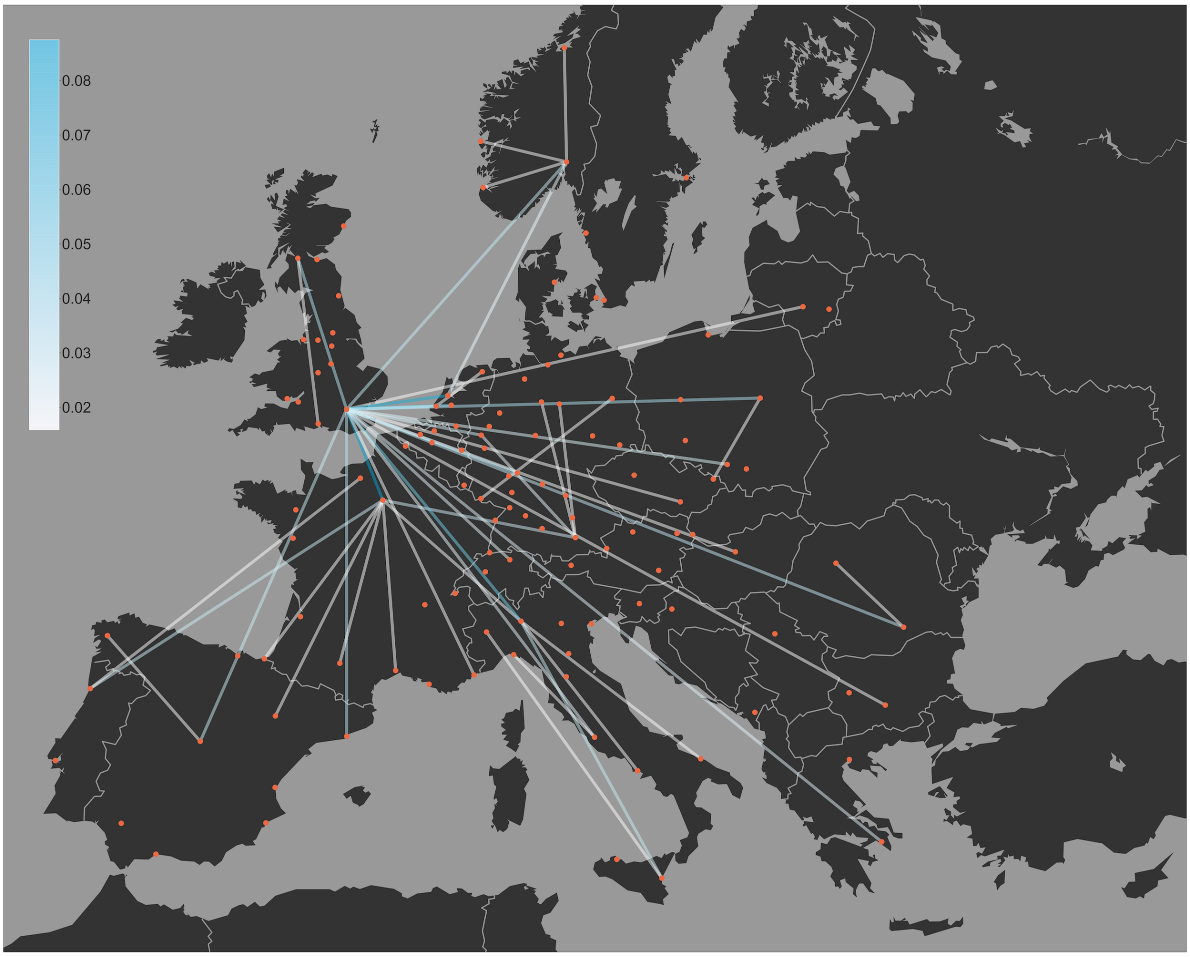


Figure 5.32: Spatial Distribution of the 50 Most Significant Connections within the Air Network

The distribution of network-wide significance across the OD pairs in the rail and air networks is provided in Figures 5.30a and 5.30b. The two histograms show how rail connections generally tend to feature higher degrees of significance. The maximum within the rail network is, in fact, almost five times that the maximum network-wide significance for air, the former standing at 0,42 and the latter at 0,09. This is due to the fact that, as explained in Section 5.1, the rail network is less dense, meaning that it features a lower number of direct connections and consequently higher average shortest path lengths. This implies that a larger number of indirect connections is required to connect all OD pairs. Consequently, links and nodes in the rail network are more often traversed by shortest paths leading to higher values of betweenness centrality. In terms of distribution, it is possible to notice that most connections of both networks feature very low degrees of significance, despite the difference in terms of magnitude of these low figures. The logarithmic scale, in fact, shows how the distribution of rail values concentrates at values considerably higher compared to air. Whilst the figures within the air network concentrate at values lower than 0,02, in the rail systems the network-wide significance of connections reaches values of 0,4.

The top 50 links with the highest network-wide significance within the rail and air network are shown

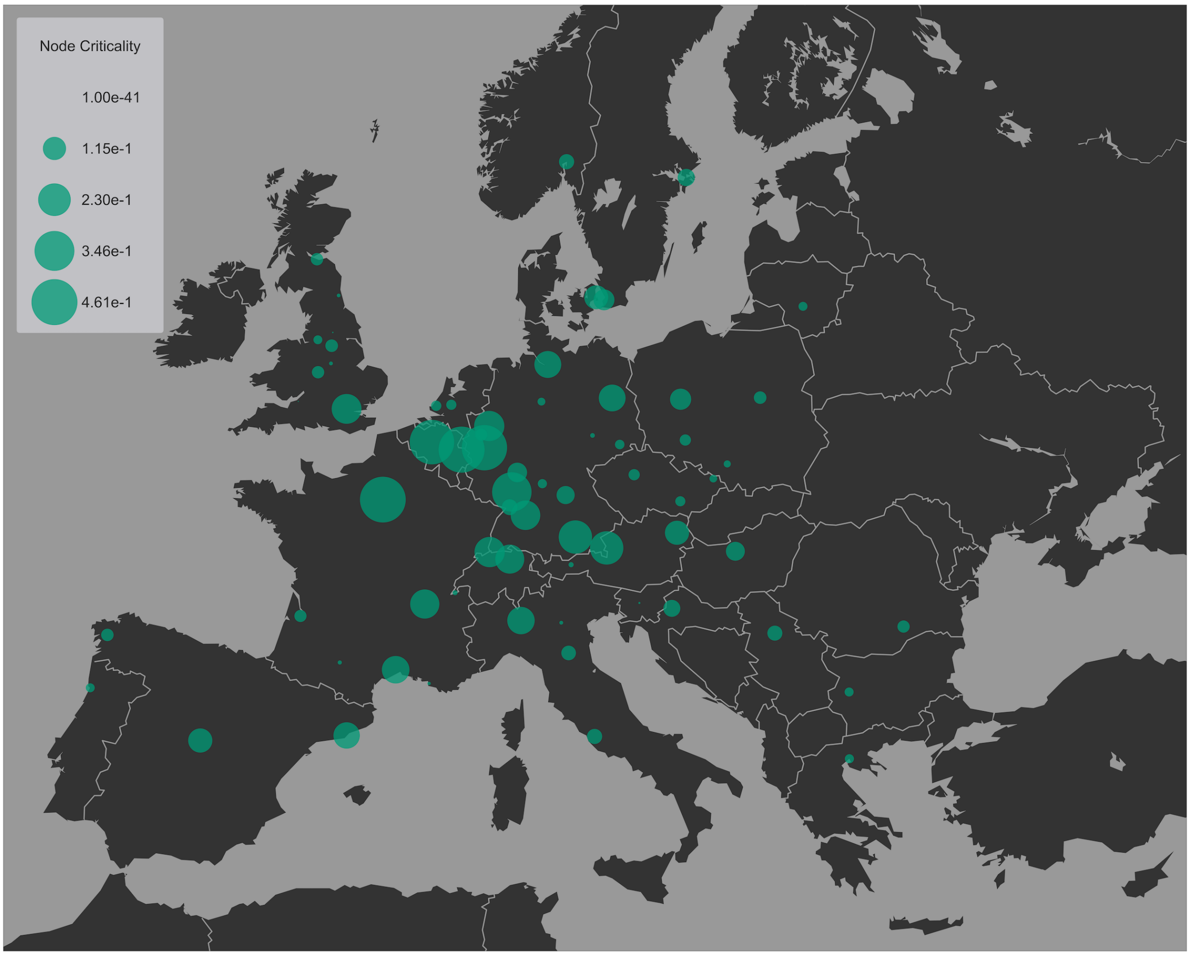


Figure 5.33: Spatial Distribution of the Network-wide Node Significance within the Rail Network

in figures 5.31 and 5.32 respectively. An interesting finding is that the 50 most significant links within the rail network construct a continuous network which represents the current backbone of European rail traffic, reflecting particularly well its infrastructural layer. In fact, the highlighted links, despite representing direct connections, almost perfectly align with the main infrastructural corridors within the network. The spatial distribution of the top 50 most critical links within the air network, highlights the crucial role of Paris and London as transfer hubs. It is worth noting that despite they appear to be more often traversed by shortest paths compared to Frankfurt and Amsterdam, this might not be true in reality. As previously mentioned, London and Paris are characterised by a polycentric airport system. This model simplifying reality by aggregating all the airports into a single node does not account for the reduced attractiveness of changing airport for transferring. Thus, it might be argued that Amsterdam and Frankfurt, being characterised by a single-airport system, could represent more attractive transfer hubs for passengers than shown by the results of this analysis. Similar patterns are noticeable also at the node level, as illustrated by figure 5.34.

In terms of network-wide node significance, it is interesting to notice the crucial role of Belgium within the entire European rail system. This is noticeable both at node and link level and is probably related to its central location within the network, connecting UK and France to the Netherlands and Germany. It is also worth noting that urban areas close to borders, connecting and relating different national

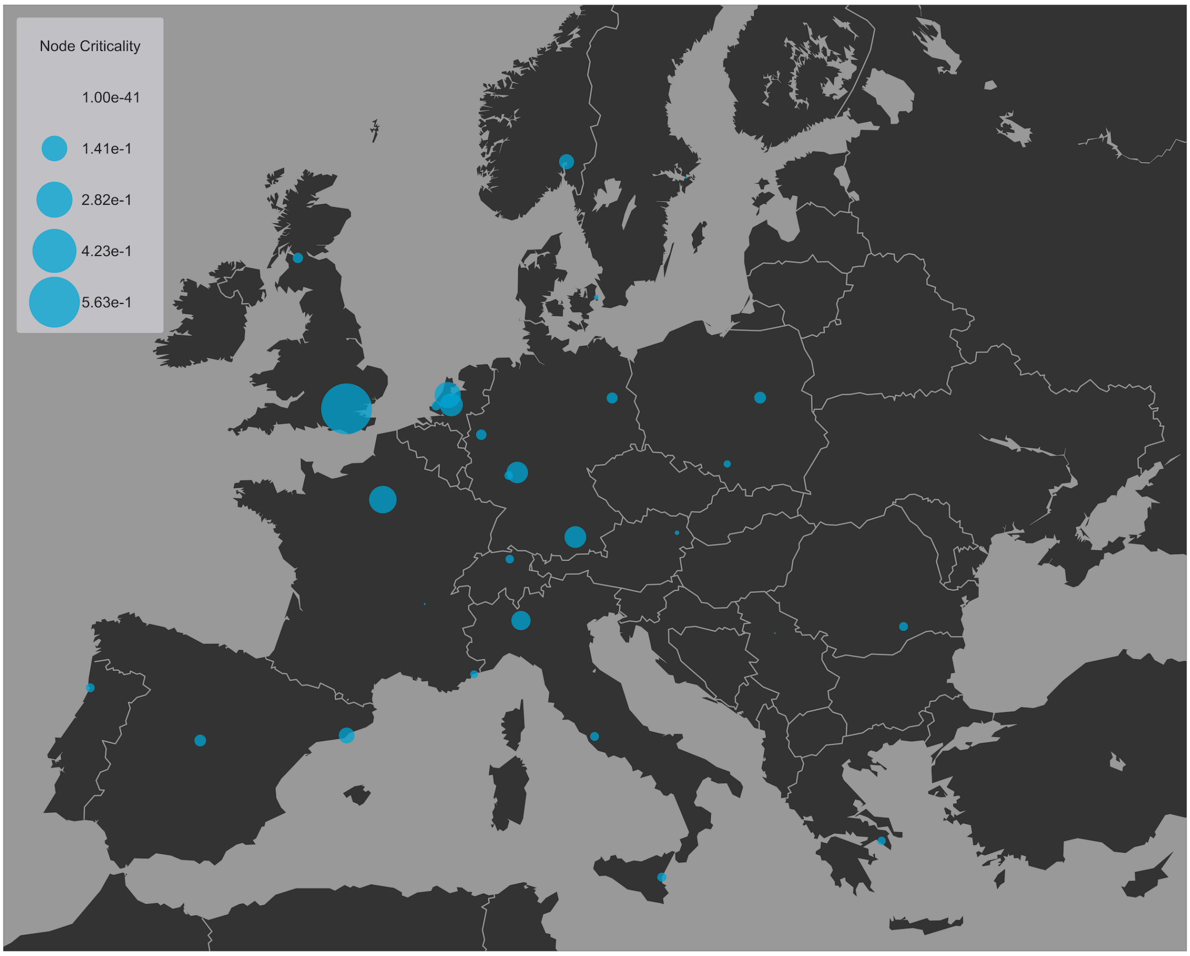


Figure 5.34: Spatial Distribution of the Network-wide Node Significance within the Air Network

systems, generally feature higher betweenness centrality degrees. Air features on average considerably lower significance values, due to the reduced amount of indirect links. However, major hubs, such as London, Paris, Amsterdam, Frankfurt and Munich, feature particularly high network-wide significance scores.

5.6. Summary of the Key Findings

Throughout Section 5.1 the network structure has been analysed and the topological properties have been identified. Comparing the results across the two modes pointed out an important difference regarding the density of the networks, which turn out to be considerably higher for the air network. In particular, within the air network more than half of the possible connections are served by a direct connection, compared to a mere 10% for rail. This implies that the air network consistently features shorter average path length figures, consequently requiring less transfers for passengers to reach any other node in the network. In particular, double the amount of transfers are required on average using the rail network. Furthermore, the fact that 65% of the routes within two transfers are served by rail highlights the crucial role of transfers in the current rail network. In terms of networks properties, both networks feature small world characteristics, despite more statistical tests required are required to fully explore this possibility. On the other hand, it has been excluded that the networks have scale free characteristics, as the degree distributions do not follow a power law distribution. These results are in line with the findings of many studies across the literature. However, it is important to highlight the limitations that this study poses in this regard. In fact, constraining the number of urban areas and aggregating the data at the urban area level, considerably impacts the structure and characteristics of the network. In particular, it is worth noting that these assumptions might not provide an accurate overview on the complete dynamics of the two networks. It is, thus, suggested to further analyse these properties specifically for each mode, considering the networks with terminals rather than urban areas as nodes. Another interesting finding, relates to the differences between the current performances of the rail network in Europe and China. In fact the Chinese rail network features shorter average path length and and lower clustering coefficient compared to the European rail network. This suggests that the former generally focuses on offering long-distance direct connections, whilst the latter more thoroughly serves connections on short-medium ranges, offering higher capillarity and coverage.

The results show that for both the air and rail networks the strongest connection in terms of frequency is a national route. This clearly implies that demand tend to be higher on national routes. At the same time, this also highlights that even at the national level rail is not always competitive enough to substitute air services. In particular, the London-Glasgow corridor is the direct connection with the highest frequency in the air network. The specific characteristics of this corridor, such as the limited OD distance (i.e. just over 500 km length) and the presence of a direct rail service, suggest that rail should be able to substitute air. The failed substitution on such a busy air link highlights the crucial role of high-speed rail in allowing rail to compete effectively with short haul flights and possibly to substitute them. The findings, in terms of travel time and frequency performances, converge with the considerations made in relation to the network structure suggesting that the current rail service supply focuses on shorter travel time connections. In fact, frequency tend to increase either linearly or even exponentially with lower travel times. In particular, an important finding relates to the overwhelming superiority of rail frequencies on routes with travel time between 100 and 300 minutes. Based on this, the maximum distance where rail could theoretically compete with air has been estimated to vary from 800 km, with an average cruise speed of 200 km/h, to 1200 km, for average cruise speeds of 300 km/h. However, this study has shown that the current supply of direct rail connections is concentrated on OD distances under 500 km, market which accounts for almost three quarters of all the direct connections provided by rail. Despite the density of the air network is five times the one of rail considering the entire network, this considerably changes when taking into account only connections ranging from 100 to 500 km, where the air-rail route supply are rather balanced. Overall, the results show that the figures within this threshold are rather similar across the two modes, as the OD pairs are quite evenly distributed between inter-modal competition, rail monopoly and air monopoly. However, despite supplying a competitive number of direct routes under 500 km, rail appears to be overwhelmingly behind on longer ranges. Direct routes longer than 500 km are in fact almost completely monopolised by air. In particular, the findings highlight that within the 500 - 1.000 km range, where rail is supposedly capable to compete with air, almost 90% of the supplied direct connections is an air monopoly.

Similar patterns are also found in terms of connectivity both at the link and node level. Link connectivity within the rail network starts plummeting for OD distances greater than 500 km, whereas in the air network it appears to be independent from the specific geodesic distance of the OD pair. At the node level, findings clearly show that air is generally independent from geographical constraints and political boundaries, whilst rail heavily relies on infrastructure and the patterns in the spatial distribution of supply seem to suggest that a single railway area is not in place yet. The local hub potential indicator highlights that areas with higher density of major urban centres generally correlate to higher train service frequencies. Furthermore, the global connection directness indicator underlines that the geographical location of a node, and especially its centrality, represent particularly important factors in determining the connectivity of a node. Finally, the low degrees of node connectivity in eastern Europe suggest that these areas might feature lower travel demand compared to western Europe, which is in line with the assumption that travel demand is a function of economic factors like GDP. Moreover, this also hints at the fact that geographical location might also play a role in the connectivity of nodes within the air network, especially due to the hub and spoke structures. Within the rail network global hub potential highlights the main cross-border gateways connecting Europe, whilst local hub potential underlines the most important urban areas in terms of local traffic. Thus, the fact that an urban area is a major European rail hub do not necessarily imply that it has the same importance also at the global level. It is worth noting that in the case of air there are more limited differences between local and global indicators due to the considerable number of direct connections and minimal average shortest path. At the link level, it is interesting to note that air does not provide higher connectivity on any of the routes under rail monopoly. On the other hand, rail actually provides higher connectivity on some of the connections shorter than 500 km under an air monopoly regime, implying that the indirect connections provided by rail in some cases are more attractive compared to the direct service provided by air, whilst the opposite is not true.

In terms of critical components, the most important nodes in the air network are London, Amsterdam, Frankfurt and Paris, despite generally good performances were noticed all over the network, with a few exceptions only. On the other hand, wider variations in terms of performances were noticed within the rail network, with node directness and connectivity being considerably higher in countries of Central Europe, such as Germany, Austria and the Netherlands. It is interesting to note that differently from air, in the rail network only minimal differences are ascribable to the urban area size and importance. The results in relation to the main rail global hubs follow the findings within the air network with the exception of Frankfurt, which appears to have relatively less importance in the rail network. Furthermore, Wien, Berlin, Hamburg, Bruxelles, Barcelona, Zürich and Milan are other important global rail hubs. At the link level, it is found that 79 out of the 954 OD pairs under a competition regime feature very high air frequencies. On 48 of these rail offers better travel time or frequencies (or both), whilst on 31 air outperforms rail. This implies that air manages to serve also connections where rail features better performances, and that there is a number of connections with a considerable demand that rail could capture. In particular, the study highlights that under the current market conditions and considering the supply performances of the two modes rail could already substitute air over 19 OD pairs in the network. Most of them are national routes between the major cities within the country, such as the case of the Milan - Rome and the Milan - Naples in Italy, the Madrid - Barcelona and Madrid - Málaga in Spain, the Lisbon - Porto in Portugal, the Marseille - Paris in France, the London - Edinburgh in the UK, the Munich - Frankfurt in Germany and the Stockholm - Malmö in Sweden. In terms of cross-border connections the London - Paris, the Amsterdam - Frankfurt and the Basel - Frankfurt appear to also feature the conditions for a complete substitution of air services with rail. Furthermore, increasing the frequencies on 16 connections and reducing the travel times on 13 others would allow to establish the aforementioned substitution conditions. Some notable examples of the former category are the London - Glasgow, the London - Amsterdam, the Madrid - Santiago de Compostela, whilst the Paris - San Sebastián, the Milan - Bari and some important connections between major German cities (e.g.

Munich - Berlin, Berlin - Frankfurt, Frankfurt - Hamburg and Hamburg - Munich) fall in the latter. Finally, on 31 connections both travel time and frequency enhancements are required for rail to become an attractive substitute to air services. Many crucial OD pairs fall into this category, interestingly enough most of them are cross-border international connections. Notable examples include the Paris - Milan, Paris - Munich, Paris - Barcelona, Amsterdam - Berlin, Stockholm - Copenhagen, Oslo - Stockholm, Zürich - Amsterdam, Wien - Düsseldorf, Madrid - Lisbon and Milan - Frankfurt. Moreover, 147 OD pairs with distance below 1.500 km that currently are under an air monopoly regime feature high levels of air demand that might justify investments on rail infrastructure and services to capture some of those market shares. The findings have highlighted that these routes, which are deemed attractive for substitution, between 100 and 500 km tend to concentrate between Italy, Switzerland, Germany, the Netherlands and the UK on a north - south axis, with a few exceptions. On the other hand, on the 500 to 1.000 km routes the spatial distribution expands from the aforementioned countries to some of their neighbours, to finally include also more peripheral areas when considering longer routes until 1.500 km. It is worth noting that most of these routes depart from the major European air hub urban areas, such as London, Paris, Amsterdam, Barcelona and Rome. In terms of infrastructure, the findings point out that the main gaps are located on the cross-border connections between Spain and Portugal, France and Spain, among France and Germany/Switzerland, the Netherlands and Germany/Denmark and between Germany and Austria. At the national level, important infrastructural gaps are found in the south of France, within Scandinavia and in the South of Italy. Furthermore also the alp corridors appear to be critical elements from an infrastructural perspective.

6

Conclusion

This study analysed and compared the structural characteristics and properties of the European rail and air networks. In particular, a connectivity indicator was proposed to provide insights into the current state and performances of the two aforementioned networks. The methodology developed within this study allowed to fill the gap identified in Section 1.3 enabling direct comparability between air and rail supply at the link, node and network levels and combining network structure and characteristics with actual performance indicators. In Section 6.1 the main research question and sub-questions presented in Section 1.5 are answered. Furthermore, the implications of the findings for practice and policy-making are highlighted in Section 6.2. Next, the limitations of the current study are discussed and some suggestions for further improvement are provided in Section 6.3. Finally, Section 6.4 provides some recommendations for further research in terms of the future endeavours required to complement the results of this research.

6.1. Research Questions

This Section aims to conclude on the key findings identified and discussed in Section 5.6 by answering the four sub-questions identified in Section 1.5.

“What are the factors determining and influencing the connectivity of rail and air transport networks in relation to modal competitiveness?”

The first question is answered through the tripartite literature review contained in Chapter 2, and especially by Section 2.1.3. The factors determining modal competitiveness and consequently hindering the creation of a level-playing field competition between the two modes have been identified and summarised in the System Dynamics illustrated by Figure 2.1. In particular, the relationship between these factors and connectivity has been discussed in Section 3.1.2. These factors and their relationships were deemed important elements to take into account during the connectivity indicator definition process. In particular, to provide further insights on the matter the factors influencing connectivity have been identified in Section 2.2.4. Interestingly enough, the factors influencing connectivity appear to overlap almost perfectly with the ones that influence modal competitiveness. Thus, the main factors directly influencing modal competitiveness in relation to the connectivity of rail and air services have been identified in the elements listed below:

- *Travel Time*: represents the total duration of the door-to-door trip, including access/egress, waiting, in-vehicle and transfer time and negatively influence modal competitiveness.
- *Frequency (Service Capacity)*: represents the number of services available on the main leg (terminal-to-terminal) within a certain time period (i.e. in the case of this study on a weekly basis) and positively influences modal competitiveness.
- *Travel Comfort*: represents the comfort level of the door-to-door trip, including both the in-vehicle, out-of vehicle and transfer part of the trip. Thus, it is negatively influenced by the number of transfers, and in turn, positively influences modal competitiveness. In particular, literature on travel behaviour and mode choice highlight that passengers' sensitivity to the time spent in transport varies based on the specific part of the trip. Travellers appear to prefer spending time travelling on the vehicle, rather than outside of it waiting or accessing/egressing the terminal, whilst transfers have the highest impact.
- *Travel Cost*: represents the total economic value of the trip, including travel fares and access/egress cost, and has a negative impact on modal competitiveness.
- *Travel Demand*: represents the total demand between each urban area OD pair and positively impacts modal competitiveness.

It is worth noting the absence of distance and velocity in this list. Although both were identified as factors determining connectivity, they were deemed to overlap with travel time, as the in-vehicle time is derived from the ratio of distance to speed. Furthermore, service reliability was considered to be one of the factors influencing travel comfort, and thus indirectly influencing modal competitiveness. It seems important to highlight that ultimately, among the factors highlighted above only the first three are included within the connectivity index. Travel costs are excluded due to their inherently dynamic nature, whilst travel demand was left out because the study and the indicator focus on service supply.

"What are the properties and performances of the European rail and air transport networks, and how do they compare?"

The second question is answered using the findings obtained from Chapter 5, and with particular regard to Section 5.1. As highlighted in Section 5.6, results suggest that the air network is denser compared to the rail network, meaning that a higher number of direct connections is supplied within the former system. This has a twofold set of implications. First of all, the lower number of transfers required in the air network highlights the crucial role of air transport on longer trips, where rail is not able to compete. In this regard, it appears paramount to underline the dimension of cooperation between the two modes, which has often been disregarded by the literature (Román & Martín, 2014; H. Yang & Zhang, 2012). The findings of this study show that currently rail and air transport appear to widely complement each other covering different sections of the market rather than merely competing within the same market section. In fact, despite both networks appear to be tightly connected, featuring small-world characteristics and consequently allowing passengers to reach any destination within a limited number of transfers some important differences are noticeable across the two networks. In particular, rail features especially high clustering coefficient degrees, suggesting that neighbour nodes in the rail network are very compactly connected, with almost all neighbours being connected with each other. On the other hand, air features extremely small average shortest path, which means that all nodes are connected within a minimum amount of steps. These findings support the idea that two modes are currently seen as complementary, with air optimally serving areas located further away (i.e. on OD distances greater than 1.000 km) and rail offering feeder services and redistributing local traffic through its capillary regional accessibility. It is possible to conclude that, considering the current structural characteristics of the two networks, moving forward more focus should be placed on inter-modality. The results, in particular, suggest that further attention should be placed on the possibility of integration between rail and air rather than merely on their competition. In this regard, it is worth noting that the system dynamics presented in Section 2.1.3 indicate the crucial importance of competition within the rail market to maintain the attractiveness of rail in the absence of inter-modal competition.

Secondly, the results highlight that currently, rail offers a limited number of direct connections, suggesting that rail might rely on transfer services rather than on direct connections to serve the long-distance European market. However, following what was discussed in Section 3.1.6, passengers appear to be especially sensitive to transfer time increases. In particular, the system dynamics suggest that the comfort level of transfers (e.g. in terms of transfer times and the quality of waiting areas in stations), appears to have a considerable influence on the modal competitiveness of rail. Thus, to improve the attractiveness of long-distance rail services two main approaches could be considered, namely the increase in transfer quality and the reduction of the number of transfers. In relation to the former, guaranteeing short transfer times by temporally coordinating the schedules of train service could prove to be an effective solution. However, this often appears to be rather complex for rail due to capacity constraints, and especially at an international level, due to the fragmented state of the market in terms of traffic management and capacity allocation. Thus, a different approach to making rail more attractive relates to the possibility to increase the number of direct connections offered. This, reducing the average number of transfers, could help to limit the dependence on schedule coordination. Furthermore, the findings of this study also highlight the crucial importance of competitive perceived door-to-door travel times, suggesting that high-speed services and night trains, due to the lower perceived travel times, might play an important role in increasing rail modal competitiveness. Finally, it is worth noting that rail infrastructural constraints not only represent a barrier but also allow the provision of additional frequencies to minor urban areas located in strategical nodes of the infrastructural network. This suggests that the rail service-infrastructure bond rather than merely hindering the mode could also be employed to make it more attractive, offering capillary coverage and widespread connectivity. In this instance, it is paramount to carefully consider the trade-off between the increase in travel times due to the inclusion of multiple stops and the increase in coverage due to their reduction. A possible solution could be offering both direct and stop services, as done on the Italian high-speed lines. In particular, a thorough evaluation of the demand patterns is required to establish the optimal stop configuration.

Furthermore, comparing the properties and performances of the Chinese and European long-distance rail networks has shown that the former offers more direct connections, relying less on transfers, and making rail more attractive and competitive with air on longer distances. In fact, the European network appears to be still considerably bound to the regional and national dimension, failing to provide effective alternatives to air on longer distances. An important reason for this relates to the lack of a comprehensive continental high-speed network, which hinders the capability of rail to offer competitive travel times on the medium and long range. At the same time, another important factor is the extreme fragmentation of the European rail system, as opposed to the integrated and harmonised Chinese rail market. In particular, Sections 5.3 and 5.4 have highlighted the considerable heterogeneity among different national rail systems in terms of network structure and characteristics. The results, in fact, point out the distinction between decentralised polycentric rail systems and centralised monocentric systems. The former is characteristic of countries such as Germany and Austria, where rail focuses more on offering widespread coverage and high frequencies rather than high speeds and low travel times. On the other hand, in the latter systems, typical of countries like France and Spain, rail mostly aims to offer competitive travel times using high-speed rail. Finally, there are some countries such as the UK and Italy, which feature a mixed system with characteristics from both the aforementioned systems. In particular, it appears that the main drivers shaping and influencing systems' characteristics are the geographical structure of urban regions, the spatial distribution of urban areas and the historical development of the lines. Furthermore, in regards to the former, it is interesting to notice the influence of the density and geographical disposition of the main urban areas.

"What is the connectivity of the European rail and air transport networks, and how do they compare?"

To answer this third question, the term connectivity was first defined and contextualised in Section 2.2. In particular, the literature review has clearly proved that the relevance of a network is inevitably related to its connectivity (Rodrigue, 2020). Connectivity through its degree/strength is, in fact, a measure that allows benchmarking and monitoring rail network performances against competitor modes, identifying critical nodes and links which require particular attention. In doing so, connectivity provides some of the information required to design more careful, accurate and future-proof strategies and policies targeted at improving the competitiveness of rail and at planning the direction of future investments. Thus, to quantitatively define and compare the degree of connectivity of the European rail and air networks, a connectivity measure composed of a connection intensity and of travel inconvenience components, the former capturing the frequency dimension whilst the latter focusing on the travel time aspect, is proposed in Section 3.3.1. Given that the two components are related through a ratio, it was expected that on a linear scale the numerator would have a greater influence on the value of the connectivity metric compared to the denominator. The results have confirmed this, showing higher connectivity figures for rail connections compared to air counterparts, following the greater magnitude of rail service frequencies. This might suggest that rail connections currently have an edge over air counterparts. However, it is worth noting that this gap is further exacerbated by the market segmentation of the two modes. Whilst rail largely serves mostly short connections, with low travel times and high frequencies, air focuses more on longer distances, with the naturally lower frequencies and higher travel times that characterise this section of the market. This finding highlights some of the complexities related to comparing the system-wide performances of modes with different core characteristics, such as rail and air.

The results in terms of connectivity widely confirm the findings in relation to the structure and performances of the two networks. In particular, the OD pair connectivity within the air network appears to be rather stable along all the different market sections, whilst rail connectivity plunges on distances longer than 500 km. It is worth noting that on longer distances rail not only provides a considerably lower number of direct connections but also poorer performances, as its frequencies and travel times appear to suffer. This highlights that performances probably play a role in the limited availability of long-distance rail services, suggesting that more endeavours should be directed at improving rail competitiveness on longer distances. As previously mentioned, rail has the potential to compete with air up to ranges of 1.000-1.200 km, however, the findings suggest that this potential has not been explored yet.

The fact that air supplies more than half of those connections indicates the considerable importance of this market suggesting that by increasing its performance rail could consequently capture wider shares of the market. Thus, the findings lead to conclude that in the long-distance market air connectivity is superior compared to rail because of a considerable gap in supply, due to its poor performance. In particular, analysing the findings through the lens of the literature review it could be argued that some important causes for that include:

- The conservative market dynamics, which do not stimulate railway undertakings to launch new international routes and offer more direct services.
- The fragmentation of the European rail infrastructure and service supply, which creates further barriers to the capacity to offer seamless cross-border connections.
- The focus of rail on regional coverage rather than long-distance services, both from a service and infrastructure perspectives. For instance, the European high-speed network is far from complete, and it appears to be a crucial condition for rail to compete with air. This relates to the second condition, as many national operators and stakeholders still tend to favour the optimisation of operations at the national rather than European level.

Furthermore, confirming what is already discussed in terms of network structure, a thorough analysis of the connectivity has pointed out that rail in central Europe is mostly employed as a high-coverage transport service for short-medium routes with lower speeds and higher frequencies that aims to provide capillary access to neighbour cities and regions. Observing the connectivity patterns typical of countries like France or Spain, where high-speed services are at the centre of the rail system, it is possible to argue that changing the paradigm and considering rail as a valid alternative to rail might allow this mode to break free from its current limitations. It is worth noting that the infrastructure layer, despite representing an important hindering factor for rail, has also some positive impacts on connectivity that could be employed to make the mode more attractive on some specific market segments. In fact, the findings have demonstrated that due to the possibility to have stops on lines rail can offer extremely high connectivity also to minor cities. In this regard, the case of Austria is particularly interesting, featuring some of the most connected urban areas in terms of available direct connections (i.e. local connection directness) within the European rail network in relation to their size and importance (e.g. Linz). Thus, it is possible to conclude that it is particularly important to strike a balance between a higher number of stops (regional coverage) and faster/more direct services (long-distance).

"What are the critical components in terms of potential and performances?"

The fourth and final sub-question has been answered by identifying the most critical components in the network using the set of indicators employed to answer the previous two questions, at the node and link level. In particular, a summary of the critical components identified within this study is provided in Section 5.6. At the node level, the most critical components appear to be slightly differently distributed across the two networks. Whilst within the air network the importance of urban areas generally depends on its importance in relation to the airport infrastructure, in the rail network the importance of nodes appears to heavily depend on their geographical centrality. Similar patterns are found at the link level, as the results show that most competitive links are in central Europe. In particular, 350 links appear to be currently competitive, suggesting that a limited improvement in performance could allow it to outperform air. Furthermore, the high urban density in central Europe also appears to be a factor that aids to create the right conditions for rail to thrive. However, although rail appears to better adapt to these conditions, the findings highlight that its performance has still space for improvement. In fact, despite the very high frequencies rail is not always particularly performing in central Europe. This is probably due to the high travel times, possibly caused by the lack of high-speed infrastructure and services, which this research points out to be particularly important in allowing substitution on long distances. Finally, some crucial routes in central Europe (e.g. Amsterdam - Frankfurt and Paris - Zürich) featuring similar frequencies across the two modes and shorter rail travel times highlight the crucial importance of feeder services between the main European air hubs to serve intercontinental destinations. Air frequencies on these routes could be captured and substituted by fostering air-rail inter-modality and offering alternative rail services between airport terminals with inter-modal fare integration. Thus, to capture feeder services and to effectively substitute them it is envisaged that rail services have to feature high frequencies and provide direct connections between airport terminals.

Findings also show that the large majority of rail monopolies are made up of shorter national routes, whilst international OD pairs appear to be less connected and with poorer performances. In particular, excluding some exceptions, such as the cross-border connections between Austria, Germany and Hungary, which feature rather good supply and performances, most of the other countries are lagging behind in terms of international rail supply. Furthermore, despite the good performances of rail on these shorter routes, rail still has not managed to substitute air on most of the routes. In this regard, Dobruszkes et al. (2022) highlight that the main reasons why super short-haul flights still exist fall into three categories:

- Hostile physical geography. In particular, this study has highlighted some corridors where considerable detours are required due to geographical conformation of the territory (e.g. Barcelona-Rome due to the sea, and Nice-Geneva due to the alps).
- Commercial reasons. These include feeder services which airlines require to feed their hubs, services targeted at wider regions, the suburbia and areas which are not necessarily centrally located within cities, and mutualising flights ad triangular routes. The former type of service is extensively highlighted throughout this study.
- Political reasons. These include PSOs and subsidised services required to provide access to remote areas. This is the least interesting for this research given that only major urban areas are included within the scope.

Following the results of this study, two more plausible sets of reasons are added to the list:

- Unavailable infrastructure. These include all those OD pairs that cannot be efficiently connected by rail due to the limitations and constraints in terms of available infrastructure (e.g. the Copenhagen-Hamburg—Bremen-Amsterdam corridor due to the infrastructural gap between Groningen and Bremen, Spain - Portugal connections and border cities in radial systems such as Spain).
- Unavailable capacity. These include all those OD pairs where rail service provision is limited by the high density of rail services and infrastructure with no further capacity available.

It is worth noting that whilst routes included in the first set of categories, as specified by Dobruszkes et al. (2022), are most likely to remain air monopolies, for the latter two categories the tide can be turned by improving/upgrading existing infrastructure or building new links. A notable exception is the case of the feeder services, which, as mentioned before, could be effectively substituted by rail through the inter-modal integration of air and rail services.

The findings have shown that the routes attractive for substitution tend to concentrate towards central Europe on shorter distances. This highlights that there might be a direct dependency between the geographical centrality of the nodes and substitution attractiveness, suggesting that substitution could probably more easily take place in central Europe compared to more peripheral areas. The analysis has also highlighted how rail attractiveness greatly depends on spatial distances and infrastructure availability, due to the considerably lower speeds compared to air, suggesting that not all OD pairs within the European continent might be effectively connected. In particular, this implies that there currently is an imbalance between central Europe and the peripheries of the continent in terms of the possibilities of rail to substitute air. This, however, also points out that there are considerable differences across European regions, suggesting that context-specific policies and solutions are required in order to solve the recurring issues of the sector. For instance, the findings suggest that central Europe might have problems related to congestion and capacity due to the considerable through traffic, whilst peripheral areas such as eastern Europe and the Iberian peninsula might face an opposite threat, with sub-optimal infrastructure utilisation rates. Thus, it appears crucial to develop specific policies for different contexts, as due to the considerable regional differences it seems hardly possible to use a one-rule-fits-all approach. Finally, the analysis of the significance has shown that the betweenness centrality used on the network with travel resistance employed link weight is a useful indicator. In particular, identifying the direct connections that are most often employed by indirect services has highlighted the backbone of European rail and air traffic in terms of supplied service networks.

"From a network connectivity perspective, what is the potential of rail to compete with air transport in the European long-distance market?"

Having answered all the sub-questions, it is possible to respond to the main research question. This thesis has provided a broad overview of the current state of the supply of rail and air services in the European market identifying the most significant nodes and links in the network. Using the connectivity index it is possible to compare the performances and the attractiveness of the two modes. Overall, this study shows that the air network connects efficiently all major European cities, whilst in the rail one there are still considerable inequalities in terms of performances across different geographical locations. From this picture, two main conclusions could be deduced. First of all, the rail network has important limitations and constraints on longer distances, and second that the sector is more fragmented and less cohesive compared to air. To minimise the environmental impact of long-distance transport, it appears beneficial to focus on capturing the demand for connections shorter than 500 km by focusing on inter-modality and air-rail integration and to focus on enhancing rail performances and opening new competitive direct routes on OD distances longer than 500 km. On the long-distance European market, currently, rail appears to be mostly complementary to air, rather than competing with, as the intermodal competition appears to be limited to shorter routes. However there is the potential for rail to be competitive with air also on longer routes, either by offering more direct services, improving schedule coordination, or increasing the supply of high-speed and night train services. In this regard, it seems important to highlight that more attention should be devoted to analysing the specific infrastructural gaps and the optimal directions for investments. From the findings, it is also possible to conclude that, the creation of a single rail market with more homogeneous characteristics is fundamental in order to increase the modal competitiveness of rail on longer distances up to 1.200 km. In particular, understanding and considering the different characteristics of each system appears to be a crucial premise to guarantee the success of the harmonisation process that is taking place within the European rail market.

Thus, the main bottlenecks hindering the capacity of rail to compete and substitute air from a network supply perspective have been identified and summarised in the following points:

- Lack of direct connections. Too many transfers are currently required to reach most destinations, making the mode less attractive for passengers.
- Lack of long-distance services. This is especially the case in the 500 - 1.000 km market segment, where air serves the overwhelming majority of connections despite theoretical studies arguing that rail should be able to compete.
- Lack of specific long-distance service types such as high-speed connections (i.e. low travel time) and/or night trains (i.e. low travel time perception).
- Lack of cross-border, international services, with a few exceptions only.
- Lack of homogeneity in the characteristics of the national systems. The ideas of a single rail area have not managed to translate into reality yet, as the current rail supply appears fragmented across different countries.
- Dependence on schedule coordination, risk of domino effect and delay propagation over all the network. This has rather severe consequences in terms of the impact of disruptions on mode attractiveness.

Finally, the findings of this study suggest that the barriers preventing rail to compete with air are not limited to the supply and the network performances, supporting the theses brought forward by Witlox et al. (2022). Other than improving rail performances and widening the supply of rail it appears fundamental to also work on offering competitive prices, improving ticketing and booking and making the mode more appealing through specific marketing policies. In this regard, an interesting example regards the use of estimated door-to-door travel times rather than in-vehicle times when selling tickets.

6.2. Implications for Practice and Policy Making

This study has a twofold set of implications for practice and policy-making, an empirical and a methodological one. The former relates to the provision of a general overview of the current supply of air and rail services, whilst the latter refers to the methodological insights provided on the applications of network science analysis and connectivity indicators for practice. In particular, this study has proposed an approach aimed at translating empirical results into straightforward knowledge that can effectively aid policy-making, transport planning and governance in general. In this regard, it could be argued that the visualisations have unearthed the added value of spatial data analysis in transport planning processes at the strategic level. It is worth noting that the methodology employed within this study represents a base from which more specific analysis can be developed. In fact, while this study represents an insightful and necessary first step in understanding the connectivity of long-distance European transport networks in relation to the modal competitiveness between rail and air, it falls short in addressing many important themes for policy-making and network planning. In particular, due to the broad geographical scope, the analysis of the results is extensive but not particularly detailed. Further research is required to make informed decisions in regard to the development of the European rail network both from a service and infrastructure perspective. In this regard, it is important to highlight that the numerical outcomes of the study can be employed to further analyse more specific case studies and/or to tackle the problem from different perspectives highlighting some aspects that have been overlooked by this study. In particular, it is worth noting the added value of building an interactive digital tool with the outcome data from this mode in providing a more extensive overview of the data that could help to effectively aid policymakers and transport planners across the industry.

Nevertheless, the findings of this study already provide a broad understanding of the connectivity distribution across all the nodes included in the research and some specific categories of connections. These insights can be employed by policymakers and industry professionals to make informed decisions related to the planning of specific policies or service supply at the link and node level (e.g. for railway undertakings to plan and schedule their services). The node analysis, in fact, allows identifying the main hubs within the network, the urban areas more directly connected with the rest of the network and the nodes with the highest connectivity. Furthermore, through the analysis of the OD relations substitution-ready, substitution potential and substitution attractive connections are identified. The former category represents OD pairs where rail is either already able to substitute air (i.e. fully substitution ready), or where a marginal increase in frequencies or marginal decrease in travel times would allow for the inter-modal substitution (i.e. partially substitution ready). Substitution potential OD pairs, on the other hand, despite being under a competition regime (i.e. both rail and air links are available), feature air services with better performances from both the frequency and travel time perspectives. In this case, direct rail services despite being available are not competitive and greater improvements are required to capture air demand. The latter category represents the links with high air demand over which rail offers no direct service. Often these OD pairs are not connected by rail infrastructure and would require considerable investments in order to become competitive. However, in some cases given the considerable air demand on these routes, it might be economically feasible to make such investments. The overview of the current state of the infrastructure, highlighting the corridors employed by substitution ready and by substitution potential/attractive OD pairs highlighting the infrastructural gaps and the areas of improvement where investments could be directed. Finally, the list of significant links provides useful insights into the connections that are most trafficked and consequently are probably more vulnerable, requiring particular attention. It is important to highlight that the findings of this study are not conclusive and further research is required to optimise investment direction and assess and quantify their economic feasibility.

6.3. Limitations

As previously mentioned this is a preliminary study that, despite providing some interesting insights, is characterised by some limitations. This section will try to capture and discuss them suggesting ways to overcome them. First, it is important to point out that the data employed within this study is not perfect. In particular, the availability of air connections might be slightly overestimated as Flight Aware records all movements of aircraft, disregarding whether they are full or empty, in service or not. In fact, sometimes airlines fly empty aircraft for scheduling reasons or to keep airport slots. Furthermore, the provision of high-speed rail services is most probably underestimated, as a few railway undertakings (e.g. Italo) do not provide their data to UIC merits.

A first limitation regards the degree of comparability of the results across the two modes and across different indices. As highlighted in the conclusion the comparability between air and rail represents a complexity that had to be faced when developing this study. Despite considerable efforts being made to maximise the degree of comparability between the two networks, the results have highlighted some minor limitations in this regard, especially at the network-wide global level. In particular, three possible solutions that could aid to increase the degree of comparability of the results across the two modes are provided below:

- Reducing the number of train service types included in the study to only those that actually compete directly with air on the long-distance market, such as high-speed and night trains. This in turn would reduce the average frequencies and the average travel time of rail, making the results more in line with air. However, it is important the important limitations of this choice, as at the current state there would be a lot of gaps in the rail network, which would prevent it from effectively using network science.
- Increasing the lower bound of the long-distance transport definition from 100km to 250km. This study has shown that air can hardly compete with rail on OD pairs located less than 250 km apart, with a few exceptions in the case of specific morphological conditions and the unavailability of rail infrastructure. However, a drawback of this approach relates to the missed
- Employing a logarithmic scale to re-scale the magnitude of the connectivity indicator. This, in fact, would lead to a wider distribution of the values between 0 and 1, which represent all the cases where the numerator is smaller than the denominator. It is important to note that in the case of the indicator developed within this study this would imply assigning more importance to variations in travel impedance rather than connection intensity, as opposed to the current state. Furthermore, using a logarithm would lead to negative values, which could create problems in using network science indicators.

In regards to the comparability of the different node connectivity indices, as pointed out throughout this report, an important limitation relates to the different magnitudes of the indicators at the local and global levels which currently prevents from comparing them. This can be solved by normalising the values between 0 and 1 using a min-max feature scaling normalisation.

Furthermore, the efforts made to allow comparability also had specific drawbacks and led to some limitations. To compare rail and air results at the node and link level this study employs urban areas as nodes rather than terminals. In this regard, it is important to highlight that using data aggregated at the terminal rather than at the urban area's level might lead to more precise results in terms of the actual availability of transfers and their duration. Changing terminal within the same urban area, in fact, requires additional travel time and effort which is not accounted for. Within this study, direct connections statistics and P-Space networks are employed to compute the shortest path and the specific characteristics of indirect connections. In particular, transfer times are considered to be a function of the frequencies on the different legs making up the shortest path connecting a certain OD pair. However, this does not accurately represent reality for two main reasons. First of all, because the distribution of services within a certain time period might not be regular. Secondly, because for lower frequencies the transfer times assumed in this study increase exponentially, despite in reality with accurate planning and schedule coordination they might actually be minimal. Thus, to increase the reliability of the model in terms of indirect connections more realistic estimates of total travel time should be included. A possible solution in this instance is to take into account the temporal coordination and schedule synchronisation of consecutive services to assess the availability and specific characteristics of indirect services, including transfer times. Some important elements to consider in this case are the complexities related to data retrieval and the risks in terms of model cumbersomeness. It is worth noting that an interesting approach in relation to the inclusion of transfers into a connectivity indicator for both rail, air and inter-modal traffic is the one employed by Zhu, Zhang and Zhang (2019).

Moreover, in order to construct networks with urban areas as nodes an additional re-aggregation step is required. It is worth noting that the assumptions and limitations related to this process, which have been discussed in Section 3.2.3, could be avoided by employing terminals as nodes. Given the considerable influence of the design choices related to the model setup and all the assumptions made throughout this study it is deemed paramount to take them into account when reading the results and in for the future development of similar models. In this regard, other than the aforementioned limitations due to the aggregation level, it seems important to highlight the limitations related to the number of nodes included in the analysis. In fact, the results suggest that the inclusion of the entirety of the nodes served nodes would probably lead to somehow different results at the network-wide and global level. In this study, the inclusion of a selection of the major European urban areas was driven by the need to limit the cumbersomeness of the model and allow direct comparability at the node and link level. However, for a more accurate analysis of the network structure and properties, it is suggested to include a larger number of nodes. It is also worth noting that some urban areas were excluded due to the considerable problems in obtaining data, especially for eastern European countries. Removing these gaps in the data availability is required to include also those cities.

Finally, another limitation relates to the use of frequencies as a measure of service capacity. Frequencies on their own provide interesting insights into the number of available services in a certain time frame, however, the amount of seats offered is another important feature that could provide additional knowledge on the current supply of long-distance services. In particular, the different characteristics of the two means of transport make this feature particularly interesting to further study. In fact, within the air market, the differences in seats per service are rather limited, given that most continental flights generally range from 80-100 seats of smaller regional jets (e.g. Airbus A220-100 and Embraer E175) to 180-200 seats of larger narrow body aircraft (e.g. Airbus A321 and Boeing 737-900). On the other hand, rail service capacity can see much greater variations of seats offered per service depending on the type of rolling stock, ranging from short regional trains consisting of a few carriages to the 508 seats of a duplex TGV, and on the possibility of coupling and uncoupling the rolling stock. Other important factors that have been disregarded by this study but that are likely to have an impact on passenger modal choice are service reliability and punctuality. Taking into account such factors would allow to provide a more complete overview of the rail-air inter-modal competition and the factors hindering rail from substituting rail.

6.4. Recommendations for Further Research

The main question that this research has tried to answer relates to the potential of rail to compete with air in the European long-distance market. As already highlighted in the previous paragraphs further research would be beneficial in translating the findings of this study into precise and accurate advice for practice. Furthermore, from a research perspective, it is believed that additional research from the demand side is required to complement and reinforce the findings of this study. This research, in fact, analysing the current supply and quantifying its performances, has neglected the demand perspective of the problem. However, this represents a fundamental aspect to provide a comprehensive answer to the main research question, which should receive special attention from researchers. In fact, to assess what are the concrete possibilities of a modal shift to rail it is necessary to evaluate what is the actual demand, what is the willingness of such demand to shift to rail and understand what are the factors influencing the aforementioned willingness and which conditions are required to ensure this shift. Furthermore, demand is also a fundamental aspect to infer the potential developments in the market and to consequently forecast how future supply could and should adapt and in which time frame. This represents a particularly important element to allow policy-makers and industry professionals to make informed decisions and to develop farsighted plans both at the governance and strategic/tactical planning levels. Moreover, further research in that sense is also required to precisely direct the considerable investments that are going to reach the rail sector. In this regard, it seems interesting to point out that the methodology and the indicators developed for this study could also be employed to compare different scenarios, for instance, by observing time variations in terms of connectivity within the same mode (rail or air) to assess the evolution of the indicators across the years. Based on these temporal patterns and the model specifications, a simulation tool for European long-distance transport could also be constructed to provide forecasts and to assess how to practically make rail more attractive, for instance by increasing frequencies or reducing travel times. This would prove particularly useful in evaluating the impact of investments and policies on the modal split, allowing policy-makers and industry professionals to fine-tune them. Simpler approaches, such as the study from Kroes and Savelberg (2019), also represent a valid alternative that might be more straightforward developed from the results of this study. Furthermore, future research in this direction should also aim to include road modes (i.e. private cars, carpooling and long-distance buses) that are not considered within this study, in order to provide a more complete overview. Finally, it is worth noting that the specific characteristics of different rail service types, including high-speed services, inter-cities and night trains, are aggregated, not allowing to understand their impact on connectivity, modal competitiveness and the overall attractiveness of rail. Further research, analysing the performances of the different rail service types and exploring how the connectivity varies across them and how each contributes to the overall rail connectivity, could provide useful insights to aid decision-making processes directed at shaping future rail supply scenarios.

References

Adler, N., Pels, E. & Nash, C. (2010). High-speed rail and air transport competition: Game engineering as tool for cost-benefit analysis [Modelling Non-urban Transport Investment and Pricing]. *Transportation Research Part B: Methodological*, *44*(7), 812–833. https://doi.org/https://doi.org/10.1016/j.trb.2010.01.001

Albalate, D., Bel, G. & Fageda, X. (2015). Competition and cooperation between high-speed rail and air transportation services in europe. *Journal of Transport Geography*, *42*, 166–174. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2014.07.003

Allroggen, F., Wittman, M. D. & Malina, R. (2015). How air transport connects the world - a new metric of air connectivity and its evolution between 1990 and 2012. *Transportation Research Part E: Logistics and Transportation Review*, *80*. https://doi.org/10.1016/j.tre.2015.06.001

Augustyniak, W. & Olipra, Ł. (2014). The potential catchment area of Polish regional airports. *Journal of International Studies*, *7*. https://doi.org/10.14254/2071-8330.2014/7-3/13

Avogadro, N., Cattaneo, M., Paleari, S. & Redondi, R. (2021). Replacing short-medium haul intra-European flights with high-speed rail: Impact on CO2 emissions and regional accessibility. *Transport Policy*, *114*. https://doi.org/10.1016/j.tranpol.2021.08.014

Bala, B. K., Arshad, F. M. & Noh, K. M. (2017). *System dynamics: Modelling and simulation*.

Barabási, A. L. (2009). Scale-free networks: A decade and beyond. *Science*, *325*. https://doi.org/10.1126/science.1173299

Barabási, A. L. & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*. https://doi.org/10.1126/science.286.5439.509

Barabási, A. L. & Pósfai, M. (2016). *Network science*. Cambridge University Press. http://barabasi.com/networksciencebook/

Barrat, A., Barthélemy, M., Pastor-Satorras, R. & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, *101*. https://doi.org/10.1073/pnas.0400087101

Behrens, C. & Pels, E. (2012). Intermodal competition in the London–Paris passenger market: High-Speed Rail and air transport. *Journal of Urban Economics*, *71*(3), 278–288. https://doi.org/https://doi.org/10.1016/j.jue.2011.12.005

Bergantino, A. S., Capozza, C. & Capurso, M. (2015). The impact of open access on intra- and intermodal rail competition. A national level analysis in Italy. *Transport Policy*, *39*, 77–86. https://doi.org/https://doi.org/10.1016/j.tranpol.2015.01.008

Bergantino, A. S. & Madio, L. (2020). Intermodal competition and substitution. HSR versus air transport: Understanding the socio-economic determinants of modal choice [Air Transport Markets, Strategies and Policies]. *Research in Transportation Economics*, *79*, 100823. https://doi.org/https://doi.org/10.1016/j.retrec.2020.100823

Beria, P., Tolentino, S., Bertolin, A. & Filippini, G. (2019). Long-distance rail prices in a competitive market. Evidence from head-on competition in Italy. *Journal of Rail Transport Planning & Management*, *12*, 100144. https://doi.org/https://doi.org/10.1016/j.jrtpm.2019.100144

Bhat, C. R. (1995). A heteroscedastic extreme value model of intercity travel mode choice. *Transportation Research Part B*, *29*. https://doi.org/10.1016/0191-2615(95)00015-6

Bombelli, A. (2020). Integrators' global networks: A topology analysis with insights into the effect of the COVID-19 pandemic. *Journal of Transport Geography*, *87*. https://doi.org/10.1016/j.jtrangeo.2020.102815

Borsati, M. & Albalate, D. (2020). On the modal shift from motorway to high-speed rail: evidence from Italy. *Transportation Research Part A: Policy and Practice*, *137*. https://doi.org/10.1016/j.tra.2020.04.006

Broido, A. D. & Clauset, A. (2019). Scale-free networks are rare. *Nature Communications*, *10*. https://doi.org/10.1038/s41467-019-08746-5

Burghouwt, G. & de Wit, J. (2005). Temporal configurations of European airline networks. *Journal of Air Transport Management*, *11*. https://doi.org/10.1016/j.jairtraman.2004.08.003

Burghouwt, G. & Redondi, R. (2013). Connectivity in air transport networks an assessment of models and applications. *Journal of Transport Economics and Policy*, *47*.

Cai, K. Q., Zhang, J., Du, W. B. & Cao, X. B. (2012). Analysis of the Chinese air route network as a complex network. *Chinese Physics B*, *21*. https://doi.org/10.1088/1674-1056/21/2/028903

Calzada-Infante, L., Adenso-Díaz, B. & García Carbajal, S. (2020). Analysis of the European international railway network and passenger transfers. *Chaos, Solitons & Fractals*, *141*, 110357. https://doi.org/https://doi.org/10.1016/j.chaos.2020.110357

Cats, O. (2017). Topological evolution of a metropolitan rail transport network: The case of Stockholm. *Journal of Transport Geography*, *62*. https://doi.org/10.1016/j.jtrangeo.2017.06.002

Chen, C., D'Alfonso, T., Guo, H. & Jiang, C. (2018). Graph theoretical analysis of the Chinese high-speed rail network over time [Long-distance passenger transport market]. *Research in Transportation Economics*, *72*, 3–14. https://doi.org/https://doi.org/10.1016/j.retrec.2018.07.030

Chen, C. & Vickerman, R. (2017). Can transport infrastructure change regions' economic fortunes? Some evidence from Europe and China. *Regional Studies*, *51*(1), 144–160. https://doi.org/10.1080/00343404.2016.1262017

Clewlow, R. R. L., Sussman, J. M. & Balakrishnan, H. (2012). Interaction of High-Speed Rail and Aviation: Exploring Air–Rail Connectivity. *Transportation Research Record*, *2266*(1), 1–10. https://doi.org/10.3141/2266-01

Clewlow, R. R. L., Sussman, J. M. & Balakrishnan, H. (2014). The impact of high-speed rail and low-cost carriers on European air passenger traffic. *Transport Policy*, *33*. https://doi.org/10.1016/j.tranpol.2014.01.015

Cornet, Y., Lugano, G., Georgouli, C. & Milakis, D. (2022). Worthwhile travel time: A conceptual framework of the perceived value of enjoyment, productivity and fitness while travelling. *Transport Reviews*, *42*(5), 580–603. https://doi.org/10.1080/01441647.2021.1983067

Danesi, A. (2006). Measuring airline hub timetable co-ordination and connectivity: Definition of a new index and application to a sample of european hubs. *European Transport*, *34*.

Dennis, N. (1998–July 16). Competition between hub airports in Europe and a methodology for forecasting connecting traffic. *Proceedings of the 8th World Conference on Transport Research*.

Derrible, S. & Kennedy, C. (2009). Network analysis of world subway systems using updated graph theory. *Transportation Research Record*. https://doi.org/10.3141/2112-03

Derrible, S. & Kennedy, C. (2010a). Characterizing metro networks: State, form, and structure. *Transportation*, *37*. https://doi.org/10.1007/s11116-009-9227-7

Derrible, S. & Kennedy, C. (2010b). The complexity and robustness of metro networks. *Physica A: Statistical Mechanics and its Applications*, *389*. https://doi.org/10.1016/j.physa.2010.04.008

Derrible, S. & Kennedy, C. (2011). Applications of graph theory and network science to transit network design. *Transport Reviews*, *31*. https://doi.org/10.1080/01441647.2010.543709

Derudder, B., Devriendt, L. & Witlox, F. (2010). A spatial analysis of multiple airport cities. *Journal of Transport Geography*, *18*. https://doi.org/10.1016/j.jtrangeo.2009.09.007

Di Pietrantonio, L. & Pelkmans, J. (2004). The Economics of EU Railway Reform. *Competition and Regulation in Network Industries*, *5*. https://doi.org/10.1177/178359170400500304

Dijkstra, L., Poelman, H. & Veneri, P. (2019). The EU-OECD definition of a functional urban area. https://doi.org/https://doi.org/https://doi.org/10.1787/d58cb34d-en

Dobruszkes, F., Lennert, M. & Van Hamme, G. (2011). An analysis of the determinants of air traffic volume for European metropolitan areas. *Journal of Transport Geography*, *19*(4), 755–762. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2010.09.003

Dobruszkes, F., Mattioli, G. & Mathieu, L. (2022). Banning super short-haul flights: Environmental evidence or political turbulence? *Journal of Transport Geography*, *104*, 103457. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2022.103457

Dupuy, G. (2013). Network geometry and the urban railway system: The potential benefits to geographers of harnessing inputs from “naive” outsiders. *Journal of Transport Geography*, *33*. https://doi.org/10.1016/j.jtrangeo.2013.09.012

Dziedzic, M., Njoya, E. T., Warnock-Smith, D. & Hubbard, N. (2020). Determinants of air traffic volumes and structure at small European airports [Air Transport Markets, Strategies and Policies]. *Research in Transportation Economics*, *79*, 100749. https://doi.org/https://doi.org/10.1016/j.retrec.2019.100749

Eurocontrol. (2020). *Data Snapshot 9 on 2020’s smaller aircraft, shorter flights and fluctuating weight*. Retrieved September 2, 2022, from https://www.eurocontrol.int/publication/eurocontrol-data-snapshot-2020s-smaller-aircraft-shorter-flights-and-fluctuating-weight

European Commission. (2013). Case No COMP/M.6663 – Ryanair/AER Lingus III. *Regulation (EC) No 139/2004 Merger Procedure*, *3*(8). https://ec.europa.eu/competition/mergers/cases/decisions/m6663_20130227_20610_3904642_EN.pdf

European Commission. (2020a). *Press Release: The journey begins – 2021 is the European Year of Rail!* Retrieved April 6, 2022, from https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2528

European Commission. (2020b). Sustainable and Smart Mobility Strategy. https://doi.org/10.2775/932

European Commission. (2020c). *Transport and the Green Deal*. Retrieved April 6, 2022, from https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal/transport-and-green-deal_en

European Commission. (2021a). *New Action Plan: boosting long-distance and cross-border passenger rail*. Retrieved April 6, 2022, from https://transport.ec.europa.eu/news/action-plan-boost-passenger-rail-2021-12-14_en

European Commission. (2021b). Seventh monitoring report on the development of the rail market under Article 15(4) of Directive 2012/34/EU of the European Parliament and of the Council. *Report from the Commission to the European Parliament and the Council*. https://op.europa.eu/en/publication-detail/-/publication/4403ba11-5588-11eb-b59f-01aa75ed71a1

European Commission. (2022a). *Fourth railway package of 2016*. Retrieved April 6, 2022, from https://transport.ec.europa.eu/transport-modes/rail/railway-packages/fourth-railway-package-2016_en

European Commission. (2022b). *Railway packages*. Retrieved April 6, 2022, from https://transport.ec.europa.eu/transport-modes/rail/railway-packages_en

European Commission & Directorate-General for Mobility and Transport. (2021). *Long-distance cross-border passenger rail services : Final report*. https://doi.org/10.2832/019365

European Commission & Eurostat. (2019). *Methodological manual on territorial typologies: 2018 edition*. Publications Office. https://doi.org/doi/10.2785/930137

European Environment Agency. (2021). Transport and Environment Report 2020 — Train or Plane? https://doi.org/10.2800/43379

European Parliament & Council of the European Union. (2020). Decision (EU) 2020/2228 of the 28.12.2020. *Official Journal of the European Union*, *L 437*, 108–115.

Eurostat. (2019). *People on the move — statistics on mobility in europe — 2019 edition*. https://doi.org/10.2785/50487

Eurostat. (2021a). *Climate change - driving forces*. Retrieved May 6, 2022, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Climate_change_-_driving_forces

Eurostat. (2021b). *Impact of the COVID-19 pandemic on rail passenger transport in Q4 2020*. Retrieved May 6, 2022, from https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210518-1

Eurostat. (2021c). *Railway passenger transport statistics - quarterly and annual data*. Retrieved May 6, 2022, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Railway_passenger_transport_statistics_-_quarterly_and_annual_data

Eurostat. (2021d). *Trends in greenhouse gas emissions*. Retrieved May 6, 2022, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Greenhouse_gas_emission_statistics_-_emission_inventories#Trends_in_greenhouse_gas_emissions

Eurostat. (2022a). *Airport traffic data by reporting airport and airlines*. Retrieved May 30, 2022, from https://ec.europa.eu/eurostat/databrowser/view/AVIA_TF_APAL/default/table?lang=EN&category=avia.avia_tf

Eurostat. (2022b). *Metropolitan Regions Background*. Retrieved May 30, 2022, from https://ec.europa.eu/eurostat/web/metropolitan-regions/background

Eurostat. (2022c). *Passenger cars per 1 000 inhabitants*. Retrieved July 27, 2022, from https://ec.europa.eu/eurostat/databrowser/view/ROAD_EQS_CARHAB/default/table?lang=en&category=road.road_eqs

Feng, X., He, S., Li, G. & Chi, J. (2021). Transfer network of high-speed rail and aviation: Structure and critical components. *Physica A: Statistical Mechanics and its Applications*, *581*. https://doi.org/10.1016/j.physa.2021.126197

Feuerstein, L., Busacker, T. & Xu, J. (2018). Factors influencing open access competition in the european long-distance passenger rail transport — a delphi study [Competition and Ownership in Land Passenger Transport (selected papers from the Thredbo 15 conference)]. *Research in Transportation Economics*, *69*, 300–309. https://doi.org/https://doi.org/10.1016/j.retrec.2018.03.001

Fisher, H. (1976). The Language and Logic of Forming an Idea. *Journal for the Theory of Social Behaviour*, *6*. https://doi.org/10.1111/j.1468-5914.1976.tb00365.x

Fontoura, W. B. & Ribeiro, G. M. (2021). System dynamics for sustainable transportation policies: A systematic literature review. *Urbe*, *13*. https://doi.org/10.1590/2175-3369.013.E20200259

Fraunhofer ISI & CE Delft. (2020). *Methodology for ghg efficiency of transport modes* (tech. rep.). European Environment Agency.

Frei, A., Kuhnimhof, T. & Axhausen, K. W. (2010). Long-Distance Travel in Europe Today: Experiences with a New Survey. *Transportation Research Board*.

Gattuso, D. & Miriello, E. (2005). Compared analysis of metro networks supported by graph theory. *Networks and Spatial Economics*, *5*. https://doi.org/10.1007/s11067-005-6210-5

Gerike, R. & Schulz, A. (2018). Workshop Synthesis: Surveys on long-distance travel and other rare events. *Transportation Research Procedia*, *32*. https://doi.org/10.1016/j.trpro.2018.10.032

Geurs, K. T. & van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: Review and research directions. *Journal of Transport Geography*, *12*. https://doi.org/10.1016/j.jtrangeo.2003.10.005

Givoni, M., Dobruszkes, F. & Lugo, I. (2012). Uncovering the real potential for air-rail substitution: An exploratory analysis. *Energy, Transport, & the Environment: Addressing the Sustainable Mobility Paradigm*, *9781447127178*. https://doi.org/10.1007/978-1-4471-2717-8_27

Guimerà, R., Mossa, S., Turtschi, A. & Amaral, L. A. (2005). The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proceedings of the National Academy of Sciences of the United States of America*, *102*. https://doi.org/10.1073/pnas.0407994102

Guo, Z., Hao, M., Yu, B. & Yao, B. (2022). Detecting delay propagation in regional air transport systems using convergent cross mapping and complex network theory. *Transportation Research Part E: Logistics and Transportation Review*, *157*. https://doi.org/10.1016/j.tre.2021.102585

Handley, J. C., Fu, L. & Tupper, L. L. (2019). A case study in spatial-temporal accessibility for a transit system. *Journal of Transport Geography*, *75*. https://doi.org/10.1016/j.jtrangeo.2019.01.005

Heufke Kantelaar, M., Molin, E., Cats, O., Donners, B. & van Wee, B. (2022). Willingness to use night trains for long-distance travel. *Travel Behaviour and Society*, *29*, 339–349. https://doi.org/https://doi.org/10.1016/j.tbs.2022.08.002

Humphries, M. D., Gurney, K. & Prescott, T. J. (2006). The brainstem reticular formation is a small-world, not scale-free, network. *Proceedings of the Royal Society B: Biological Sciences*, *273*. https://doi.org/10.1098/rspb.2005.3354

Janic, M. (2003). High-speed rail and air passenger transport: A comparison of the operational environmental performance. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *217*(4), 259–269. https://doi.org/10.1243/095440903322712865

Kaplan, S., Popoks, D., Prato, C. & Ceder, A. (2014). Using connectivity for measuring equity in transit provision. *Journal of Transport Geography*, *37*. https://doi.org/10.1016/j.jtrangeo.2014.04.016

Koppelman, F. S. & Wen, C. H. (2000). The paired combinatorial logit model: Properties, estimation and application. *Transportation Research Part B: Methodological*, *34*. https://doi.org/10.1016/S0191-2615(99)00012-0

Kotegawa, T., Fry, D., DeLaurentis, D. & Puchaty, E. (2014). Impact of service network topology on air transportation efficiency. *Transportation Research Part C: Emerging Technologies*, *40*, 231–250. https://doi.org/https://doi.org/10.1016/j.trc.2013.11.016

Kroes, E. & Savelberg, F. (2019). Substitution from air to high-speed rail: The case of amsterdam airport. *Transportation Research Record*, *2673*(5), 166–174. https://doi.org/10.1177/0361198119839952

Li, W. & Cai, X. (2007). Empirical analysis of a scale-free railway network in china. *Physica A: Statistical Mechanics and its Applications*, *382*. https://doi.org/10.1016/j.physa.2007.04.031

Li, X., Tang, J., Hu, X. & Wang, W. (2020). Assessing intercity multimodal choice behavior in a Touristy City: A factor analysis. *Journal of Transport Geography*, *86*, 102776. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2020.102776

Lieshout, R. (2012). Measuring the size of an airport's catchment area [Special Section on Accessibility and Socio-Economic Activities: Methodological and Empirical Aspects]. *Journal of Transport Geography*, *25*, 27–34. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2012.07.004

Lieshout, R., Malighetti, P., Redondi, R. & Burghouwt, G. (2016). The competitive landscape of air transport in Europe [Air transport liberalisation and airline network dynamics: Investigating the complex relationships]. *Journal of Transport Geography*, *50*, 68–82. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2015.06.001

Limtanakool, N., Dijst, M. & Schwanen, T. (2006). The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium- and longer-distance trips. *Journal of Transport Geography*, *14*(5), 327–341. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2005.06.004

Lin, J. & Ban, Y. (2013). Complex Network Topology of Transportation Systems. *Transport Reviews*, *33*. https://doi.org/10.1080/01441647.2013.848955

Luo, D., Cats, O., van Lint, H. & Currie, G. (2019). Integrating network science and public transport accessibility analysis for comparative assessment. *Journal of Transport Geography*, *80*, 102505. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2019.102505

Majima, T., Katuhara, M. & Takadama, K. (2007). Analysis on transport networks of railway, subway and waterbus in Japan. *Studies in Computational Intelligence*, *56*. https://doi.org/10.1007/978-3-540-71075-2_8

Malichová, E., Cornet, Y. & Hudák, M. (2022). Travellers' use and perception of travel time in long-distance trips in Europe. *Travel Behaviour and Society*, *27*, 95–106. https://doi.org/https://doi.org/10.1016/j.tbs.2021.12.003

Malighetti, P., Paleari, S. & Redondi, R. (2008). Connectivity of the European airport network: "Self-help hubbing" and business implications. *Journal of Air Transport Management*, *14*(2), 53–65. https://doi.org/https://doi.org/10.1016/j.jairtraman.2007.10.003

Malina, R. (2010). Competition in the German airport market: An empirical investigation.

Marcucci, E. & Gatta, V. (2011). Regional airport choice: Consumer behaviour and policy implications. *Journal of Transport Geography*, *19*(1), 70–84. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2009.10.001

Martín, J. C., Román, C., García-Palomares, J. C. & Gutiérrez, J. (2014). Spatial analysis of the competitiveness of the high-speed train and air transport: The role of access to terminals in the Madrid–Barcelona corridor. *Transportation Research Part A: Policy and Practice*, *69*, 392–408. https://doi.org/https://doi.org/10.1016/j.tra.2014.09.010

Moeckel, R., Fussell, R. & Donnelly, R. (2015). Mode choice modeling for long-distance travel. *Transportation Letters*, *7*. https://doi.org/10.1179/1942787514Y.0000000031

Moreno-Monroy, A. I., Schiavina, M. & Veneri, P. (2021). Metropolitan areas in the world. delineation and population trends [Delineation of Urban Areas]. *Journal of Urban Economics*, *125*, 103242. https://doi.org/https://doi.org/10.1016/j.jue.2020.103242

National Research Council. (2005). *Network Science*. The National Academies Press. https://doi.org/10.17226/11516

Newman, M. E. J. (2010). *Networks: An introduction*. Oxford University Press.

Nieße, H. & Grimme, W. (2015). How to measure airport connectivity? - average shortest travel time and average highest path velocity as indicators. *International Journal of Aviation Management*, *2*. https://doi.org/10.1504/ijam.2015.072382

OECD. (2022). *Metropolitan Areas*. Retrieved May 30, 2022, from https://www.oecd.org/regional/regional-statistics/metropolitan-areas.htm

Ortega Pérez, E., Mancebo Quintana, S. & Otero Pastor, I. (2011). Road and railway accessibility atlas of Spain. *Journal of Maps*, *7*. https://doi.org/10.4113/jom.2011.1167

Paleari, S., Redondi, R. & Malighetti, P. (2010). A comparative study of airport connectivity in China, Europe and US: Which network provides the best service to passengers? *Transportation Research Part E: Logistics and Transportation Review*, *46*. https://doi.org/10.1016/j.tre.2009.08.003

Pavlyuk, D. (2009). Spatial competition pressure as a factor of European airports' efficiency. *Transport and Telecommunication*, *10*.

Peer, S., Knockaert, J., Koster, P., Tseng, Y.-Y. & Verhoef, E. T. (2013). Door-to-door travel times in RP departure time choice models: An approximation method using GPS data. *Transportation Research Part B: Methodological*, *58*, 134–150. https://doi.org/https://doi.org/10.1016/j.trb.2013.10.006

Petersen, M., Sessa, C., Enei, R., Ulied, A., Larrea, E. & Obisco, O. (2009). *Transvisions: Report on Transport Scenarios with a 20 and 40 Year Horizon* (tech. rep.). European Commission.

Poelman, H. (2013). Measuring accessibility to passenger flights in Europe: Towards harmonised indicators at the regional level. *European Commission Regional Focus*.

Psaltoglou, A. & Calle, E. (2018). Enhanced connectivity index – a new measure for identifying critical points in urban public transportation networks. *International Journal of Critical Infrastructure Protection*, *21*. https://doi.org/10.1016/j.ijcip.2018.02.003

Rich, J. & Mabit, S. L. (2012). A Long-Distance Travel Demand Model for Europe. *European Journal of Transport and Infrastructure Research*, *12*(1). https://doi.org/10.18757/ejtir.2012.12.1.2946

Rodrigue, J.-P. (2020). *The Geography of Transport Systems* (5th ed.). Routledge. https://doi.org//10.4324/9780429346323

Román, C., Espino, R. & Martín, J. C. (2010). Analyzing competition between the high speed train and alternative modes. The case of the Madrid-Zaragoza-Barcelona corridor. *Journal of Choice Modelling*, *3*. https://doi.org/10.1016/S1755-5345(13)70030-7

Román, C. & Martín, J. C. (2014). Integration of hsr and air transport: Understanding passengers' preferences. *Transportation Research Part E: Logistics and Transportation Review*, *71*, 129–141. https://doi.org/https://doi.org/10.1016/j.tre.2014.09.001

Rothengatter, W., Zhang, J., Hayashi, Y., Nosach, A., Wang, K. & Oum, T. H. (2021). Pandemic waves and the time after Covid-19 – Consequences for the transport sector. *Transport Policy*, *110*, 225–237. https://doi.org/https://doi.org/10.1016/j.tranpol.2021.06.003

Rothfeld, R., Straubinger, A., Paul, A. & Antoniou, C. (2019). Analysis of European airports' access and egress travel times using Google Maps. *Transport Policy*, *81*, 148–162. https://doi.org/https://doi.org/10.1016/j.tranpol.2019.05.021

Rotoli, F., Malavasi, G. & Ricci, S. (2016). Complex railway systems: Capacity and utilisation of interconnected networks. *European Transport Research Review*, *8*. https://doi.org/10.1007/s12544-016-0216-6

Seidenglanz, D., Taczanowski, J., Król, M., Horňák, M. & Nigrin, T. (2021). Quo vadis, international long-distance railway services? Evidence from Central Europe. *Journal of Transport Geography*, *92*, 102998. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2021.102998

Sen, A. K. & Morlok, E. K. (1976). On Estimating Effective Frequencies And Average Waiting Times For Indirect Connections. *Transportation Planning and Technology*, *3*. https://doi.org/10.1080/03081067608717104

Sen, P., Dasgupta, S., Chatterjee, A., Sreeram, P. A., Mukherjee, G. & Manna, S. S. (2003). Small-world properties of the Indian railway network. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *67*. https://doi.org/10.1103/PhysRevE.67.036106

Serafimova, T., Finger, M. & Montero-Pascual, J. (2022). Modal shift: the moment of truth. Main takeaways and lessons learnt from the European year of rail. *RSC Policy Briefs*, 97–111. https://doi.org/10.2870/26564

Sienkiewicz, J. & Hołyst, J. A. (2005). Statistical analysis of 22 public transport networks in Poland. *Phys. Rev. E*, *72*, 046127. https://doi.org/10.1103/PhysRevE.72.046127

Silva, T. C., Dias, F. A., dos Reis, V. E. & Tabak, B. M. (2022). The role of network topology in competition and ticket pricing in air transportation: Evidence from Brazil. *Physica A: Statistical Mechanics and its Applications*, *601*, 127602. https://doi.org/https://doi.org/10.1016/j.physa.2022.127602

Socorro, M. P. & Viecens, M. F. (2013). The effects of airline and high speed train integration. *Transportation Research Part A: Policy and Practice*, *49*, 160–177. https://doi.org/https://doi.org/10.1016/j.tra.2013.01.014

Soh, H., Lim, S., Zhang, T., Fu, X., Lee, G. K. K., Hung, T. G. G., Di, P., Prakasam, S. & Wong, L. (2010). Weighted complex network analysis of travel routes on the Singapore public transportation system. *Physica A: Statistical Mechanics and its Applications*, *389*. https://doi.org/10.1016/j.physa.2010.08.015

Suau-Sanchez, P. & Voltes-Dorta, A. (2019). Drivers of airport scheduled traffic in European winter tourism areas: Infrastructure, accessibility, competition and catchment area. *Journal of Air Transport Management*, *81*, 101723. https://doi.org/https://doi.org/10.1016/j.jairtraman.2019.101723

Sun, X., Wandelt, S., Hansen, M. & Li, A. (2017). Multiple airport regions based on inter-airport temporal distances. *Transportation Research Part E: Logistics and Transportation Review*, *101*, 84–98. https://doi.org/https://doi.org/10.1016/j.tre.2017.03.002

Takebayashi, M. (2015). Multiple hub network and high-speed railway: Connectivity, gateway, and airport leakage. *Transportation Research Part A: Policy and Practice*, *79*. https://doi.org/10.1016/j.tra.2015.03.024

Tardivo, A., Zanuy, A. C. & Martín, C. S. (2021). COVID-19 Impact on Transport: A Paper from the Railways' Systems Research Perspective. *Transportation Research Record*, *2675*(5), 367–378. https://doi.org/10.1177/0361198121990674

Telesford, Q. K., Joyce, K. E., Hayasaka, S., Burdette, J. H. & Laurienti, P. J. (2011). The Ubiquity of Small-World Networks. *Brain Connectivity*, *1*. https://doi.org/10.1089/brain.2011.0038

Travel Time. (2022). *Create thousands of free isochrones with an API*. Retrieved May 30, 2022, from https://traveltime.com/features/isochrones

Trudeau, R. J. (1993). *Introduction to Graph Theory*. Dover Publications, Inc.

van Steen, M. (2010). *Graph Theory and Complex Networks: An Introduction*.

van Acker, V., van Wee, B. & Witlox, F. (2010). When Transport Geography Meets Social Psychology: Toward a Conceptual Model of Travel Behaviour. *Transport Reviews*, *30*(2), 219–240. https://doi.org/10.1080/01441640902943453

von Ferber, C., Holovatch, T., Holovatch, Y. & Palchykov, V. (2007). Network harness: Metropolis public transport. *Physica A: Statistical Mechanics and its Applications*, *380*. https://doi.org/10.1016/j.physa.2007.02.101

von Ferber, C., Holovatch, T., Holovatch, Y. & Palchykov, V. (2009). Public transport networks: Empirical analysis and modeling. *European Physical Journal B*, *68*. https://doi.org/10.1140/epjb/e2009-00090-x

Wang, J., Mo, H., Wang, F. & Jin, F. (2011). Exploring the network structure and nodal centrality of China's air transport network: A complex network approach. *Journal of Transport Geography*, *19*. https://doi.org/10.1016/j.jtrangeo.2010.08.012

Wang, W., Cai, K., Du, W., Wu, X., Tong, L., Zhu, X. & Cao, X. (2020). Analysis of the Chinese railway system as a complex network. *Chaos, Solitons and Fractals*, *130*. https://doi.org/10.1016/j.chaos.2019.109408

Wang, X. F. & Chen, G. (2003). Complex networks: Small-world, scale-free and beyond. *IEEE Circuits and Systems Magazine*, *3*. https://doi.org/10.1109/MCAS.2003.1228503

Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*. https://doi.org/10.1038/30918

Weeks, J. R. (2010). Defining Urban Areas. In T. Rashed & C. Jürgens (Eds.), *Remote Sensing of Urban and Suburban Areas* (pp. 33–45). Springer Netherlands. https://doi.org/10.1007/978-1-4020-4385-7_3

Wei, P., Spiers, G. & Sun, D. (2014). Algebraic connectivity maximization for air transportation networks. *IEEE Transactions on Intelligent Transportation Systems*, *15*. https://doi.org/10.1109/TITS.2013.2284913

Welch, T. F., Mishra, S. & Wang, F. (2015). Interrelationship between Airport Enplanements and Accessibility: Case of Three Airports in Metropolitan Washington, D.C., Region. *Transportation Research Record*, *2501*(1), 46–55. https://doi.org/10.3141/2501-07

Wen, C. H. & Koppelman, F. S. (2001). The generalized nested logit model. *Transportation Research Part B: Methodological*, *35*. https://doi.org/10.1016/S0191-2615(00)00045-X

Wikimedia Commons. (2022). *High Speed Railroad Map of Europe*. Retrieved October 13, 2022, from https://commons.wikimedia.org/wiki/File:High_Speed_Railroad_Map_of_Europe.svg

Williams, M. J. & Musolesi, M. (2016). Spatio-temporal networks: Reachability, centrality and robustness. *Royal Society Open Science*, *3*. https://doi.org/10.1098/rsos.160196

Wilson, R. J. (1972). *Introduction to Graph Theory*. Oliver & Boyd.

Witlox, F., Zwanikken, T., Jehee, L., Donners, B. & Veeneman, W. (2022). Changing tracks: Identifying and tackling bottlenecks in european rail passenger transport. *European Transport Research Review*, *14*(7). https://doi.org/10.1186/s12544-022-00530-9

World Tourism Organization. (2018). *European Union Tourism Trends*. UNWTO. https://doi.org/10.18111/9789284419470

WTO. (2018). *European Union Tourism Trends*. United Nations World Tourism Organisation. https://doi.org/10.18111/9789284419470

Xu, W., Zhou, J., Yang, L. & Li, L. (2018). The implications of high-speed rail for Chinese cities: Connectivity and accessibility. *Transportation Research Part A: Policy and Practice*, *116*. https://doi.org/10.1016/j.tra.2018.06.023

Xu, Z., Zhang, Q., Chen, D. & He, Y. (2020). Characterizing the Connectivity of Railway Networks. *IEEE Transactions on Intelligent Transportation Systems*, *21*. https://doi.org/10.1109/TITS.2019.2909120

Yang, H. & Zhang, A. (2012). Effects of high-speed rail and air transport competition on prices, profits and welfare. *Transportation Research Part B: Methodological*, *46*(10), 1322–1333. https://doi.org/https://doi.org/10.1016/j.trb.2012.09.001

Yang, S. & Chen, Z. (2022). The Impact of COVID-19 on High-Speed Rail and Aviation Operations. *Sustainability*, *14*(3). https://doi.org/10.3390/su14031683

Zanin, M. & Lillo, F. (2013). Modelling the air transport with complex networks: A short review. *European Physical Journal: Special Topics*, *215*. https://doi.org/10.1140/epjst/e2013-01711-9

Zanin, M., Sun, X. & Wandelt, S. (2018). Studying the Topology of Transportation Systems through Complex Networks: Handle with Care. *Journal of Advanced Transportation*, *2018*. https://doi.org/10.1155/2018/3156137

Zhang, A., Wan, Y. & Yang, H. (2019). Impacts of high-speed rail on airlines, airports and regional economies: A survey of recent research. *Transport Policy*, *81*, A1–A19. https://doi.org/https://doi.org/10.1016/j.tranpol.2019.06.010

Zhang, Z., Li, X. & Li, H. (2015). A quantitative approach for assessing the critical nodal and linear elements of a railway infrastructure. *International Journal of Critical Infrastructure Protection*, *8*, 3–15. https://doi.org/https://doi.org/10.1016/j.ijcip.2014.11.001

Zhou, Y., Wang, J. & Sheu, J. B. (2019). On connectivity of post-earthquake road networks. *Transportation Research Part E: Logistics and Transportation Review*, *123*. https://doi.org/10.1016/j.tre.2019.01.009

Zhu, Z., Zhang, A. & Zhang, Y. (2018). Connectivity of intercity passenger transportation in China: A multi-modal and network approach. *Journal of Transport Geography*, *71*. https://doi.org/10.1016/j.jtrangeo.2017.05.009

Zhu, Z., Zhang, A. & Zhang, Y. (2019). Measuring multi-modal connections and connectivity radiations of transport infrastructure in China. *Transportmetrica A: Transport Science*, *15*. https://doi.org/10.1080/23249935.2019.1642966

Zhu, Z., Zhang, A., Zhang, Y., Huang, Z. & Xu, S. (2019). Measuring air connectivity between China and Australia. *Journal of Transport Geography*, *74*. https://doi.org/10.1016/j.jtrangeo.2018.12.017

A

Appendix A: Average Transfers per Node

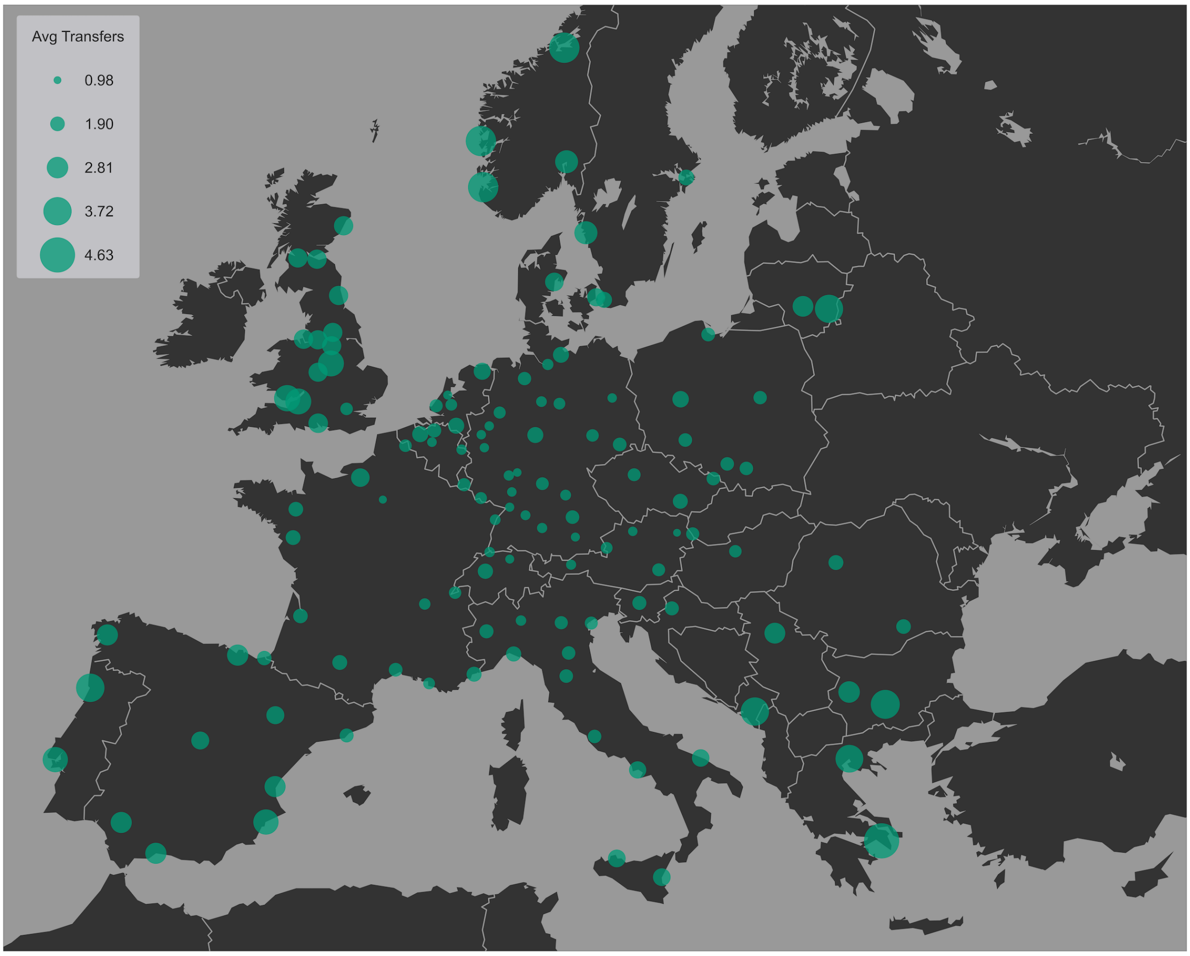


Figure A.1: Spatial Distribution of the Average Transfers per Node within the Rail Network

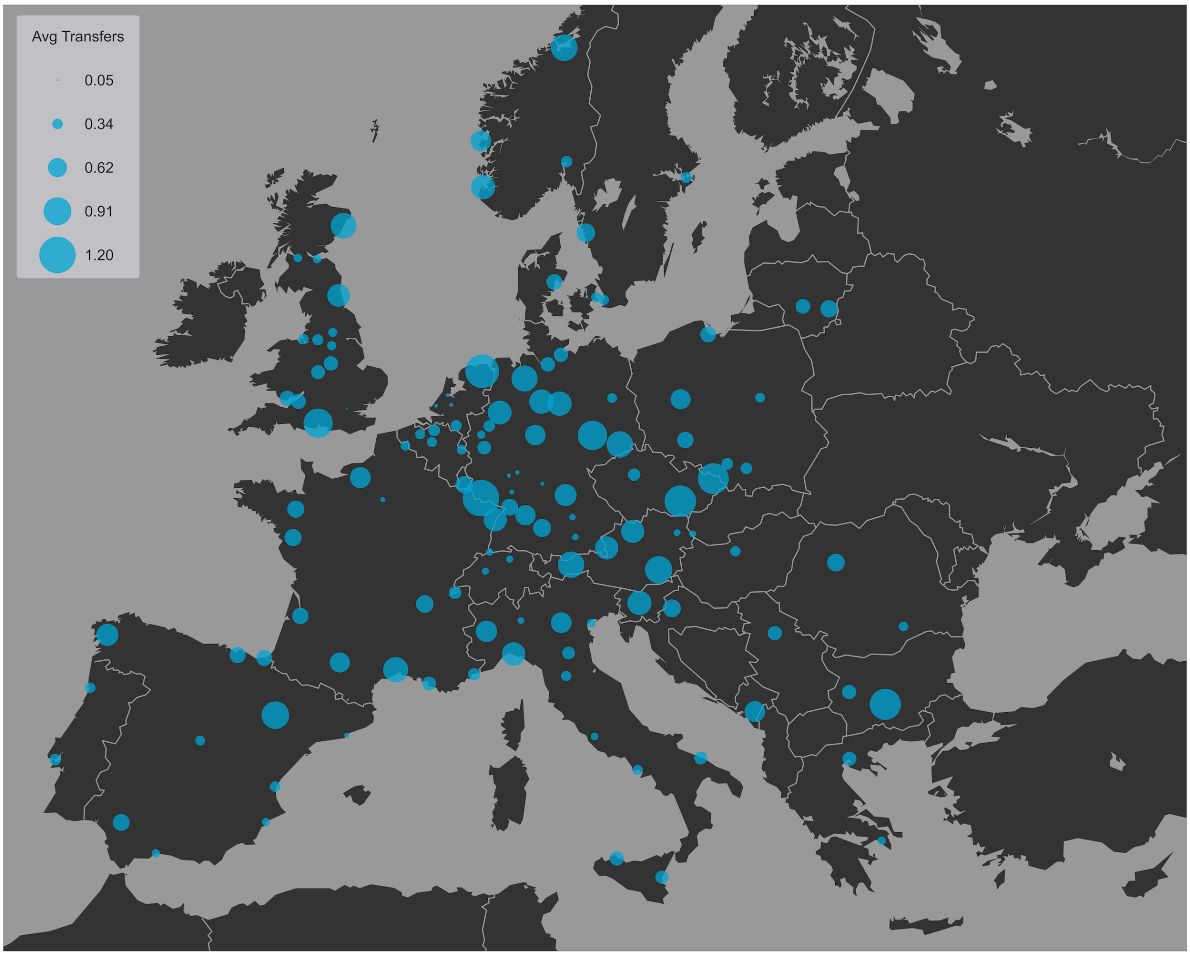


Figure A.2: Spatial Distribution of the Average Transfers per Node within the Air Network

B

Appendix B: Local Node Connectivity

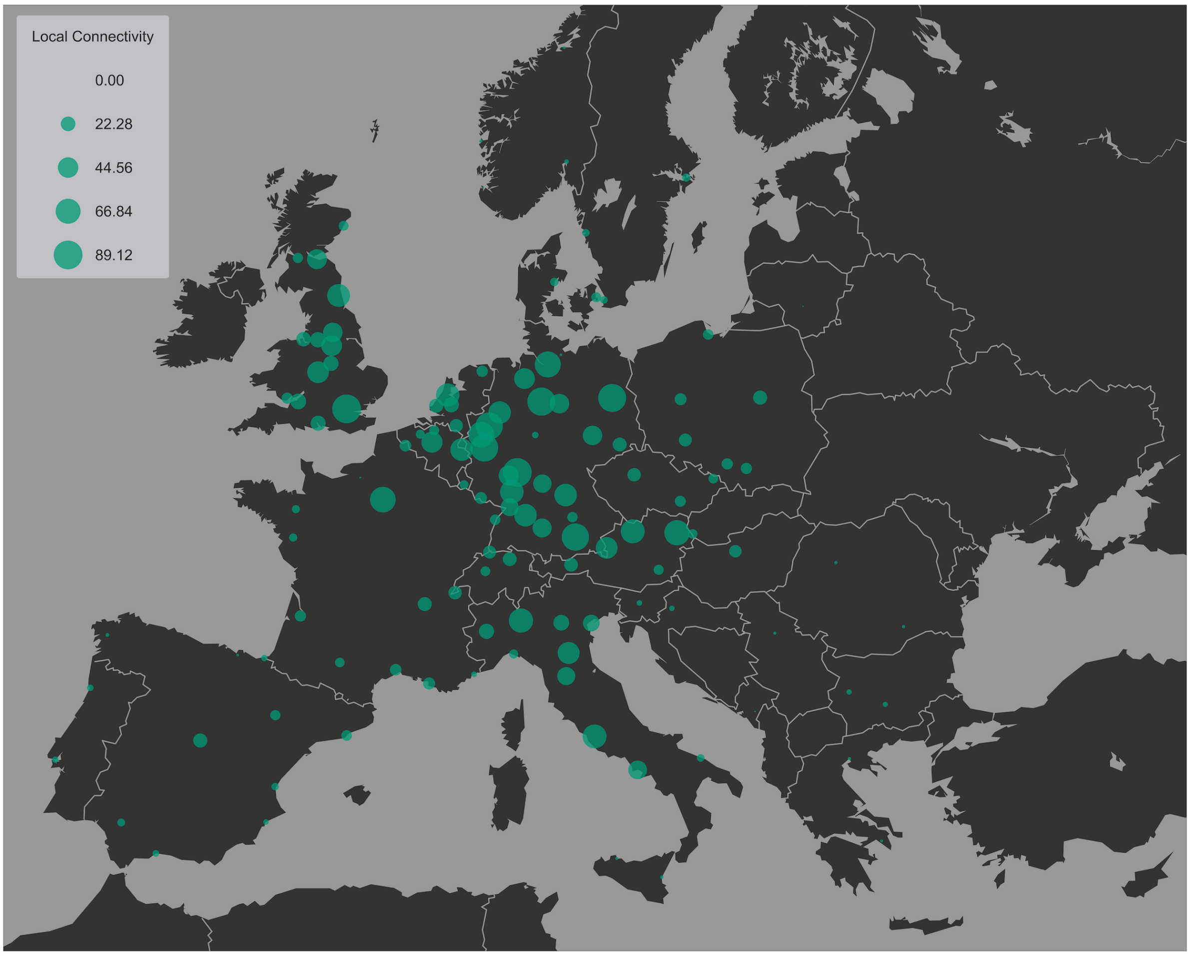


Figure B.1: Spatial Distribution of the Local Node Connectivity within the Rail Network

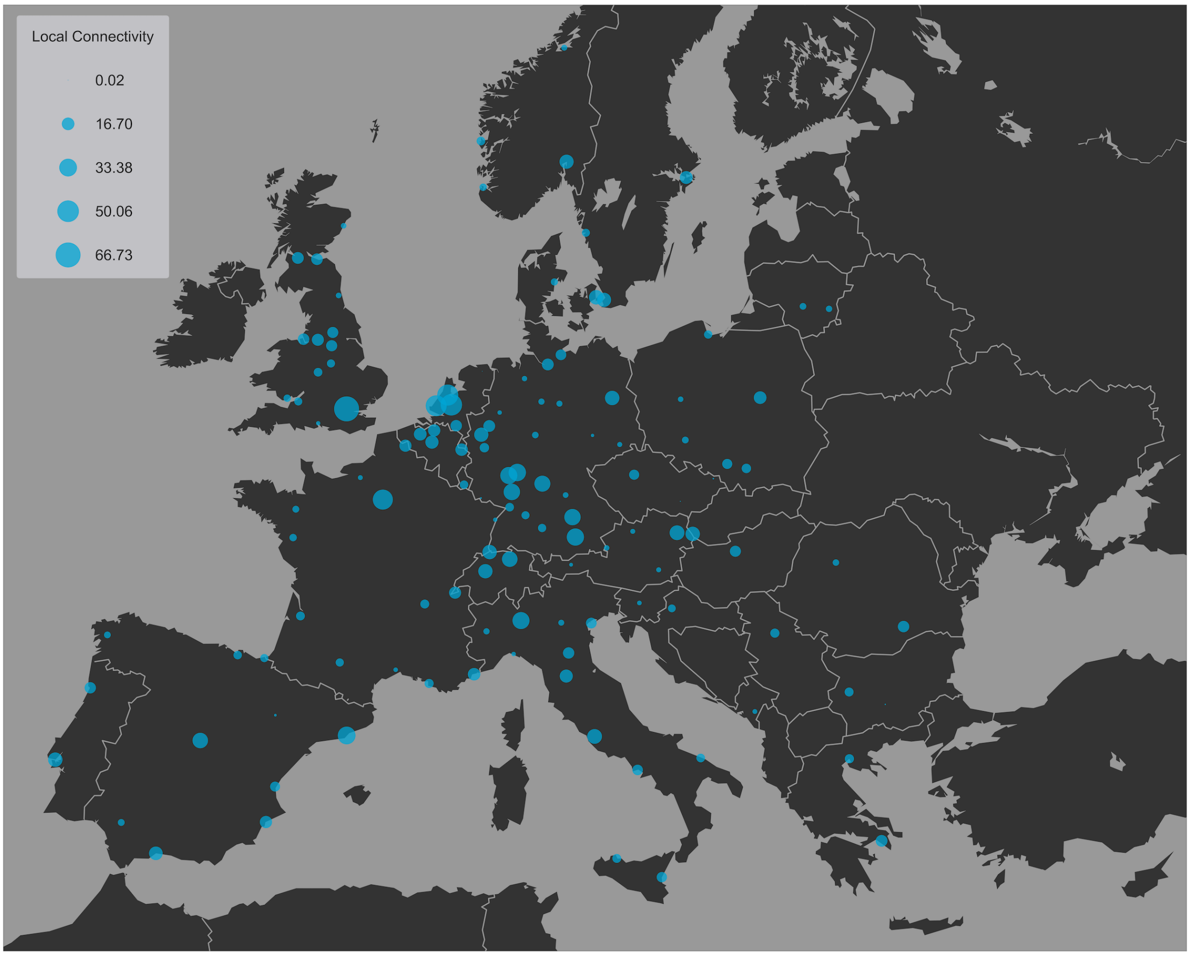


Figure B.2: Spatial Distribution of the Local Node Connectivity within the Air Network

C

Appendix C: Attractive OD pairs for Substitution

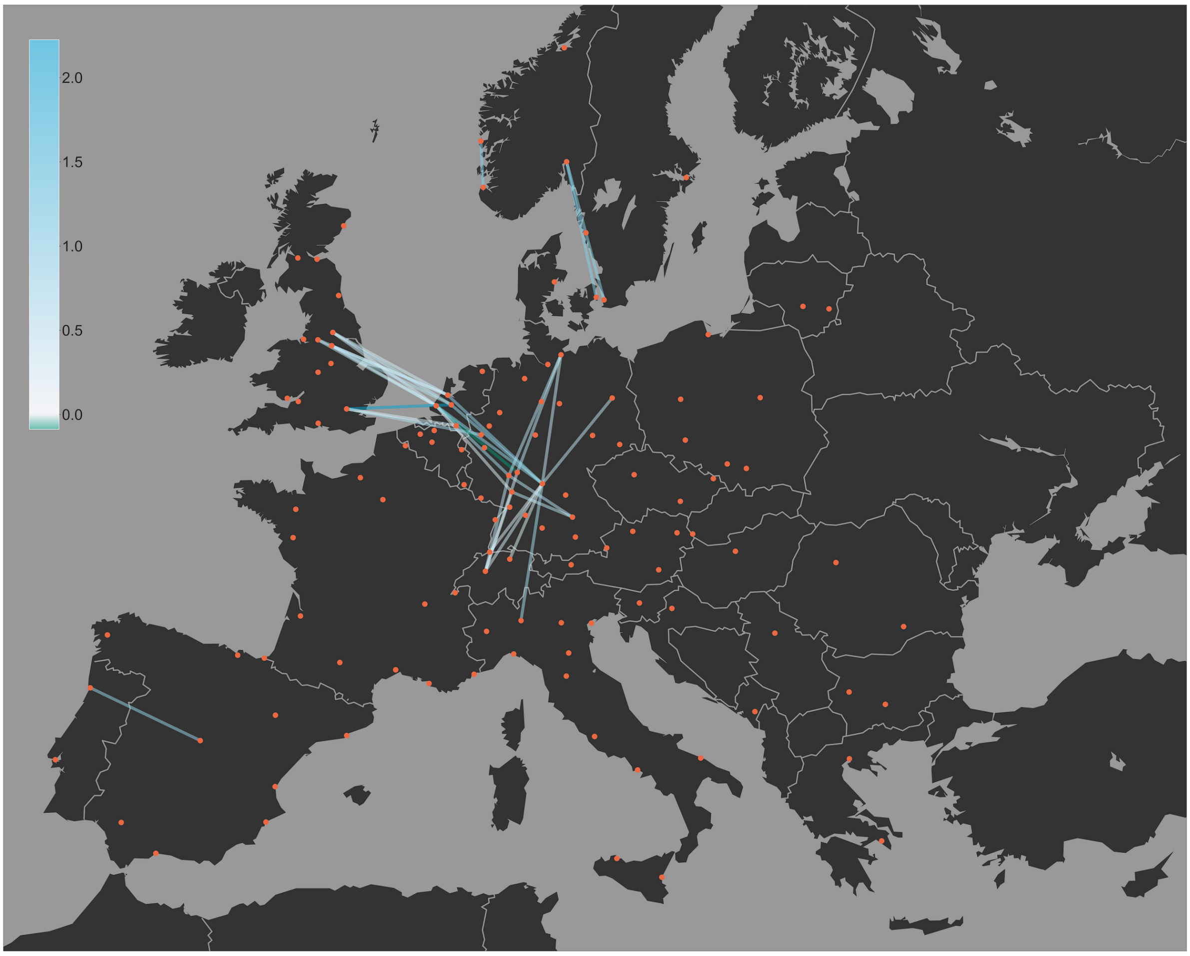


Figure C.1: Spatial Distribution of the 100-500 km Links Attractive for Substitution

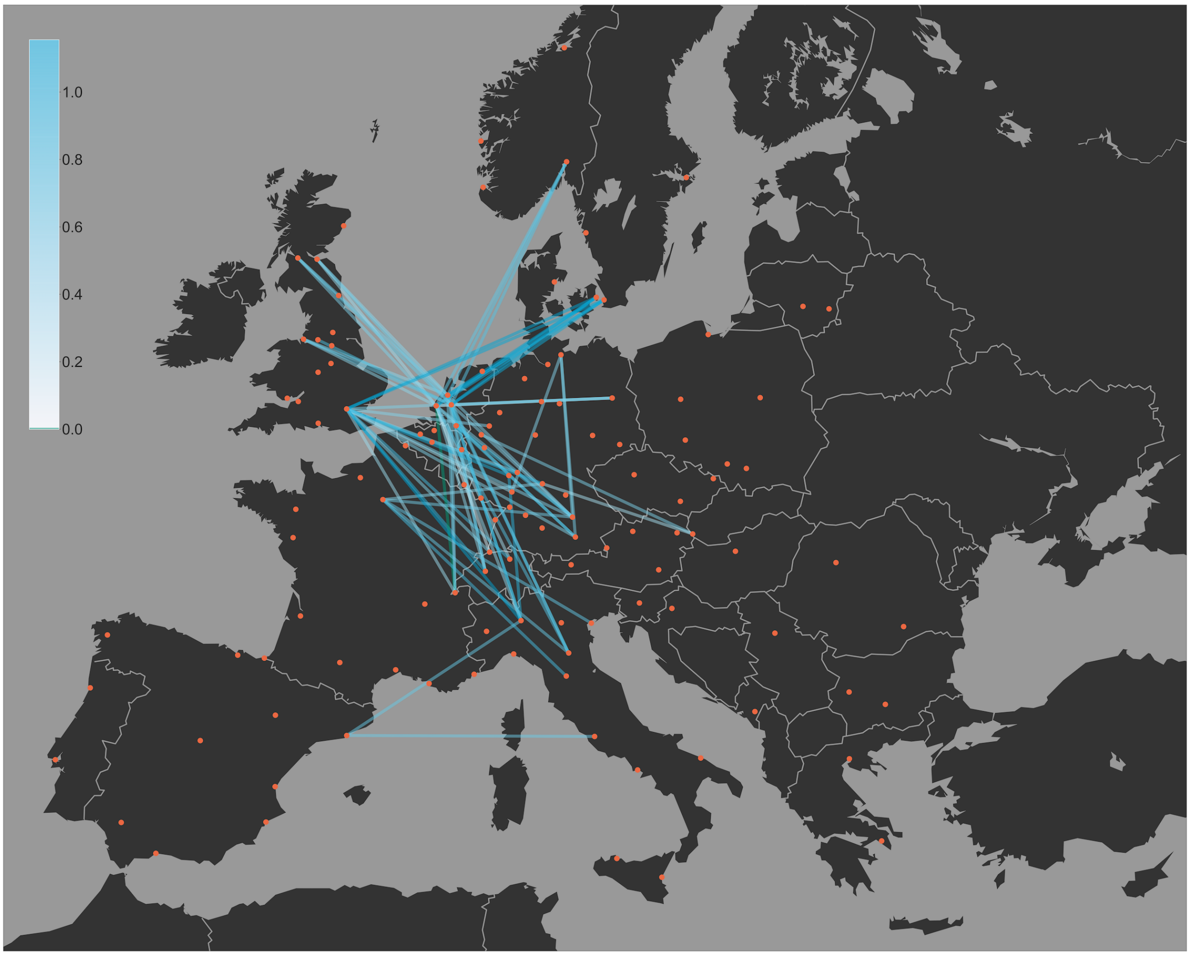


Figure C.2: Spatial Distribution of the 500-1.000 km Links Attractive for Substitution

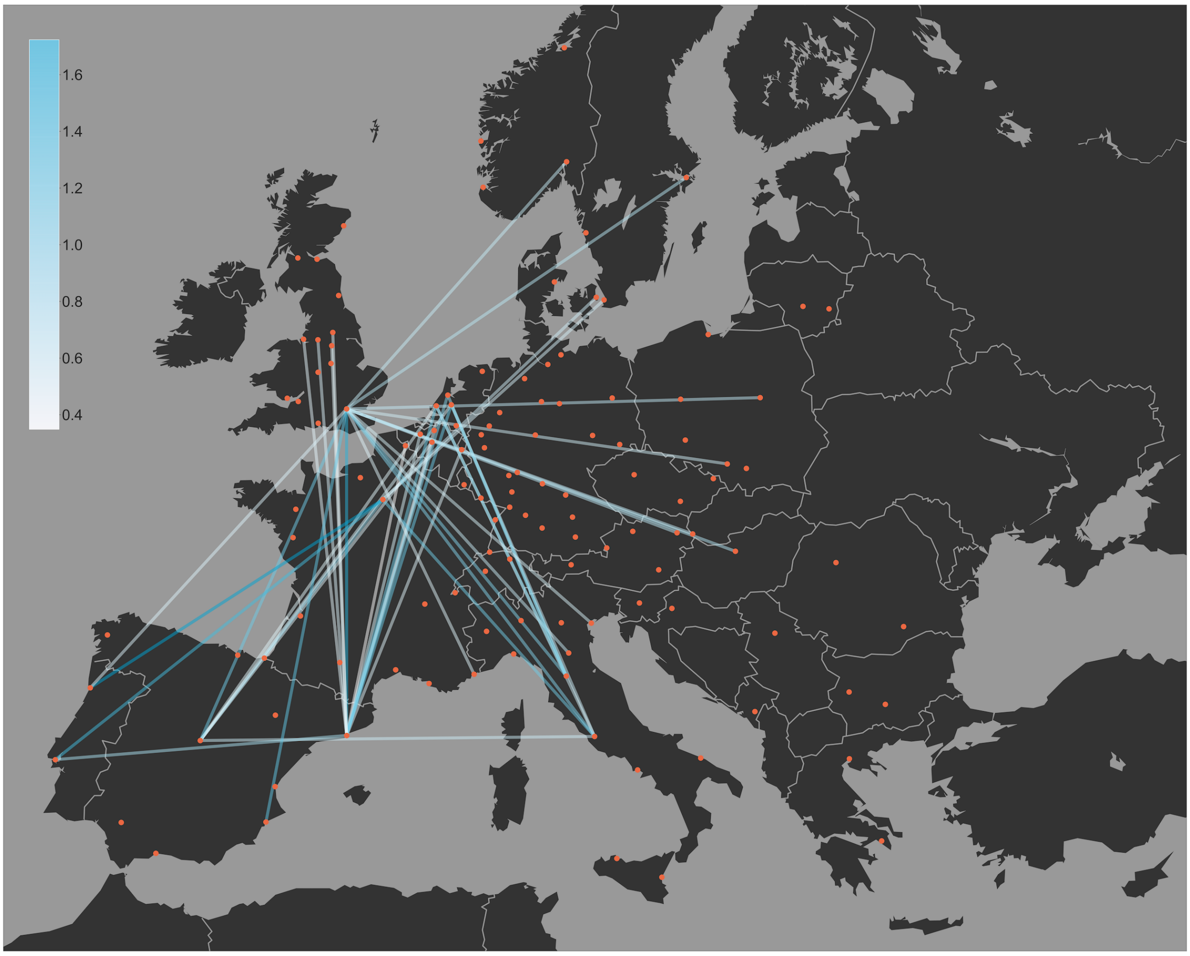


Figure C.3: Spatial Distribution of the 1.000-1.500 km Links Attractive for Substitution

D

Appendix D: Scientific Paper

The Connectivity of the Long-distance Rail and Air Transport Networks in Europe

Francesco Bruno

Department of Transport and Planning, Faculty of Civil Engineering and Geosciences (CEG), Delft University of Technology, Delft, the Netherlands

Keywords: Connectivity, Long-distance travel, Air to rail modal shift, European transport network, Substituitability, Competitiveness

Abstract

Demand for long-distance transport has been steadily increasing over the last years, worsening the environmental impact of the transport sector. Despite European governmental bodies and policy-makers directed considerable efforts towards the substitution of short-haul flights with greener rail alternatives, the desired results do not appear to have materialised yet. This study contributes to the literature by providing a broad overview of the current state of rail and air service supply in the European market and identifying the most significant nodes and links within the network. Furthermore, the connectivity index proposed by this research allows to compare the performances and the attractiveness of the two modes. Overall, this study shows that air efficiently connects most major European cities, whilst, in the rail network, considerable inequalities in terms of performances across different geographical locations are still present. From this picture, derive two main conclusions. First, the rail network has substantial limitations and constraints on longer distances, and second, the sector is more fragmented and less cohesive than air. To minimise the environmental impact of long-distance transport, it appears beneficial to focus on capturing the demand for connections shorter than 500 km by fostering inter-modality and air-rail integration, enhancing rail performances and opening new competitive direct routes on OD distances longer than 500 km. Currently, rail appears to be more complementary to air, rather than competing with, as the inter-modal competition is limited to shorter routes. However, the results suggest that rail has the potential to be competitive with air also on longer routes, either by offering more direct services, improving schedule coordination, or increasing the supply of high-speed and night train services. In this regard, it is worth noting that more attention should be devoted to analysing the specific infrastructural gaps and the optimal directions for investments. Finally, this paper identifies the main bottlenecks hindering the capacity of rail to compete and substitute air, suggesting that these barriers are not limited but go beyond the network performances and service supply perspectives.

1 Introduction

In the last decades, the demand for long-distance transport [1] has exponentially grown all over the globe, and especially in Europe (World Tourism Organization, 2018). And this trend is not projected to cease, as literature generally agrees in saying that an increase in the share of medium- and long-distance trips in Europe is to be expected (Limtanakool et al., 2006). Despite figures having suddenly dropped in 2020 due to the advent of the Covid-19 pandemic, they appear to have rapidly bounced back, as foreseen by Gudmundsson et al. (2021). In particular, the future increase in demand within the European long-distance market appears to be inter-sectional, encompassing different trip purposes, such as tourism and commuting, among others. The WTO (2018) projects arrivals in EU destinations from European source markets to grow by 1.9% a year on average through 2030. Furthermore, the faster and more frequent transport connections and the increasingly flexible work conditions have extended the maximum commuting distance from the urban level to the regional and even national ones, as highlighted by the growing size of these markets (Eurostat, 2019). These figures are the outcome of the considerable infrastructural investments, the radical market changes and the evolution of travel behaviours that have reshaped the entire transport sector over the last decades of the twentieth century. This revolution, providing a more "democratic" and widespread access to long-distance transport, made travel across the continent increasingly easy, opening to a world of new

---

[1] Long-distance transport is a versatile term that does not have a clear, univocal, and truly comprehensive definition. In general, the literature defines it either by taking into account the distance between origin and destination (OD) or the length of the trip in terms of overnight stay (Gerike & Schulz, 2018). Following the former approach, within this study travel is defined to be long-distance when it involves trips longer than 100 km, as many researchers and the European Commission employ this threshold (Frei et al., 2010; Limtanakool et al., 2006; Malichová et al., 2022; Petersen et al., 2009; Rich & Mabit, 2012). On the other hand, the upper bound is set to 1.500 km representing the maximum distance where rail might be able to compete with air, following Givoni et al. (2012).

possibilities. However, the consequent boost in traffic and movement of passengers and freight has created a wide range of problems and negative externalities, which have been further exacerbated by the extensive dominion of fossil fuel modes. In fact, the two protagonists and key players in this revolution have been road and air transport, whilst rail has been a victim of these trends seeing its passenger market shares plummeting from the 1970s until the 2000s when they have stabilised around 6% (Di Pietrantonio & Pelkmans, 2004). In 2018 almost 70% of the private trips across the European Union were made by car, 14% by air and only 9% by rail (Eurostat, 2019). As a consequence of these trends, the transportation industry today has to face its increasing limitations in terms of service capacity and environmental impact.

Whilst greenhouse gas (GhG) emission levels are generally declining, the emissions caused by the transport sector are still on the rise. Eurostat (2021b) highlights that in 2019 the GhG emission in the European Union, despite recording a 24% drop compared to the overall levels of 1990, have increased by more than 30% within the transport sector, which now accounts for over 25% of the total GhG emissions of the EU (Eurostat, 2021a). In particular, the long-distance market is widely regarded as the main source of transport-related GhG emissions, highlighting its crucial role in reducing the environmental impact of the entire transport sector (Rich & Mabit, 2012). As a reference, in 2009, trips longer than 100 km, despite representing only 2.5% of all the European trips, accounted for 55% of the passenger-km travelled across the continent (Petersen et al., 2009). The worsening of the climate crisis over the last years and the consequent raising of environmental concerns have triggered the response from public organisations, which saw rail as a possible solution. As highlighted by the European Commission (2020), rail represents one of the most environmentally friendly and energy-efficient transport modes available in the long-distance market, featuring consistently lower GhG emissions per passenger-km compared to the two main competitors on long-distance transport, rail and air (Fraunhofer ISI & CE Delft, 2020). Thus, to tackle this issue the European Union has directed considerable efforts towards enhancing the attractiveness of the former, to increase its modal shares at the expense of the latter two transport modes European Parliament and Council of the European Union (2020). A notable example is the Fourth Railway Package, approved in 2016 and implemented in 2020, which, projecting towards the creation of a Single European Railway Area, aims to revitalise the rail sector and to increase its competitiveness with other modes, such as road transport on the short-medium range (0-400 km) and air transport on the medium-long range (400-1000 km) (European Commission, 2022).

Researchers generally converge in deeming high-speed and night trains capable of becoming a sustainable alternative to short-haul flights on distances up to 1.000-1.500 km if appropriate and thorough political support is provided (Li et al., 2020; Seidenglanz et al., 2021; Serafimova et al., 2022; Witlox et al., 2022; Zhu et al., 2018). Considering that the average flight distance in 2020 for intra-continental European flights is 981 km, it is clear that rail could potentially compete with air services on a considerable share of routes (Eurocontrol, 2020). However, despite the robust and comprehensive set of measures and initiatives taken by the European Union, especially in terms of the harmonisation of technical systems and operational regulations, and in relation to the opening up of the international rail market to competition, data shows that the modal shift from air to rail is still not taking place at the desired pace. Currently, a large majority of the long-distance market's passenger-km still consists of road and air transport (Malichová et al., 2022). At the same time, the European Environment Agency (EEA) points out that over the last decades the passenger-km of the aviation sector has risen much more sharply and steadily in comparison to the rail in the European market (European Environment Agency, 2021). Many scholars have investigated the reasons behind the underwhelming results that are preventing rail from becoming a key player in the European long-distance transport market. However, it is currently not clear what is the actual potential of the modal shift between rail and air in terms of competitiveness and substitution and how it varies across the continent.

To this hand, this paper aims to fill this research gap by reconciling scientific and academic cutting-edge research with practice, by applying the theoretical connectivity paradigms to the European case study. The contribution of this study is twofold, including both an empirical and a methodological facet. In regards to the former, it categorises European origin-destination (OD) pairs in terms of competitiveness and substitutability, assessing and comparing their performances through a connectivity indicator. In regards to the latter, this study proposes a novel and reproducible methodology to measure connectivity and identify critical links in transport networks, ensuring the comparability between modes. Thus, the overarching goal of this study is to provide a comprehensive comparative overview of the current performances of the rail and air service networks in the long-distance European market. The added value of this study is twofold. First, approaching the matter using a system-wide approach that considers the entire continent, complements the results of studies on limited geographical scopes (e.g. Kroes and Savelberg, 2019 and Avogadro et al., 2021). Second, including specific metrics related to the service supply performances by integrating transport system characteristics, service supply data and passenger behaviour, this paper allows to provide further insights into the actual potential for competition and substitution of routes complementing previous analyses, such as the one from Givoni et al.

(2012). Finally, this study aims to reflect on the continental trends and patterns rising some critical issues for future research and paving the way for the formulation of further policy reflections and the quantification of competitiveness and substitutability potential.

The remainder of this paper is organised as follows. Section 2 presents a brief overview of the previous research on competitiveness and substitutability potential and connectivity, framing this study within the literature. The model setup, data collection process and methodology are specified in Section 3. The results are, thus, presented and discussed in Section 4. Finally, Section 5 concludes the study by summarising the key findings, highlighting the limitations and providing suggestions for further research.

2 Related Research

Nowadays, rail is increasingly taking centre stage in the transport debate, attracting the attention of many scholars and policymakers. The considerable developments of the European rail sector over the last decades have radically revolutionised its role within the wider transport context, especially within the long-distance market (Seidenglanz et al., 2021). In fact, governmental bodies, such as the European Commission, saw this as a possible solution to the recurrent and persistent environmental issues embedded in the structure of the current transport system, pushing towards the substitution of short-haul flights with rail alternatives. In an effort to shed light on the feasibility of such a project and the possible challenges that this process could bring about, the research on the topic has been pointed towards two main directions, following two diverging approaches. On the one hand, some efforts have been targeted at identifying the crucial barriers and bottlenecks that are still hindering this process (Bergantino et al., 2015; Bergantino & Madio, 2020; Clewlow et al., 2014; Dobruszkes et al., 2011; Serafimova et al., 2022; Witlox et al., 2022). On the other, further endeavours have been directed at assessing the eventual potential for competition, substitution and cooperation between the two modes (Adler et al., 2010; Albalate et al., 2015; Avogadro et al., 2021; Behrens & Pels, 2012; Beria et al., 2019; Givoni et al., 2012; Román & Martín, 2014; Socorro & Viecens, 2013; Zhang et al., 2019). The former efforts focus on allowing to remove the obstacles that are still preventing the aforementioned modal shift, pushing towards a more rail-centric long-distance transport market and revitalising the momentum of this process. Conversely, the latter endeavours attempt to more radically question if, and under which conditions, railways are actually able to compete or cooperate with air transport in this specific market. Overall, literature appears to generally share a rather positive outlook on the potential of rail, and especially high-speed rail, to compete with air transport (Adler et al., 2010; Behrens & Pels, 2012). Nevertheless, researchers converge in deeming rail to be currently partially unable to compete on a level playing field with its competitors (e.g. road in the medium-distance and air in the long-distance market), due to its weak “modal competitiveness” and to the many barriers that are still preventing rail to be truly attractive for passengers (Martín et al., 2014; Román & Martín, 2014; Witlox et al., 2022). Concerning the substitutability of air routes with a rail alternative, Avogadro et al. (2021) foresees that, due to the recent adoption of policies aimed at reducing the emissions of the aviation sector, the cancellation of all the air routes where an effective alternative exists is likely to be forthcoming. In particular, the authors estimate at 26.5 million (or 3.02% of total intra-European) the currently offered air seats that could be replaced by rail assuring similar travel times for travellers (Avogadro et al., 2021).

Despite the literature having widely analysed the possible causes for the failed expectations, identifying the factors influencing the modal competitiveness of rail in the current European scenario through a diverse set of perspectives, the network dimension of the problem appears to require more attention. In particular, it is not clear what is the actual potential of rail to compete with air and between which OD pairs. Some researchers, such as Givoni et al. (2012) and Kroes and Savelberg (2019), have partially filled this gap, the former employing a GIS-based methodology to examine the potential for air-rail substitution and the latter assessing the potential for high-speed rail to substitute air services at Amsterdam Schiphol Airport using a forecasting model based on a modal split model and demand growth factors. However, both studies have a few important limitations. The former manages to capture the broader geographical context, providing a global overview of the routes with potential substitution. However, considering only the geographical distance and travel demand the study fails to consider the service supply attributes and performances on those corridors, which represents a fundamental aspect to assess substitution and competition potential (Avogadro et al., 2021; Luo et al., 2019; Zhu et al., 2018). The latter study, on the other hand, manages to take into account service supply but, focusing on a limited scope, does not allow to capture the continental dimension of the problem, which is particularly important, especially in light of the EC goals in terms of the creation of a single railway area (European Commission, 2022). In this regard, Clewlow et al. (2014) highlights the paramount importance of considering wider trends using a system perspective in developing policies aimed at reducing the environmental impact of the transport sector. Thus, this study aims to bridge this gap by providing a broad overview of the current state of the long-distance European market, highlighting the most critical links in terms of competitiveness and substitution and providing a reproducible methodology to compare the performances of rail and air service supply.

To capture and aggregate the performances of transport service supply, connectivity [2] indicators are often used across the literature and within the industry, especially in the aviation sector (Burghouwt & Redondi, 2013). On the other hand, a thorough literature review has revealed a poignant scarcity of studies that assess the connectivity of rail transport systems, with a few exceptions, such as the study from Xu et al. (2020). Comparative research on the connectivity of rail and air appears to be even more limited. Nevertheless, Zhu et al. (2018) and Zhu et al. (2019) developed an interesting connectivity index for the Chinese rail and air networks that can be employed to compare the performances of two modes and to understand the role of nodes and links within each network and in the entire transport system. This highlights a second research gap, which relates to the fact that most research in the field has focused on the Chinese case study, whilst the connectivity of the European rail network has neither been quantified nor thoroughly studied yet, as also pointed out by the European Commission (European Parliament & Council of the European Union, 2020). Furthermore, no study appears to have bench-marked rail and air connectivity in an effort to provide insights into the competitiveness between the two modes from a network perspective, identifying the role of different routes and their performances within the network. Consequently, this study aims to fill this gap by proposing a connectivity indicator for the European case study that could be employed by researchers and policy-makers to assess the state of specific long-distance routes and of the market as a whole. The added value of an aggregated connectivity metric relates to its capability to communicate complex dynamics with clarity and simplicity, which assumes a crucial role in practice and governance. In particular, (Zhu et al., 2019), highlight the importance of connectivity to efficiently direct investments in systems where resources are scarce and limited, such as the rail system. This assumes particular importance considering the substantial investments that the European rail sector is going to receive in the foreseeable future (Avogadro et al., 2021).

3 Data and Methods

3.1 Model Setup

The geographical scope of this research includes all the countries within the European continent, connected to the European international rail network. Thus, the three insular EU countries (i.e. Ireland, Cyprus and Malta) are excluded, whereas a number of European countries not part of the EU but still connected to the TEN-T rail networks are included (i.e. the UK, Ukraine, Belarus, Moldova, Serbia, Bosnia-Herzegovina, Albania, North Macedonia, Montenegro and Kosovo). Although most of them and three EU countries (i.e. Finland, Estonia and Latvia) are excluded at a later stage due to data availability. In particular, to allow direct comparability between rail and air the analysis is developed at the urban area [3]. Urban areas are chosen in place of regions or grid-based systems following the assumption that, in order to be practically sustainable, international long-distance services require a considerable demand, which can only be assured by major urban centres with specific characteristics, in terms of population, GDP, labour market attractiveness, touristic attractiveness, and intra-EU migrations. In fact, the supply of international long-distance services, especially in the aviation and high-speed markets, is generally liberalised and not subject to PSOs, with a few exceptions only (i.e. mostly services used to serve remote and infrastructurally isolated areas such as islands). The aforementioned criteria are employed to define the urban area attractiveness in terms of long-distance international transport and to select which to include in the analysis. Population and GDP aim to capture the general production and attraction of international cross-border trips, following the assumption that demand levels are higher in urban areas with more population and higher GDPs (Zhang et al., 2019; Zhu et al., 2018). Furthermore, labour market attractiveness, touristic attractiveness, and intra-EU migrations aim to capture the specific travel behaviour for long-distance trips by reflecting the three main trip purposes as defined by Rich and Mabit (2012). Labour market attractiveness measures the intensity of business trips, touristic attractiveness the intensity of holiday trips, and intra-EU migrations the intensity of personal trips, following the assumption that most passengers moving internationally within Europe for personal reasons are expats.

Following the literature review, the factors determining and influencing the modal competitiveness between rail and air are identified and summarised in Figure 1, a causal loop diagram that highlights the system dynamics and the relationships between different elements. It was decided to exclude travel costs due to their inherent volatility and dynamic nature. In fact, prices for long-distance travel tend to fluctuate

---

[2] Connectivity appears to be a multi-faceted term, acquiring precise meanings depending on the specific context into which it is employed. The common thread that appears to connect the many definitions and metrics employed across the literature is that connectivity aims to measure the strength or degree wherewith two elements are connected.

[3] Dijkstra et al. (2019) highlight how comparing cities on an international level can be extremely challenging, as the definition of city widely varies across different countries, even within the EU. Whilst it might be argued that the concept of the urban centre as a node is rather clear, it is much more problematic to establish the actual extent of the urban area (Weeks, 2010). Thus, to retrieve comparable datasets and avoid biases urban areas are defined within this research as "densely inhabited urban centres and their surrounding and interconnected lower-density areas", reflecting the definition of metropolitan areas given by Moreno-Monroy et al. (2021).

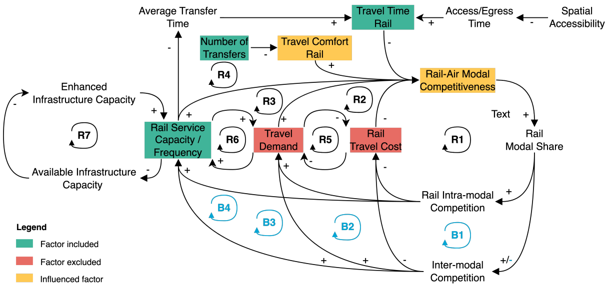


Figure 1: *Factors Influencing Rail Modal Competitiveness*

considerably based on many different factors, such as advance purchase and travelling period (i.e. seasonality). Furthermore, given that this study approaches the problem from a service supply perspective, travel demand is also not directly modelled. The frequency component in this study is included as weekly frequencies. The decision to have a weekly time frame rather than a daily one refers to the inherent characteristics of long-distance transport, where service scheduling and planning are generally developed on a weekly rather than daily basis as the frequencies are rather low, in some extreme cases sinking to one or two weekly connections. On the other hand, travel comfort is not included as assessing the quality of transfers, due to the extent of the scope, was not deemed practically feasible. Although it can be argued that also this component is, at least partially, considered as the number of transfers/transfer time, which directly influences travel comfort, is included. Finally, travel time within this study is modelled as door-to-door travel time, as illustrated in Figure 2. This, in fact, is the travel time measure upon which passengers generally rely in order to make their travel decisions (Peer et al., 2013). To further simplify the model some components are fixed to standardised values based on practical experience and on the recommendations provided by airports, airlines, train stations, and train companies. The parameters are then checked and refined based on the assumptions made by different studies across the literature such as Avogadro et al. (2021) and Zhang et al. (2019). In regards to access/egress times, the values are computed manually for each airport-urban area pair using Google Maps to take into account the average traffic conditions (the fastest access/egress option between PT and car is selected). On the other hand, rail stations being generally located in highly accessible areas, oftentimes in the proximity of city centres, imply that a fixed 20 min access/egress time can be safely assumed. Waiting time at departure for air is fixed at 120 min to allow 90 min for check-in, luggage loading, security controls and boarding procedures and further 30 minutes for eventual congestion and to compensate for the longer times required in many cases to access airports (access/egress times are computed based on the fastest modal alternative). The waiting at departure for rail is fixed at 25 minutes, a time which is generally suggested as the standard time to reach stations when travelling with international trains within Europe. Waiting time at arrival is assumed to be 5 min for rail, considering the time to disembark the train and exit the station and 30 min for air as the processes are longer due to the bigger size of terminals and the longer disembarkment and luggage unloading procedures. Finally, in-vehicle times are obtained from actual data of scheduled transport services, as further explained in Section 3.2. For indirect connections, the shortest weighted path in terms of in-vehicle time is substituted for the in-vehicle time of the direct connection. This implies that transfer times are not directly included within the in-vehicle time components. Thus, an additional transfer time parameter is added for indirect connections. The in-vehicle time is weighted by 0,44 for night train, 0,75 for day train and 1 for plane to take into account passenger's preferences and travel time sensitivity based on the studies from Román et al. (2010) and Heufke Kantelaar et al. (2022), further adding to the consideration of the travel comfort component.

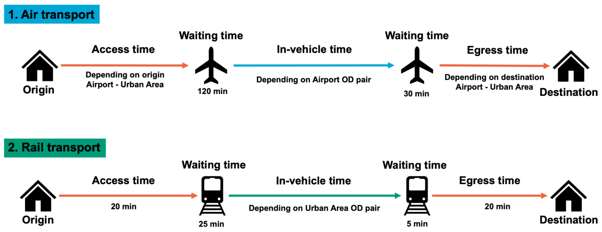


Figure 2: *Overview of the Travel Time Composition per Mode*

3.2 Data Collection

Following the identification of the urban areas, each is associated with accessible air and rail terminals. To select the accessible airports 90 min or 1,5 h isochrones are employed as catchment areas, following Poelman (2013), European Commission (2013), Malina (2010) and Augustyniak and Olipra (2014). On the other hand, major rail stations, providing intercity and long-distance services within the city boundaries are selected in the latter case. Thus, the weekly frequency and average travel time between each possible terminal OD pair are retrieved separately through air and rail service schedules. For the former, the data has been manually collected through the use of the publicly available real-time aviation data provider Flight Aware. In particular, the flight schedules between each airport OD pair are retrieved through its "search by route" engine. This tool reports all the flights departing from a certain origin airport and arriving at a certain destination airport for a week time, from five days before to two days after the current date, in the case of this analysis July 22nd 2022. To make the data collection process faster and less cumbersome a list with a limited number of airports was defined based on the requirements of this specific research, meaning that not all European airports have been included. For each origin airport, the data has been collected by iterating over all the possible destinations included in the airport list using the ICAO codes of the two airports. Despite the output data being pre-filtered using the filters made available by Flight Aware, the data is also manually cleaned to ensure that only commercial services are included. Finally, each unique terminal OD-pair is grouped, averaging the in-vehicle time and counting the number of occurrences of that OD-pair in the dataset (weekly frequencies of flights). On the other hand, the data used to create rail networks has been collected through the UIC MERITS [4] (Multiple East-West Railways Integrated Timetable Storage), a database containing the integrated timetable data of many European countries which are used to provide information to journey planners and ticket-selling websites. The decision to use this commercial database, rather than manually retrieve the data from freely accessible journey planners, relates to the many inconsistencies and considerable discrepancies detected in the data available throughout the different journey planners. Notable journey planners that have been queried and whose results have been compared include DB (Deutsche Bahn), Trainline, Rail Europe, Eurail/Interrail, Omio, Rome2Rio, and ÖBB (Österreichische Bundesbahnen). The data collected includes all the train services available between each station OD pair in the week between June 27th 2022 and July 3rd 2022. Furthermore, data for the previous week, from June 20th 2022 to June 26th 2022 is retrieved to check that there are no gaps in the data. Also in this case the data is cleaned to filter out the regional and suburban services, given that the scope of the study is limited to long-distance transport. Due to some gaps in the available data, some cross-border connections have been manually added to ensure that the entire network is interconnected, using data retrieved either through the aforementioned journey planners or using the websites of the national railway undertakings. It is important to note that some of these services have been suspended due to the Covid-19 pandemic and have not yet resumed, whilst others are not direct connections but include a transfer at the border. In these cases, the total travel time from the origin station to the destination station, including the specific transfer time based on schedule synchronisation is employed as the average in-vehicle time.

Finally, the travel times and frequencies between each terminal OD-pair od are re-aggregated to each urban area OD-pair ij. The travel times are computed using a weighted average of all available direct routes r_{od} connecting origin i and destination j through the respective terminals o and d, whereas frequencies are summed over all the available routes. In the case of air, the attractiveness of specific terminals is considered in an effort to capture the influence of different airport characteristics (i.e, size and location/accessibility) on route choice and the inconvenience of using routes that are longer and less frequent. To model this, frequencies on a specific route are used as weights to average in-vehicle time and the ratio of the minimum total travel time and the total travel time of each route is used as a measure to discount the frequencies of longer routes. This implies that among all the direct available routes between i and j, the route r_{od}, using origin terminal o and destination terminal d, with the shortest total travel time will have all the frequencies included, whilst longer routes will add to that only up to an extent. This is particularly useful to penalise routes that use distant terminals when terminals located closer to the urban area are available.

3.3 Methodology

The networks analysed in this study are represented as undirected weighted graphs (network) $G = (N, L)$, composed of a set of vertices (nodes) $N = \{n_1, n_2, ..., n_{|N|}\}$ and of a set of edges (links) $L = \{l_1, l_2, ..., l_{|L|}\}$ using a space-of-service or P-Space topology as described by Luo et al. (2019) and von Ferber et al. (2009). Nodes N represent the 125 European urban areas selected, whilst links L represent the eventual existence of at least a direct connection between each pair of urban areas. The methodology employed to compute link connectivity, for direct and indirect connections, is illustrated in Figure 3. To compute the connectivity of indirect links the shortest paths per mode are computed using the Dijkstra method on networks with link resistance, as per formula 2, employed as weight. Link connectivity LC_{ij} between an origin i and a destination j is defined, within this study, as the ratio of connection intensity

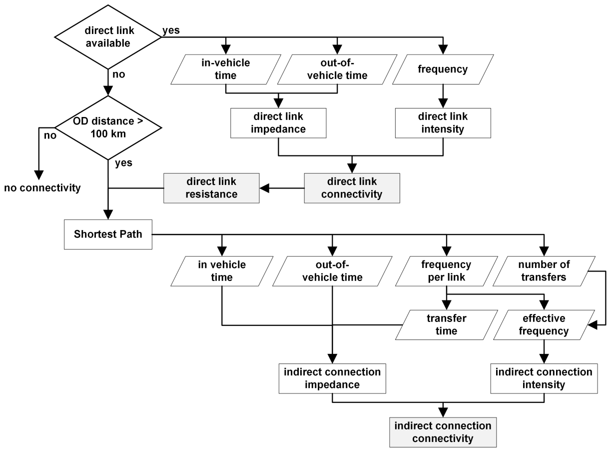


Figure 3: *Link Connectivity Overview*

---

[4] The access to this commercial UIC-owned database is provided by Hacon within the framework of the "Rail connectivity index" project developed in consortium with Ecorys for the Directorate-General for Mobility and Transport (DG MOVE) of the European Commission.

to travel inconvenience, as per Formula 1, where the former is a function of travel time and the latter is a function of frequency.

$$LC_{ij} = \frac{CI_{ij}}{TI_{ij}} \tag{1}$$

Furthermore, link resistance LR_{ij} is defined as the reciprocal of link connectivity, as per Formula 2.

$$LR_{ij} = (LC_{ij})^{-1} = \frac{TI_{ij}}{CI_{ij}} \tag{2}$$

The two components CI_{ij} and TI_{ij} are put into relation using a ratio in order to provide a measure of the relationship incurring between their magnitudes. The use of a ratio allows to enhance the influence of the former component on connectivity, and of the latter on resistance, in line with the literature on the matter, most notably Zhu et al. (2018) for connectivity and Luo et al. (2019) for resistance. The connection intensity component CI_{ij} represents the intensity of the connection between two urban areas and represents a function of the OD frequency, as illustrated by Equation 3. For indirect connections, it is measured by the effective frequency, the ratio of the minimum frequency among all the shortest path's legs to the number of transfers. The number of transfers is included to penalise frequencies of indirect connections compared to direct equivalents.

$$CI_{ij} = \theta_f \, f_{ij} \tag{3}$$

The travel inconvenience component TI_{ij} represents the degree of inconvenience that passengers incur when using a certain link, in case of direct connections, or a series of links, in case of indirect connections. This parameter is measured by three main components, namely in-vehicle time, out-of-vehicle time and transfer time (only for indirect connections), as shown by Formula 4.

$$\begin{aligned} TI_{ij} = \theta_{in-veh} \, tt_{ij}^{in-veh} + \theta_{out-veh} \, (tt_{ij}^{acc} + tt_{ij}^{wait} + \\ + tt_{ij}^{egr}) + \theta_{trf} \, tt_{ij}^{trf} \end{aligned} \tag{4}$$

To estimate transfer time tt_{ij}^{trf} of indirect connections, the formula defined by Sen and Morlok (1976) for intercity public transport systems is employed, as per equation 5. Where K represents the set of transfers on the shortest path between i and j, T is a constant representing the period, whilst f^{inc} and f^{out} represent frequencies on the incoming and outgoing link respectively. The period is set to 7560 to account for weekly, rather than daily, frequencies, and to consider that long-distance services rarely depart or arrive after midnight and before 06:00 in the morning, meaning that 18-hour operational days are considered rather than complete 24-hour days.

$$tt_{ij}^{trf} = \sum_{k \in K} \frac{T}{f^{inc} + f^{out}} \tag{5}$$

It is worth noting that this formula does not always guarantee an accurate estimation of transfer times. Nonetheless, it is employed to give an indication of the magnitude of the travel impedance, assuming that lower frequencies tend to negatively impact the ease of transferring between consecutive services even when transfer times are optimised (e.g. in case of cancellations or delays). Finally, each component is associated with a specific coefficient θ, which represents the degree of influence that each element has on passengers' behaviour for long-distance travel. The parameters considered are not mode-specific as in-vehicle time modal sensitivity is already taken into account in the data set. Furthermore, differences across modes for the remaining parameters (i.e. out-of-vehicle time, transfer time and frequency) are not deemed as marked and are thus neglected. The chosen parameters are derived from ratios between model parameters provided by Moeckel et al. (2015) in the case of travel time and by Wen and Koppelman (2001) for frequency and are illustrated in Table 1. Finally, the node connectivity nc_i represents the node counterpart of link connectivity and aims to provide a broad indicator of the connectivity of a node. It is measured to represent the average connectivity for all direct and indirect available connections from node i, as per formula 6.

Table 1: *Overview of the θ coefficients' value per component*

Component	In-vehicle time	Out-of-vehicle time	Transfer time	Frequency
Parameter	θ_{in-veh}	$\theta_{out-veh}$	θ_{trf}	θ_f
Value	1	2	4	9

$$nc_i = \sum_{j \in N} \frac{LC_{ij}}{N_i} \tag{6}$$

Where, N_i represents the total number of nodes connected with node i, and is computed by subtracting the number of connections with no connectivity with i (i.e. OD distance <= 100 km) to the total number of nodes in the network N. This allows to capture the connectivity degree of a node within the entire network. Higher values of node connectivity imply that the node is well connected to a large number of other nodes.

The methodology employed to categorise direct links is schematised in Figure 4. A first distinction is made between links under a monopoly regime of either rail or air services and links under a competition regime, with both modes operating services on the link. Thus, a second distinction is made within the links under competition, between competitive links and substitution-ready links. The former group identifies all those links over which travelling by air or rail does not imply consistent differences for passengers, whilst the latter identifies all those links where rail could possibly substitute air services. It is worth noting that these two categories are not mutually exclusive and that all the routes are checked for com-

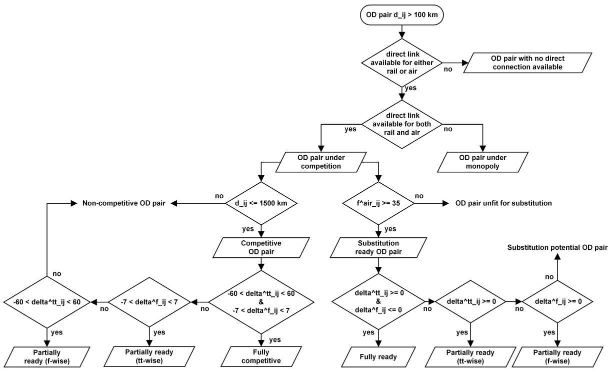


Figure 4: *Link Categorisation Methodology Overview*

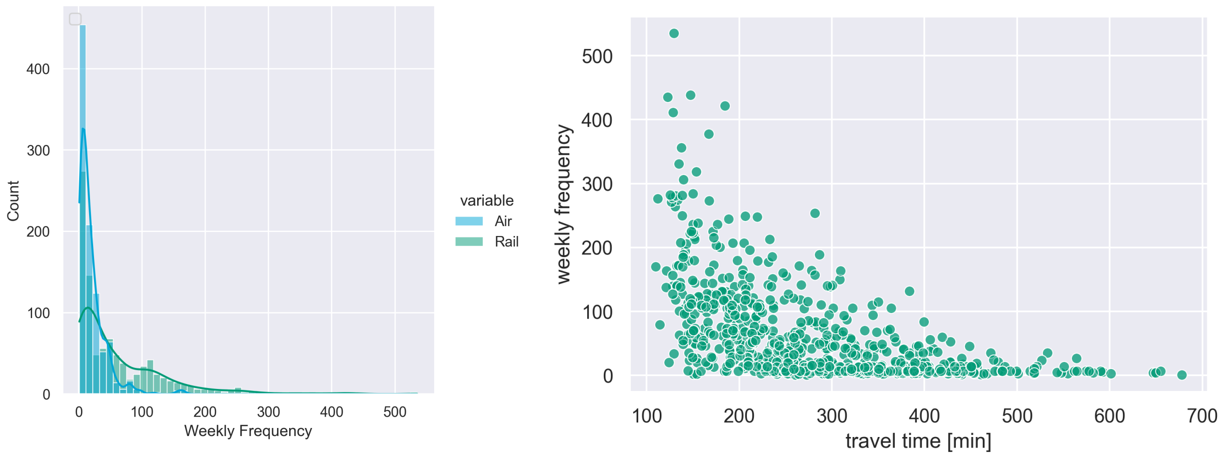
petitiveness and substitutability. Following Givoni et al. (2012), the fundamental condition for competition is that the OD geodesic distance d_{ij} is lower than 1500km, whilst the one for substitution relates to the weekly frequency of air services f_{ij}^{air}, which are required to be greater than 35. The main determinant for substitutability is represented by the level of demand rather than the distance. In particular, the threshold value is taken from the studies of Givoni et al. (2012) and Kroes and Savelberg (2019) which consider as high-demand routes all the connections with more than five daily flights. To evaluate the performances of services on each link the differences in terms of travel time and weekly frequencies between air and rail are computed as $delta_{tt}$ and $delta_f$ respectively. Thus, the thresholds for competitive links have been set to ±60 minutes in terms of travel time and to ±7 weekly frequencies (i.e. ±1 daily frequency). This implies that links are assumed to be competitive when the travel time difference between the two modes is lower than an hour long and the difference in daily frequencies is 1 or less. Furthermore, to capture links ready for substitution $delta_{tt}$ is required to be greater or equal to zero (i.e. travel time for air greater or equal compared to rail) and $delta_f$ is required to be lower or equal to zero (i.e. the number of frequencies for air is lower or equal compared to rail). In particular, for both competitive and substitution-ready links a threefold categorisation is made, between links that fulfil both criteria, defined as fully competitive and fully substitution ready respectively, and links that fulfil only one of the criteria (i.e. either travel time or frequency), defined as partially competitive and partially substitution-ready respectively. The links under competition that do not fulfil any of these criteria are defined as non-competitive links in the former case and links with substitution potential in the latter. Links with substitution potential represent all those OD pairs under a competition regime with a considerable demand (i.e. more than 35 weekly air services), where rail is currently not competitive either from a travel time or frequency perspective. Finally, all the air monopolies with frequent air service connections (i.e. more than 35 weekly air services) on the 100 - 1.500 km market segment not directly connected by rail services are categorised as OD pairs attractive for substitution.

4 Results

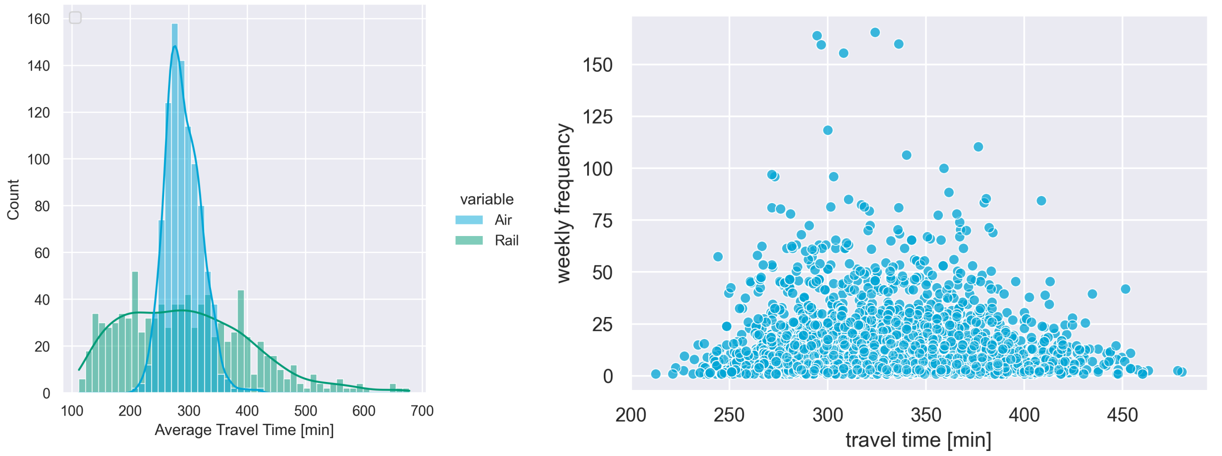
This section presents and discusses the results of the methodology described in Section 3 applied to the European case study. Section 4.1 dives into link connectivity, analysing the characteristics of its components (i.e. travel time and frequency) and the connectivity of all the links across the network. Section 4.2 highlights the relationships between OD pairs in the networks, building on the qualitative analysis of connections in terms of connectivity and capturing the quantity and distribution of connections across the aforementioned categories. Furthermore, the spatial distribution of the OD pairs is also described. Finally, Section 4.3 aggregates the connectivity at the node level, analysing the distribution of this metric across the two networks.

4.1 Link Connectivity

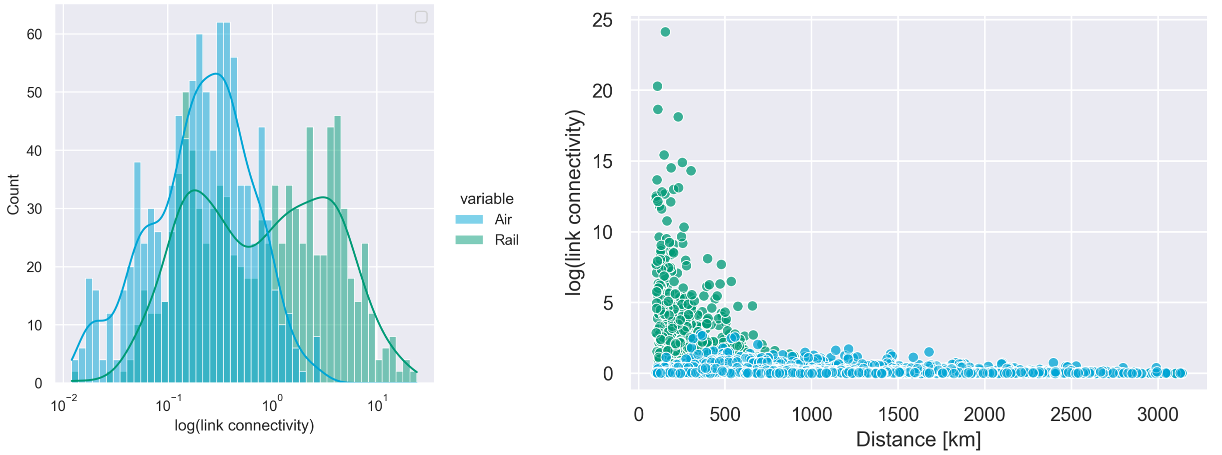
To explore and compare the characteristics of the two networks in terms of travel time and frequency, the distributions of direct travel time and direct frequency are plotted in Figures 5a and 5c respectively. In terms of frequency, the air network shows a considerably higher concentration of low-frequency connections, with a maximum of 165.5 weekly frequency equivalents between London and Glasgow. Rail, on the other hand, has a much wider distribution of frequencies which peaks at 535 on the corridor between Wien and Linz. It is interesting to notice how for both the air and rail networks the strongest connection in terms of frequency is a national route. The distributions of travel times follow a similar pattern. Air travel times are concentrated at around 300 min, with a minimum of 212,5 min between Geneva and Lyon and a maximum travel time of 480 min on the Lubeck-San Sebastián, Barcelona-Kaunas, Glasgow-Thessaloniki and Porto-Bucarest routes. On the other hand, rail travel times are more evenly distributed ranging from a minimum of 112 min between Brussels and Liège to a maximum of 677,5 min (more than 11 hours) between Hamburg and Salzburg. It is worth noting that the trends and patterns identified within this paragraph are strongly influenced by the travel time composition of the two modes. Air total door-to-door travel time is mostly influenced by the fixed out-of-vehicle times, as the high speeds allow for a low variability of travel time across different distances. On the other hand, rail service total travel time mostly depends on the in-vehicle time, which is highly variable considering its dependence on spatial distance due to the lower average speeds.



(a) *Frequency distribution for direct links* **(b)** *Relationship between travel time and frequency in direct rail connections*



(c) *Travel time distribution for direct links* **(d)** *Relationship between travel time and frequency in direct air connections*



(e) *Link connectivity distribution* **(f)** *Relationship between link connectivity and distance*

Figure 5: *Analysis of link connectivity, its components and distance*

Furthermore, some interesting patterns are also noticeable comparing the relationship between travel time and frequency in the rail and air network, illustrated in Figures 5b and 5d respectively. The former features high weekly frequencies (over 300) only for short travel times (under 200) min. The latter, on the other hand, features the highest frequencies for medium travel times, ranging between 275 and 375 minutes. This suggests that the current rail service supply tends to focus on OD pairs reachable within short travel times. In fact, with increasing travel times rail frequencies generally tend to decrease, in some cases linearly and in others exponentially. The overwhelming superiority of rail on these shorter routes is possibly due to the inability of air to offer competitive travel times on such connections (i.e. travel time < 200 min), due to the considerable access/egress and waiting time that are required to use the mode. On the medium travel time range, the frequencies of rail and air appear to be similar. Thus, the market of trips with perceived door-to-door travel time between 250 and 400 minutes (i.e. 4 and 6,5 hours circa respectively) appears to have comparable sizes for both rail and air. In this regard, it is worth noting that comparable frequencies within the European market generally suggest that rail might feature higher service capacity in terms of the offered seat compared to air, due to the higher capacity on average of trains compared to narrow-body aircraft. The distribution of travel times in the air network further suggests that in order to be really attractive long-distance rail transport within Europe should aim to offer perceived door-to-door travel times under 450 minutes (7,5 hours). Furthermore, it is worth noting that air does not provide any service with a travel time lower than 200 minutes, being hardly able to offer any service with a travel time lower than 250 minutes. Thus, using this as a maximum threshold and considering the rail access/egress times and waiting time of 70 minutes and in-vehicle time sensitivity of 0,75 (compared to 1 for air) the maximum distance range where rail is theoretically able to compete with air is estimated as:

- 1200km with an average speed of 300km/h
- 1000km with an average speed of 250km/h
- 800km with an average speed of 200km/h

The link connectivity in the air network ranges from a minimum of 0,00037 between Ostrava and Southampton to a maximum of 2,76 between Amsterdam and London, whilst in the rail network it varies from a minimum of 0,00014 on the Dresden-Athens corridor to a maximum of 24,13 on the Wien-Linz corridor. In particular, the histogram of link connectivity, provided by figure 5e, highlights that rail generally appears to have an edge over air in terms of connectivity. This is due to the formulation of link connectivity which places particular importance on high frequencies and short travel times, both extremes where air is not particularly present. Figures 5b and 5d, in fact, show that the magnitude of the highest rail frequencies figures is more than three times over the air counterparts. At the same time, travel times ranging between 100 and 200 are only available for rail. Thus, the important advantage of rail in terms of connectivity can be explained by highlighting the large number of high-frequency low-travel time corridors, which thanks to these combinations of characteristics greatly outperform all other links. In particular, it should be highlighted the prominent importance of extremely high frequencies in establishing connectivity, as the Wien-Linz corridor, featuring the highest figure for frequency, is also the most connected link across both networks. In relation to spatial distance Figure 5f shows that link connectivity (on a logarithmic scale) is generally evenly distributed for values between 0 and 2,5 with the exclusion of an important cluster of outliers composed of short distance (i.e. <500km) rail connections. Whilst link connectivity within the air network appears to be independent of distance, connectivity in the rail network plummets for connections over 500 km. This is probably due to the structural characteristics of rail networks and its infrastructural dimension. Direct services between OD pairs located further apart generally stop in urban areas on the path providing additional frequencies to the surrounding cities, meaning that it is reasonably foreseeable that the number of frequencies for rail services tends to grow exponentially for nodes with strategic infrastructural roles between major cities and in areas with higher densities of urban areas.

4.2 OD Relation Analysis

Table 2 highlights that the air network features a considerably higher number of direct connections, and is consequently more densely connected, compared to the rail network. However, when considering only OD pairs with spatial distances within a threshold of 500km this difference almost disappears, with the two networks featuring a similar number of routes. This implies that for geodesic distances up to 500 km the air-rail route supply shares are rather balanced. The 100-500 km market, in fact, appears to be, by far, the most important market for rail service supply, ac-

Table 2: *Influence of distance on the total number of direct connections*

Distance Threshold (km)	Possible OD pairs		Supplied Direct Connections Air			Supplied Direct Connections Rail		
	Link number	Cumulative	Link number	Cumulative	% possible OD pairs	Link number	Cumulative	% possible OD pairs
500	2.636	2.636	1.328	1.328	50,38	1.090	1.090	41,35
1.000	5.256	7.892	2.854	4.182	54,3	360	1.450	0,68
1.500	4.128	12.020	2.246	6.428	54,4	4	1.454	0,01
∞	3.364	15.384	1560	7.988	46,37	0	1.454	0,00

Table 3: *Influence of distance on the number of direct connections under competition and monopoly*

Distance Threshold (km)	Unique Supplied OD pairs		Inter-modal Competition Connections		Rail Monopoly Connections		Air Monopoly Connections	
	Link number	% possible OD pairs	Link number	% unique connections	Link number	% unique connections	Link number	% unique connections
500	1.774	67,30	644	36,30	446	25,14	684	38,56
1.000	2.908	55,33	306	10,52	54	1,86	2.548	87,62
1.500	2.246	54,41	4	0,18	0	0	2.242	99,82
∞	1.560	46,37	0	0	0	0	1560	100

counting for almost 75% of all the direct connections provided. In particular, the remaining share of direct connections is almost exclusively made up of links between 500 and 1.000 km long, with the exclusion of 4 links, which exceed in route length the upper threshold. In contrast, the availability of direct air connections tends to be the highest between 500 and 1.500 km, serving more than 50% of all the possible OD pair connections. Thus, rail despite supplying a competitive number of direct routes under 500km, appears to be overwhelmingly behind on longer ranges. This is particularly interesting in light of the widely supported argument that rail can offer competitive services on routes up to 800 - 1.500 km (Avogadro et al., 2021; Givoni et al., 2012; Li et al., 2020; Seidenglanz et al., 2021; Serafimova et al., 2022; Witlox et al., 2022; Zhu et al., 2018). In fact, despite most of the possible routes (34,16%) in the network having a distance between 500 and 1000 km, rail currently supplies an extremely limited number of services in this segment of the market. This, together with the fact that air serves more than half of those connections, indicates the importance of this market and suggests that there is an unexplored potential for rail to capture this demand. Some plausible causes of this are the fragmentation of the European rail market and the infrastructural and capacity constraints. The former, in fact, hinder the supply of cross-border international trains, whilst the latter limits the supply of long-distance services in denser areas (i.e. the Netherlands), where large shares of capacity are consumed by local and regional services. Furthermore, the structure of the rail service supply, with a considerable focus on shorter routes, suggests that rail might rely on transfer services rather than on direct connections to cover longer-distance markets. This has a twofold set of implications. On the one hand, the considerable number of transfers might negatively impact the attractiveness of the mode, as transfer time is often perceived in a particularly negative way by passengers. Consequently, imposing transfer services on passengers might greatly reduce the appeal of the mode, leading to them choosing direct air alternatives. On the other hand, relying on transfers imposes strict and ambitious requirements in terms of schedule coordination and disruption management, to allow the feasibility of transfers, ensure their attractiveness in terms of waiting time and avoid the propagation of delays along the network. Failing in doing that could further damage the attractiveness of the mode, leading travellers to prefer comparable air services. These matters are certainly problematic also in the aviation field, but the specific structure of the European rail supply network appears to create the conditions to make the issue even more crucial.

The first distinction is made between direct links under a regime of monopoly and inter-modal competition. The figures provided in Table 3 highlight the total number of unique supplied OD pairs (i.e. OD pairs which have at least a direct connection either in the rail or air network), of links with inter-modal competition (i.e. both air and rail provide direct service on the route), and of links with a monopoly of either rail or air. It is possible to notice how the number of unique supplied OD pairs, despite decreasing with the growth of OD distances, tends to be rather balanced, especially between 500 and 1500 km. This, however, is allowed by the supply of air services that account for the quasi-totality of unique supplied OD pairs on routes longer than 500 km. In particular, under 500 km the shares of links with inter-modal competition, rail monopoly and air monopoly appears to be quite evenly distributed, with a slight majority of links being supplied by only air services. Over the 500 km, the market appears to be almost exclusively served by air services, with rail contributing only in a minimum part. Figure 6 also points out that more than 50% of all the possible OD pairs (i.e. with a distance greater than 100km) are connected either by air or rail. In particular, the large majority (82,87%) is connected

Table 4: *Relationship between distance and distribution of OD pairs by category*

Distance Threshold (km)	Competitive			Substitution-ready			Substitution Potential	Substitution Attractive
	Fully	Part. (tt-wise)	Part. (f-wise)	Fully	Part. (tt-wise)	Part. (f-wise)		
500	27 (58,70%)	147 (71,36%)	50 (51,02%)	13 (68,42%)	15 (93,75%)	8 (61,54%)	11 (35,48%)	34 (23,13%)
1000	19 (41,30%)	59 (28,64%)	48 (48,98%)	6 (31,58%)	1 (6,25%)	5 (38,46%)	19 (61,29%)	69 (46,94%)
1500	0 (0,00%)	0 (0,00%)	0 (0,00%)	0 (0,00%)	0 (0,00%)	0 (0,00%)	1 (3,23%)	44 (29,93%)
Total	46 (100,00%)	206 (100,00%)	98 (100,00%)	19 (100,00%)	16 (100,00%)	13 (100,00%)	31 (100,00%)	147 (100,00%)

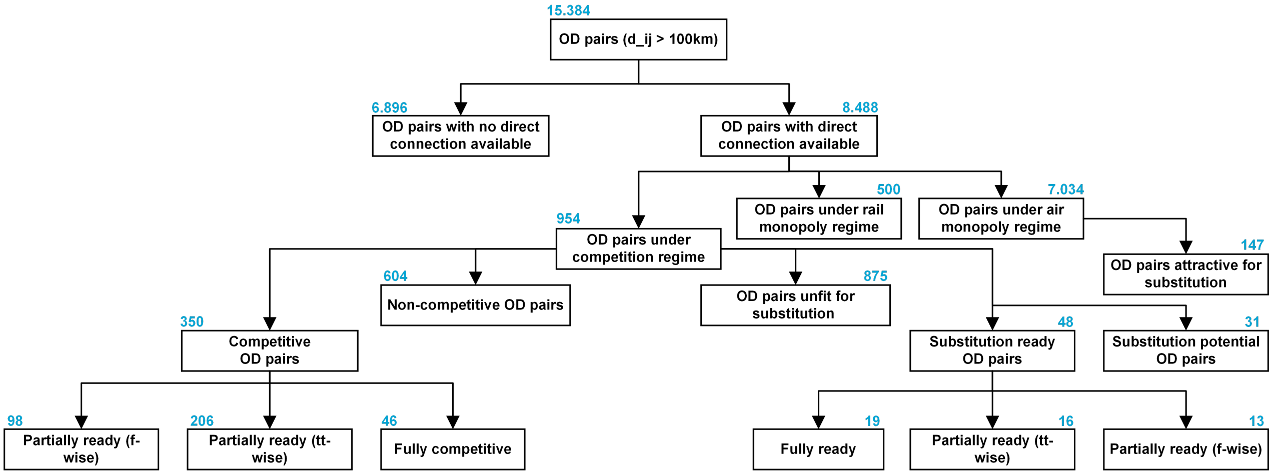
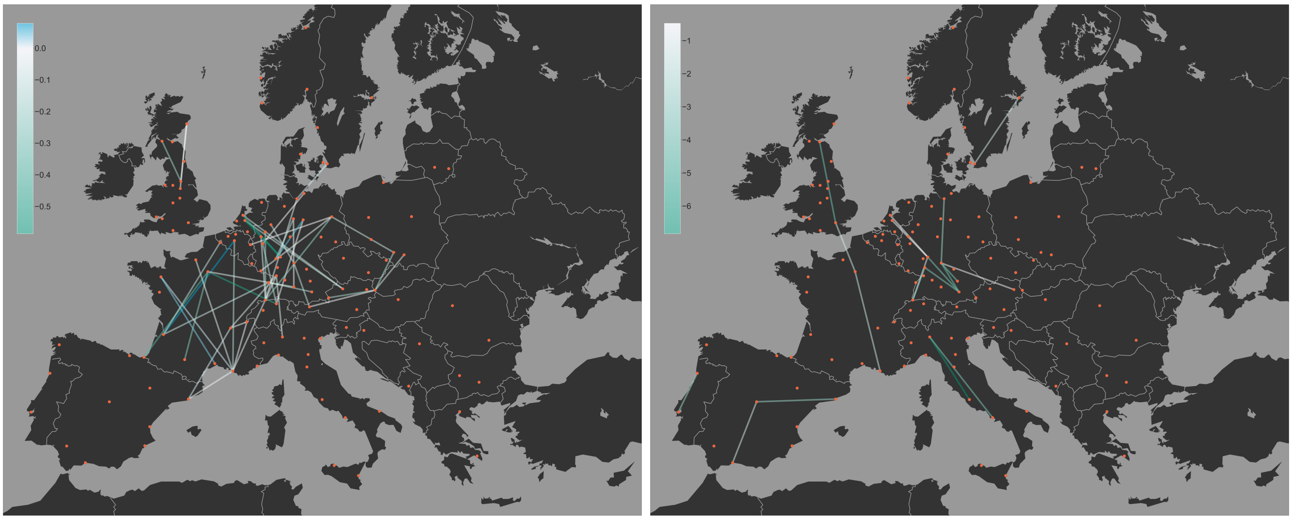
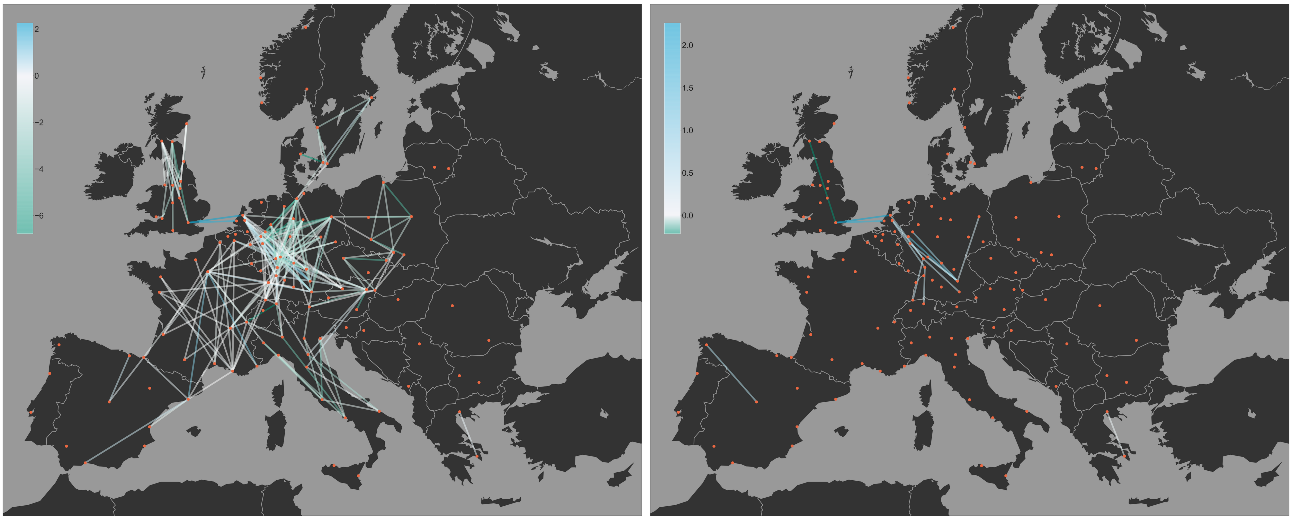


Figure 6: *Distribution of the OD pairs across the categories identified within this study*

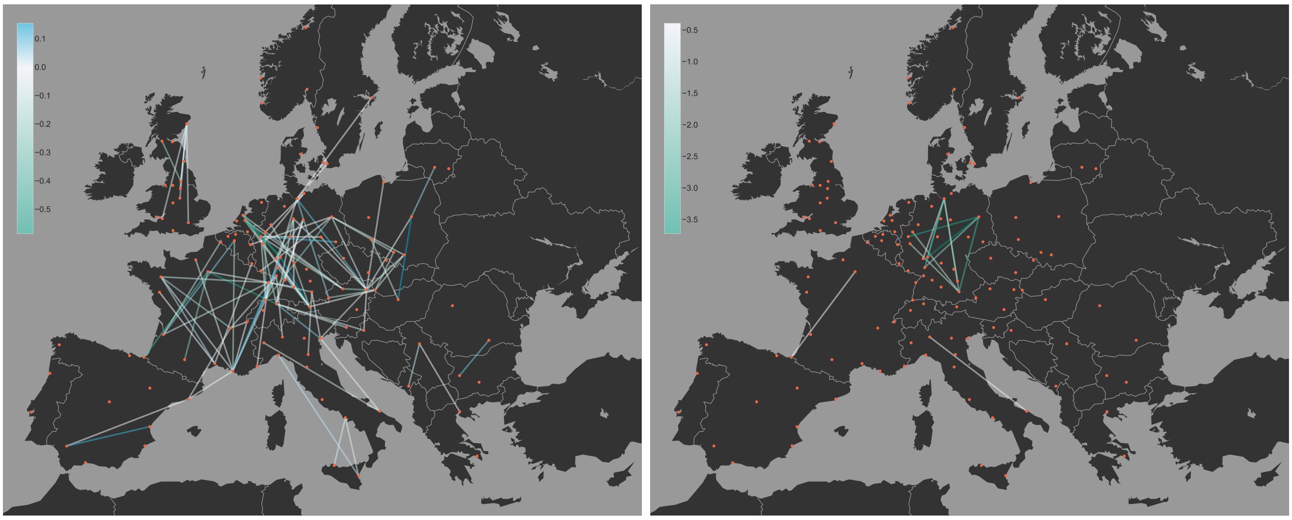
only by air, with rail monopolies accounting for a mere 5,89% and the remaining 11,24% of OD pairs being served by both modes. Furthermore, more than half (i.e. 63,31%) of the OD pairs under competition rail and air are non-competitive, so one has an edge over the other in terms of both travel time and frequency. This finding is particularly interesting, as it suggests that other important factors might influence passengers' modal choices. Finally, the distribution of direct links under a competition regime between competitive and substitution-ready connections and their relationship with distance are described in Table 4. It is worth noting the limited but non-negligible number of routes where currently rail is theoretically able to substitute air services, offering better services. Furthermore, some others could become part of this selection by either reducing the travel time thresholds or by increasing the frequency thresholds employed for the selection of the links. However, the number of competitive links is considerably higher, indicating that rail can effectively compete with air on a wide set of routes, even on links without sufficient demand to justify substitution. The amount of competitive links appears to be more evenly distributed across distance thresholds compared to substitution-ready ones, which are mostly concentrated on 100-500 km OD pairs. Interestingly enough most substitution potential and attractive routes are in the 500-1.000 threshold, suggesting that a considerable amount of air demand lies within ranges where rail can be competitive with air. Furthermore, the limited number of substitution-ready OD pairs suggests that in many cases demand might not be not large enough to allow for the considerable investments in rail infrastructure and services required to guarantee the modal competitiveness of rail with air. This calls for more specific demand estimation models to forecast the evolution of the market and the economic feasibility of future investments. To provide some additional insights into the quality of the links the difference in connectivity (air-rail) is projected on the spatial distribution maps of the corridors using a scale of colour ranging from green, to white and blue. Darker shades of green indicate negative difference values, representing the links where rail most comprehensively dominates air in terms of connectivity. Lighter colours represent links with connectivity differences around nil, where the degree of connectivity is similar across both modes. Finally darker shades of blue represent increasingly air-dominated links in terms of connectivity. These twofold visualisations aim to provide an overall overview of the quantity, quality and spatial locations of links, allowing to understand how the connectivity on each link compares across modes and providing interesting insights for policy-making. In terms of connectivity, an interesting finding relates to the contrasting patterns found in France and Germany highlighted by Figures 7a, 7c and 7e. Whilst in the former country, routes with competing travel times generally feature higher air frequencies, in the latter rail frequencies appear to considerably exceed the ones of air. At the same time, on routes with similar frequencies rail appears to be more competitive in terms of travel times in France whereas in Germany air appears to have an edge. These differences can be explained by the type of rail services that the two country offer. Whilst in Germany, rail is focused on offering wide-spread coverage and high frequencies, in France rail aims to offer competitive travel times using high-speed rail. Overall, it could be concluded that on links with similar travel times generally, rail provides more frequencies, whilst on links with similar frequency air features lower travel times. Finally, in terms of the spatial distribution of lines, it is interesting to notice the Amsterdam-Frankfurt and Paris-Zürich corridors, where frequencies are similar across the two modes, despite rail travel times being shorter. The presence of high air frequencies despite the less competitive travel times is probably related to the provision of feeder services towards the main European air hubs to serve intercontinental destina-



(a) *Spatial Distribution of the Fully Competitive Links* **(b)** *Spatial Distribution of the Fully Substitution-ready Links*



(c) *Spatial Distribution of the Partially Competitive Links (travel time wise)* **(d)** *Spatial Distribution of the Partially Substitution-ready Links (travel time wise)*

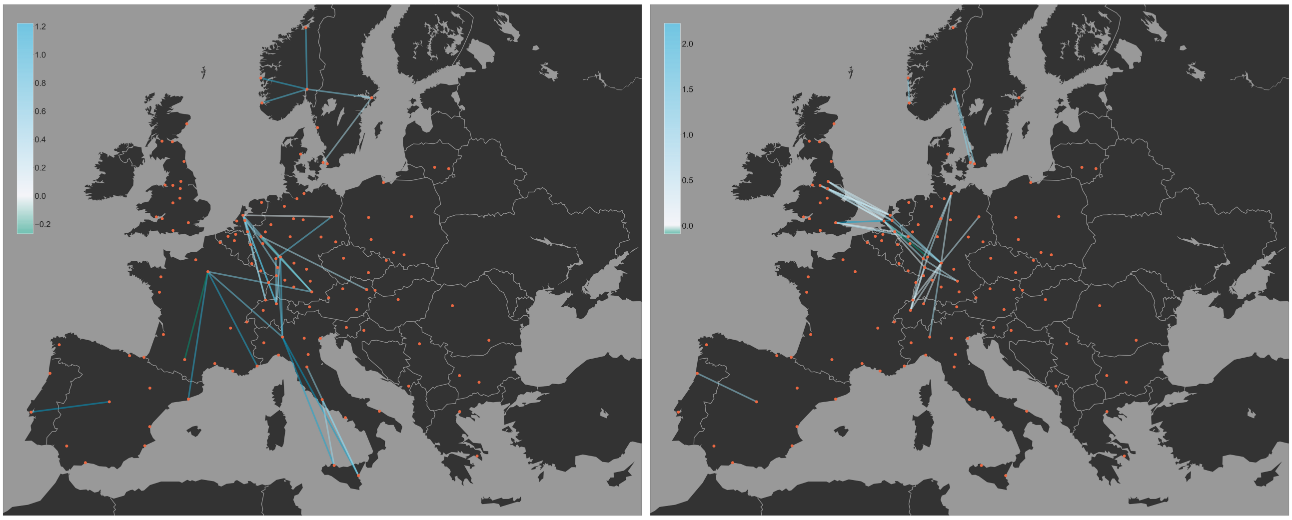


(e) *Spatial Distribution of the Partially Competitive Links (frequency wise)* **(f)** *Spatial Distribution of the Partially Substitution-ready Links (frequency wise)*

Figure 7: *Spatial Distribution of Competitive and Substitution-ready OD Pairs*

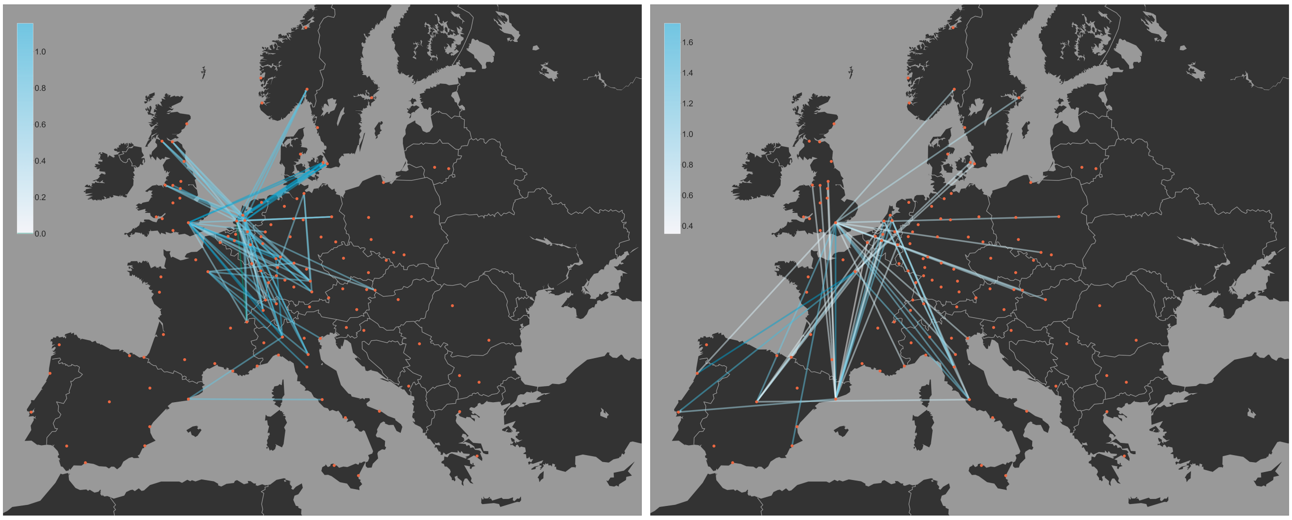
tions. Air frequencies on these routes could be thus decreased fostering air-rail inter-modality.

In terms of substitutability, Figure 6 shows that 79 out of the 954 OD pairs under a competition regime



(a) *Spatial Distribution of the Routes with Substitution Potential*

(b) *Spatial Distribution of the 100-500 km Links Attractive for Substitution*



(c) *Spatial Distribution of the 500-1.000 km Links Attractive for Substitution*

(d) *Spatial Distribution of the 1.000-1.500 km Links Attractive for Substitution*

Figure 8: *Spatial Distribution of Substitution Potential and Attractive OD Pairs*

feature, particularly high air frequencies. On 48 of these rail offers better travel time or frequencies (or both), whilst on 31 air outperforms rail. This implies that air manages to serve also connections where rail features better performances and that there is a number of connections with a considerable demand that rail could capture. In particular, the results point out that under the current market conditions and considering the supply performances of the two modes rail might already substitute air over 19 OD pairs in the network. Figure 7b highlights that most of them are national routes between the major cities within the country, such as the case of the Milan - Rome and the Milan - Naples in Italy, the Madrid - Barcelona and Madrid - Málaga in Spain, the Lisbon - Porto in Portugal, the Marseille - Paris in France, the London - Edinburgh in the UK, the Munich - Frankfurt in Germany and the Stockholm - Malmö in Sweden. In terms of cross-border connections London - Paris, Amsterdam - Frankfurt and Basel - Frankfurt appear to also feature the conditions for complete substitution of air services with rail. Furthermore, increasing the frequencies on 16 connections and reducing the travel times on 13 others would allow to establish the aforementioned substitution conditions. Some notable examples of the former category are the London - Glasgow, the London - Amsterdam, the Madrid - Santiago de Compostela, whilst the Paris - San Sebastián, the Milan - Bari and some important connections between major German cities (e.g. Munich - Berlin, Berlin - Frankfurt, Frankfurt - Hamburg and Hamburg - Munich) fall in the latter, as shown by Figures 7d and 7f respectively. In terms of trends, it is worth noting that Germany appears to be the country which currently features the highest share of substitution-ready routes. Considering the current rail supply scenario in the country it could be argued that the eventual implementation of high-speed rail services, providing shorter travel times, might lead to increased attractiveness of rail which could in turn contribute to the sub-

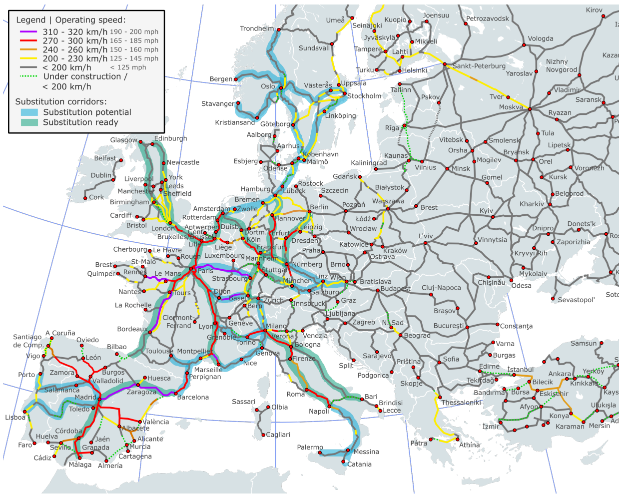


Figure 9: *Substitution Potential and Substitution-ready Routes at the Infrastructure Level (Source: adaptation from Wikimedia Commons, 2022)*

stitution of a consistent portion of the national flights currently available in the market. However, more specific studies are required to assess the actual costs and benefits of such a scenario and its implications for the entire European long-distance transport market. Finally, on 31 connections both travel time and frequency enhancements are required for rail to become an attractive substitute for air services. Many crucial OD pairs fall into this category, interestingly enough most of them are cross-border international connections. Notable examples highlighted in Figure 8a include the Paris - Milan, Paris - Munich, Paris - Barcelona, Amsterdam - Berlin, Stockholm - Copenhagen, Oslo - Stockholm, Zürich - Amsterdam, Wien - Düsseldorf, Madrid - Lisbon and Milan - Frankfurt. Moreover, 147 OD pairs with distances below 1.500 km that currently are under an air monopoly regime feature high levels of air demand that might justify investments in rail infrastructure and services to capture some of those market shares. Figure 8b highlights that these routes, which are deemed attractive for substitution, between 100 and 500 km tend to concentrate between Italy, Switzerland, Germany, the Netherlands and the UK on a north-south axis, with a few exceptions. On the other hand, on the 500 to 1.000 km routes, the spatial distribution expands from the aforementioned countries to some of their neighbours, as illustrated by Figure 8c. Finally, Figure 8d suggest that more peripheral areas tend to be increasingly included when considering longer routes until 1.500 km. It is worth noting that most of these routes depart from the major European air hub urban areas, such as London, Paris, Amsterdam, Barcelona and Rome. In terms of infrastructure, Figure 9 evidence an important lack of rail infrastructure between Spain and Portugal, which is most probably hindering the capacity to offer competing services on the Lisbon-Madrid corridor. Moreover, between Barcelona and Paris, there appears to be a gap in the high-speed infrastructure, which also relates to the connectivity of the Toulouse-Paris corridor. Another considerable infrastructure gap is the one between Scandinavia/Northern Germany and the Netherlands/Belgium, as no rail infrastructure currently connects Bremen and Groningen. Other critical gaps are the two alp transit corridors, the Frejus, connecting Turin to Lyon and the Gotthard, connecting Milan to Zürich, the connection between Lyon and Nice which requires a detour to go around the alps and the standard rail lines with lower speeds between Germany and Austria and between Sicily and Naples. It is worth noticing that two high-speed corridors, the Paris-Basel and Paris-Strasbourg lines are currently not employed by substitution-ready links. However, due to the state of the infrastructure, it is possible to argue that the reasons for this are most probably related to gaps in the service supply rather than to infrastructural deficiencies. In that regard, offering high-speed cross-border services from Paris to a wider set of destinations within Switzerland and Germany would probably allow reducing the total number of flights currently connecting the capital of France to central Europe.

4.3 Node Connectivity

Figure 10 highlight that node connectivity is generally higher for rail compared to air. This is in line with the results at the link level and with the expectations. The indicator is based on a ratio that ascribes more importance to high frequencies (i.e. numerator of the ratio) compared to low travel times (i.e. denominator of the ratio). Furthermore, it is worth noting that the gap in connectivity widens when considering indirect connections due to the transfer times formulation being directly related to frequencies, which are considerably higher within the rail network. This could be adjusted by taking into account schedule coordination and including actual transfer times, rather than using shortest path and transfer time estimations. Finally, the lower tail of the distribution shows that rail features

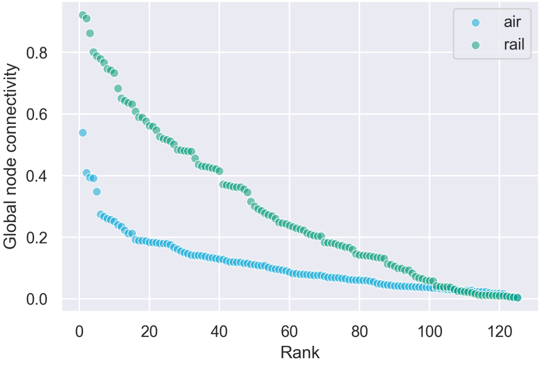
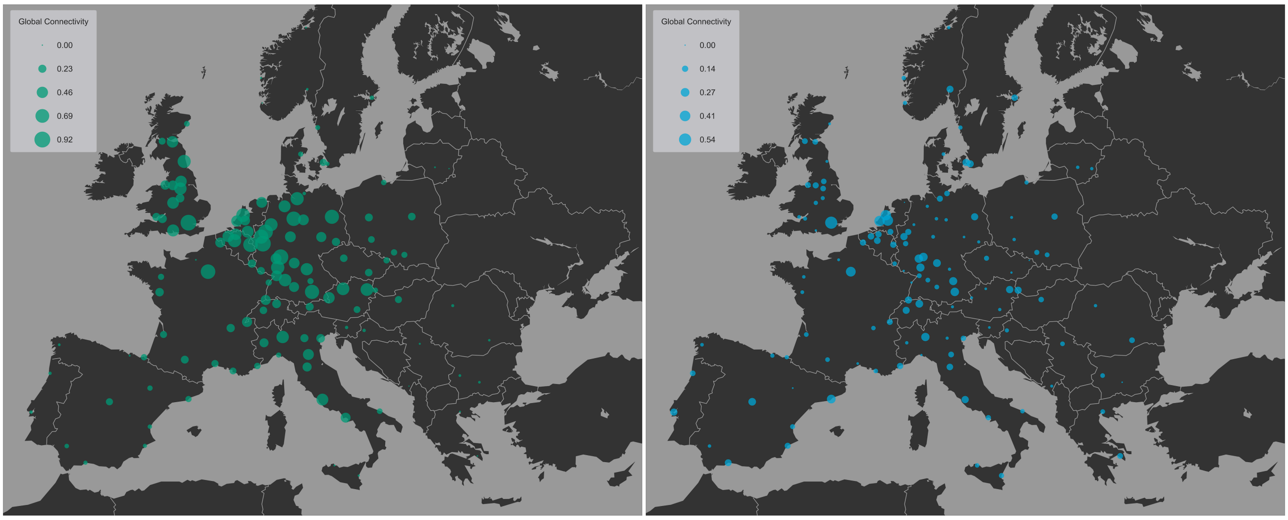


Figure 10: *Node Connectivity Distribution*



(a) *Rail*

(b) *Air*

Figure 11: *Spatial Distribution of the Node Connectivity*

poorer connectivity for the least connected nodes in the network. Thus, the rail network is characterised by higher heterogeneity in connectivity, compared to the air network, where connectivity is more evenly distributed across the nodes. In terms of critical components, Figure 11b highlights that the most important nodes in the air network are London, Amsterdam, Frankfurt and Paris, despite the generally good performances that characterise the entirety of the network with a few exceptions only. On the other hand, Figure 11a points out the wider variations in terms of performances within the rail network, with node directness and connectivity being considerably higher in countries of Central Europe, such as Germany, Austria and the Netherlands. An interesting finding is that node connectivity within the rail network appears to be correlated with the urban area density, as areas with a higher concentration of urban areas show higher values for node connectivity. In contrast, airport size might replace urban area density as the main influence factor in determining the degree of node connectivity in the air network. It is also worth noting the generalised poor connectivity of eastern Europe, despite an important distinction that must be made between Poland, Czech Republic and Hungary and the remaining eastern European countries, with the former group of countries featuring considerably higher degrees of node connectivity in the rail network. The main reason for this probably relates to the limited demand for long-distance transport and the absence of important airport hubs, suggesting that geographical centrality might play a role also for air connectivity. The node connectivity spatial distribution, however, is particularly useful to identify the specific transport network characteristics of each national system. In regards to rail, most countries (i.e. Germany, the Netherlands, Belgium, Austria, Poland, Czech Republic and the UK) appear to have a decentralised and polycentric structure, with low heterogeneity in connectivity degrees across most nodes. It is worth noting that all these countries feature generally high urban area densities and have a considerable focus on conventional train services. In contrast, France and Spain are characterised by a centralised and monocentric structure, with a radial infrastructure topology. Both countries also have a focus on high-speed and have lower densities of urban areas. An interesting case is Italy, which, despite sharing the decentralised characteristics of the former group of countries, focused mostly on high-speed rail. This can be probably explained by the linear shape of the country, which allows to equally serve all urban areas located on the north-south axis. In terms of the air network national air network systems generally appear to be very centralised across all of Europe with each country featuring a few hubs. The most polycentric countries are Germany and Italy, whilst France, Spain, and Austria are very centralised systems.

5 Conclusion

Overall, the results from this study highlight that rail connections feature higher connectivity figures compared to air counterparts. This might suggest that rail connections currently have an edge over air counterparts. However, it is worth noting that this gap is further exacerbated by the market segmentation of the two modes. Whilst rail largely serves mostly short connections, with low travel times and high frequencies, air focuses more on longer distances, with the naturally lower frequencies and higher travel times that characterise this section of the market. This first finding highlights some of the complexities related to comparing the system-wide performances of modes with different core characteristics, such as rail and air.

A second finding relates to the spatial distribution of the most significant links and nodes. Whilst within the air network the importance of urban areas generally mostly relies on airport infrastructure accessibility, in the rail network the nodes' significance appears to heavily depend on their geographical centrality. Thus the infrastructure layer, despite representing an important hindering factor for rail, has also some positive impacts on connectivity that could be employed to make the mode more attractive to some specific market segments. By increasing its capillarity rail is able to provide good accessibility also to smaller centres, which are not directly served by a major airport. In this regard, it is worth noting the importance of the trade-off between the higher number of stops (i.e. regional coverage) and faster/more direct services (i.e. long-distance) when scheduling rail services. Furthermore, the distribution of substitutable routes suggests that substitution could probably more easily take place in central Europe as opposed to more peripheral areas. The results also point out the considerable differences across European regions. Central Europe might have problems related to congestion and capacity due to the considerable through traffic, whilst peripheral areas such as eastern Europe and the Iberian peninsula might face an opposite threat, with sub-optimal infrastructure utilisation rates. The considerable regional differences highlight the need for context-specific policies and solutions, suggesting that it would be hardly possible to use a one-rule-fits-all approach. It is possible to conclude that the creation of a single rail market with more homogeneous characteristics is fundamental in order to increase the modal competitiveness of rail on longer distances up to 1.200 km. In particular, understanding and considering the different characteristics of each system appears to be a crucial premise to guarantee the success of the harmonisation process that is taking place within the European rail market.

This study confirms that rail has the potential to compete with air up to ranges of 1.000-1.200 km, despite suggesting that this potential has not been explored yet. In fact, whilst the OD pair connectivity within the air network appears to be rather stable along all the different market sections, rail connectivity plunges on distances longer than 500 km. The findings lead to conclude that on these routes air connectivity is superior compared to rail due to a considerable gap in supply related to the poor performances in terms of frequencies, travel times and the number of direct connections. Contextualising the results within the broader literature suggest that important causes for that include:

- The conservative market dynamics, which do not stimulate railway undertakings to launch new international routes and offer more direct services.
- The fragmentation of the European rail infrastructure and service supply, which creates further barriers to the capacity to offer seamless cross-border connections. The findings show that the large majority of rail monopolies are made up of shorter national routes, whilst international OD pairs appear to be less connected featuring poorer performances.
- The focus of rail on regional coverage rather than long-distance services, both from a service and infrastructure perspectives. Many national operators and stakeholders still appear to favour the optimisation of operations at the national rather than European level, with a considerable focus on conventional rather than long-distance service types (high-speed and night trains). In terms of infrastructure the European high-speed network, despite being a crucial condition for rail to compete with air, is still fragmented with a considerable number of gaps.

Finally, the fact that air supplies more than half of those connections, indicates the considerable importance of this market suggesting that more endeavours should be directed at improving rail competitiveness on routes over 500 km.

Despite the good performances of rail on shorter routes between 100 and 500 km, air still offers many super short-haul flights on this range. Dobruszkes et al. (2022) highlight three main reasons for this:

- Hostile physical geography. In particular, this study has highlighted some corridors where considerable detours are required due to geographical conformation of the territory (e.g. Barcelona-Rome due to the sea, and Nice-Geneva due to the alps).
- Commercial reasons. These include feeder services which airlines require to feed their hubs, services targeted at wider regions, the suburbia and areas which are not necessarily centrally located within cities, and mutualising flights ad triangular routes. This study, in particular, highlights the crucial importance of feeder routes, especially between the main European air hubs, where services feature similar frequencies across the two modes despite the shorter rail travel times (e.g. Amsterdam - Frankfurt and Paris - Zürich).
- Political reasons. These include PSOs and subsidised services required to provide access to remote areas. This is the least interesting for this research given that only major urban areas are included within the scope.

Following the results of this study, two more plausible sets of reasons are added to the list:

- Unavailable infrastructure. These include all those OD pairs that cannot be efficiently connected by rail due to the limitations and constraints in terms of available infrastructure (e.g. the Copenhagen-Hamburg—Bremen-Amsterdam corridor due to the infrastructural gap between

Groningen and Bremen, Spain - Portugal connections and border cities in radial systems such as Spain).

- Unavailable capacity. These include all those OD pairs where rail service provision is limited by the high density of rail services and infrastructure with no further capacity available.

It is worth noting that despite routes included in the first three categories being most likely to remain air monopolies, for the latter two categories the tide can be turned by improving/upgrading existing infrastructure or building new links. A notable exception is represented by feeder services. Air frequencies on these routes could be captured and substituted by fostering air-rail inter-modality and offering alternative rail services between airport terminals with inter-modal fare integration. To capture feeder services and to effectively substitute them it is envisaged that rail services should provide high service frequencies and direct connections between airport terminals.

Finally, the main bottlenecks hindering the capacity of rail to compete and substitute air from a network supply perspective have been identified and summarised in the following points:

- Lack of direct connections. Too many transfers are currently required to reach most destinations, making the mode less attractive for passengers.
- Lack of long-distance services. This is especially the case in the 500 - 1.000 km market segment, where air serves the overwhelming majority of connections despite theoretical studies arguing that rail should be able to compete.
- Lack of specific long-distance service types such as high-speed connections (i.e. low travel time) and/or night trains (i.e. low travel time perception).
- Lack of cross-border, international services, with a few exceptions only.
- Lack of homogeneity in the characteristics of the national systems. The ideas of a single rail area have not managed to translate into reality yet, as the current rail supply appears fragmented across different countries.
- Dependence on schedule coordination, risk of domino effect and delay propagation over all the network. This has rather severe consequences in terms of the impact of disruptions on mode attractiveness.

The results further confirm that this represents a preliminary study that, despite providing some interesting insights, should be further developed to provide precise and accurate advice for practice and policy-making. In particular, given that the demand perspective is outside of the scope of this paper, it is believed that additional research from that side would be beneficial to complement and reinforce the findings of this study. It is worth noting that the proposed methodology could also be employed to compare different scenarios and assess the evolution of the indicators over the years. Based on these temporal patterns and the model specifications, a basic simulation tool for European long-distance transport could also be constructed to provide forecasts and to assess how to practically make rail more attractive. This would prove particularly useful in evaluating the impact of investments and policies on the modal split, allowing policy-makers and industry professionals to fine-tune them. Simpler approaches, such as the study from Kroes and Savelberg (2019), also represent a valid alternative that might be more straightforward developed from the results of this study. Furthermore, future research in this direction should also aim to include road modes (i.e. private cars, carpooling and long-distance buses) that are not considered within this study. Finally, further research could analyse the disaggregated performances of the different rail service types (i.e. conventional rail, high-speed services and night trains) to explore how each contributes to the overall rail connectivity, providing useful insights to aid decision-making processes directed at shaping future rail supply scenarios.

References

Adler, N., Pels, E., & Nash, C. (2010). High-speed rail and air transport competition: Game engineering as tool for cost-benefit analysis [Modelling Non-urban Transport Investment and Pricing]. *Transportation Research Part B: Methodological*, *44*(7), 812–833. https://doi.org/https://doi.org/10.1016/j.trb.2010.01.001

Albalate, D., Bel, G., & Fageda, X. (2015). Competition and cooperation between high-speed rail and air transportation services in europe. *Journal of Transport Geography*, *42*, 166–174. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2014.07.003

Augustyniak, W., & Olipra, Ł. (2014). The potential catchment area of Polish regional airports. *Journal of International Studies*, *7*. https://doi.org/10.14254/2071-8330.2014/7-3/13

Avogadro, N., Cattaneo, M., Paleari, S., & Redondi, R. (2021). Replacing short-medium haul intra-European flights with high-speed rail: Impact on CO2 emissions and regional accessibility. *Transport Policy*, *114*. https://doi.org/10.1016/j.tranpol.2021.08.014

Behrens, C., & Pels, E. (2012). Intermodal competition in the London–Paris passenger market: High-Speed Rail and air transport. *Journal of Urban Economics*, *71*(3), 278–288. https://doi.org/https://doi.org/10.1016/j.jue.2011.12.005

Bergantino, A. S., Capozza, C., & Capurso, M. (2015). The impact of open access on intra- and inter-modal rail competition. A national level analysis in Italy. *Transport Policy*, *39*, 77–86. https://doi.org/https://doi.org/10.1016/j.tranpol.2015.01.008

Bergantino, A. S., & Madio, L. (2020). Intermodal competition and substitution. HSR versus air transport: Understanding the socio-economic determinants of modal choice [Air Transport Markets, Strategies and Policies]. *Research in Transportation Economics*, *79*, 100823. https://doi.org/https://doi.org/10.1016/j.retrec.2020.100823

Beria, P., Tolentino, S., Bertolin, A., & Filippini, G. (2019). Long-distance rail prices in a competitive market. Evidence from head-on competition in Italy. *Journal of Rail Transport Planning & Management*, *12*, 100144. https://doi.org/https://doi.org/10.1016/j.jrtpm.2019.100144

Burghouwt, G., & Redondi, R. (2013). Connectivity in air transport networks an assessment of models and applications. *Journal of Transport Economics and Policy*, *47*.

Clewlow, R. R. L., Sussman, J. M., & Balakrishnan, H. (2014). The impact of high-speed rail and low-cost carriers on European air passenger traffic. *Transport Policy*, *33*. https://doi.org/10.1016/j.tranpol.2014.01.015

Di Pietrantonio, L., & Pelkmans, J. (2004). The Economics of EU Railway Reform. *Competition and Regulation in Network Industries*, *5*. https://doi.org/10.1177/178359170400500304

Dijkstra, L., Poelman, H., & Veneri, P. (2019). The EU-OECD definition of a functional urban area. https://doi.org/https://doi.org/https://doi.org/10.1787/d58cb34d-en

Dobruszkes, F., Lennert, M., & Van Hamme, G. (2011). An analysis of the determinants of air traffic volume for European metropolitan areas. *Journal of Transport Geography*, *19*(4), 755–762. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2010.09.003

Dobruszkes, F., Mattioli, G., & Mathieu, L. (2022). Banning super short-haul flights: Environmental evidence or political turbulence? *Journal of Transport Geography*, *104*, 103457. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2022.103457

Eurocontrol. (2020). *Data Snapshot 9 on 2020's smaller aircraft, shorter flights and fluctuating weight*. Retrieved September 2, 2022, from https://www.eurocontrol.int/publication/eurocontrol-data-snapshot-2020s-smaller-aircraft-shorter-flights-and-fluctuating-weight

European Commission. (2013). Case No COMP/M.6663 – Ryanair/AER Lingus III. *Regulation (EC) No 139/2004 Merger Procedure*, *3*(8). https://ec.europa.eu/competition/mergers/cases/decisions/m6663_20130227_20610_3904642_EN.pdf

European Commission. (2020). *Press Release: The journey begins – 2021 is the European Year of Rail!* Retrieved April 6, 2022, from https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2528

European Commission. (2022). *Fourth railway package of 2016*. Retrieved April 6, 2022, from https://transport.ec.europa.eu/transport-modes/rail/railway-packages/fourth-railway-package-2016_en

European Environment Agency. (2021). Transport and Environment Report 2020 — Train or Plane? https://doi.org/10.2800/43379

European Parliament & Council of the European Union. (2020). Decision (EU) 2020/2228 of the 28.12.2020. *Official Journal of the European Union*, *L 437*, 108–115.

Eurostat. (2019). *People on the move — statistics on mobility in europe — 2019 edition*. https://doi.org/10.2785/50487

Eurostat. (2021a). *Climate change - driving forces*. Retrieved May 6, 2022, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Climate_change_-_driving_forces

Eurostat. (2021b). *Trends in greenhouse gas emissions*. Retrieved May 6, 2022, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Greenhouse_gas_emission_statistics_-_emission_inventories#Trends_in_greenhouse_gas_emissions

Fraunhofer ISI & CE Delft. (2020). *Methodology for ghg efficiency of transport modes* (tech. rep.). European Environment Agency.

Frei, A., Kuhnimhof, T., & Axhausen, K. W. (2010). Long-Distance Travel in Europe Today: Experiences with a New Survey. *Transportation Research Board*.

Gerike, R., & Schulz, A. (2018). Workshop Synthesis: Surveys on long-distance travel and other rare events. *Transportation Research Procedia*, *32*. https://doi.org/10.1016/j.trpro.2018.10.032

Givoni, M., Dobruszkes, F., & Lugo, I. (2012). Uncovering the real potential for air-rail substitution: An exploratory analysis. *Energy, Transport, & the Environment: Addressing the Sustainable Mobility Paradigm*, *9781447127178*. https://doi.org/10.1007/978-1-4471-2717-8_27

Gudmundsson, S., Cattaneo, M., & Redondi, R. (2021). Forecasting temporal world recovery in air transport markets in the presence of large economic shocks: The case of covid-19. *Journal of Air Transport Management*, *91*, 102007. https://doi.org/https://doi.org/10.1016/j.jairtraman.2020.102007

Heufke Kantelaar, M., Molin, E., Cats, O., Donners, B., & van Wee, B. (2022). Willingness to use night trains for long-distance travel. *Travel Behaviour and Society*, *29*, 339–349. https://doi.org/https://doi.org/10.1016/j.tbs.2022.08.002

Kroes, E., & Savelberg, F. (2019). Substitution from air to high-speed rail: The case of amsterdam airport. *Transportation Research Record*, *2673*(5), 166–174. https://doi.org/10.1177/0361198119839952

Li, X., Tang, J., Hu, X., & Wang, W. (2020). Assessing intercity multimodal choice behavior in a Touristy City: A factor analysis. *Journal of Transport Geography*, *86*, 102776. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2020.102776

Limtanakool, N., Dijst, M., & Schwanen, T. (2006). The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium- and longer-distance trips. *Journal of Transport Geography*, *14*(5), 327–341. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2005.06.004

Luo, D., Cats, O., van Lint, H., & Currie, G. (2019). Integrating network science and public transport accessibility analysis for comparative assessment. *Journal of Transport Geography*, *80*, 102505. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2019.102505

Malichová, E., Cornet, Y., & Hudák, M. (2022). Travellers' use and perception of travel time in long-distance trips in Europe. *Travel Behaviour and Society*, *27*, 95–106. https://doi.org/https://doi.org/10.1016/j.tbs.2021.12.003

Malina, R. (2010). Competition in the German airport market: An empirical investigation.

Martín, J. C., Román, C., García-Palomares, J. C., & Gutiérrez, J. (2014). Spatial analysis of the competitiveness of the high-speed train and air transport: The role of access to terminals in the Madrid–Barcelona corridor. *Transportation Research Part A: Policy and Practice*, *69*, 392–408. https://doi.org/https://doi.org/10.1016/j.tra.2014.09.010

Moeckel, R., Fussell, R., & Donnelly, R. (2015). Mode choice modeling for long-distance travel. *Transportation Letters*, *7*. https://doi.org/10.1179/1942787514Y.0000000031

Moreno-Monroy, A. I., Schiavina, M., & Veneri, P. (2021). Metropolitan areas in the world. delineation and population trends [Delineation of Urban Areas]. *Journal of Urban Economics*, *125*, 103242. https://doi.org/https://doi.org/10.1016/j.jue.2020.103242

Peer, S., Knockaert, J., Koster, P., Tseng, Y.-Y., & Verhoef, E. T. (2013). Door-to-door travel times in RP departure time choice models: An approximation method using GPS data. *Transportation Research Part B: Methodological*, *58*, 134–150. https://doi.org/https://doi.org/10.1016/j.trb.2013.10.006

Petersen, M., Sessa, C., Enei, R., Ulied, A., Larrea, E., & Obisco, O. (2009). *Transvisions: Report on Transport Scenarios with a 20 and 40 Year Horizon* (tech. rep.). European Commission.

Poelman, H. (2013). Measuring accessibility to passenger flights in Europe: Towards harmonised indicators at the regional level. *European Commission Regional Focus*.

Rich, J., & Mabit, S. L. (2012). A Long-Distance Travel Demand Model for Europe. *European Journal of Transport and Infrastructure Research*, *12*(1). https://doi.org/10.18757/ejtir.2012.12.1.2946

Román, C., Espino, R., & Martín, J. C. (2010). Analyzing competition between the high speed train and alternative modes. The case of the Madrid-Zaragoza-Barcelona corridor. *Journal of Choice Modelling*, *3*. https://doi.org/10.1016/S1755-5345(13)70030-7

Román, C., & Martín, J. C. (2014). Integration of hsr and air transport: Understanding passengers' preferences. *Transportation Research Part E: Logistics and Transportation Review*, *71*, 129–141. https://doi.org/https://doi.org/10.1016/j.tre.2014.09.001

Seidenglanz, D., Taczanowski, J., Król, M., Horňák, M., & Nigrin, T. (2021). Quo vadis, international long-distance railway services? Evidence from Central Europe. *Journal of Transport Geography*, *92*, 102998. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2021.102998

Sen, A. K., & Morlok, E. K. (1976). On Estimating Effective Frequencies And Average Waiting Times For Indirect Connections. *Transportation Planning and Technology*, *3*. https://doi.org/10.1080/03081067608717104

Serafimova, T., Finger, M., & Montero-Pascual, J. (2022). Modal shift: the moment of truth. Main takeaways and lessons learnt from the European year of rail. *RSC Policy Briefs*, 97–111. https://doi.org/10.2870/26564

Socorro, M. P., & Viecens, M. F. (2013). The effects of airline and high speed train integration. *Transportation Research Part A: Policy and Practice*, *49*, 160–177. https://doi.org/https://doi.org/10.1016/j.tra.2013.01.014

von Ferber, C., Holovatch, T., Holovatch, Y., & Palchykov, V. (2009). Public transport networks: Empirical analysis and modeling. *European Physical Journal B*, *68*. https://doi.org/10.1140/epjb/e2009-00090-x

Weeks, J. R. (2010). Defining Urban Areas. In T. Rashed & C. Jürgens (Eds.), *Remote Sensing of Urban and Suburban Areas* (pp. 33–45).

Springer Netherlands. https://doi.org/10.1007/978-1-4020-4385-7_3

Wen, C. H., & Koppelman, F. S. (2001). The generalized nested logit model. *Transportation Research Part B: Methodological*, *35*. https://doi.org/10.1016/S0191-2615(00)00045-X

Wikimedia Commons. (2022). *High Speed Railroad Map of Europe*. Retrieved October 13, 2022, from https://commons.wikimedia.org/wiki/File:High_Speed_Railroad_Map_of_Europe.svg

Witlox, F., Zwanikken, T., Jehee, L., Donners, B., & Veeneman, W. (2022). Changing tracks: Identifying and tackling bottlenecks in european rail passenger transport. *European Transport Research Review*, *14*(7). https://doi.org/10.1186/s12544-022-00530-9

World Tourism Organization. (2018). *European Union Tourism Trends*. UNWTO. https://doi.org/10.18111/9789284419470

WTO. (2018). *European Union Tourism Trends*. United Nations World Tourism Organisation. https://doi.org/10.18111/9789284419470

Xu, Z., Zhang, Q., Chen, D., & He, Y. (2020). Characterizing the Connectivity of Railway Networks. *IEEE Transactions on Intelligent Transportation Systems*, *21*. https://doi.org/10.1109/TITS.2019.2909120

Zhang, A., Wan, Y., & Yang, H. (2019). Impacts of high-speed rail on airlines, airports and regional economies: A survey of recent research. *Transport Policy*, *81*, A1–A19. https://doi.org/https://doi.org/10.1016/j.tranpol.2019.06.010

Zhu, Z., Zhang, A., & Zhang, Y. (2018). Connectivity of intercity passenger transportation in China: A multi-modal and network approach. *Journal of Transport Geography*, *71*. https://doi.org/10.1016/j.jtrangeo.2017.05.009

Zhu, Z., Zhang, A., & Zhang, Y. (2019). Measuring multi-modal connections and connectivity radiations of transport infrastructure in China. *Transportmetrica A: Transport Science*, *15*. https://doi.org/10.1080/23249935.2019.1642966