European air/rail intermodality network potential: a classification of hub integration

D. (Daan) Nijhof

Delft University of Technology

European air/rail intermodality network potential: a classification of hub integration

by

D. (Daan) Nijhof

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday November 10th, 2023.

Student number:	4455088	
Project duration:	February - November, 2023	
Faculty:	Faculty of Civil Engineering, Delft	
Field of study:	MSc Transport, Infrastructure and Logistics	
Supervision:	Prof. dr. O. (Oded) Cats,	TU Delft, Committee Chair
	Dr. A. (Alessandro) Bombelli,	TU Delft, Daily Supervisor
	Dr. W.W. (Wijnand) Veeneman,	TU Delft, Daily Supervisor
	Ir. B.J.H.F. (Barth) Donners,	Royal HaskoningDHV, Daily Supervisor

Cover: Adapted stock image retrieved from: www.andersreizen.nu

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Preface

With this master's thesis, my studies at TU Delft come to an end. Starting my Bachelor's in Civil Engineering, I've slowly gained more and more interest in the transport sector. I'm happy that I could perform my thesis in a topic, which I'm really enthusiastic, and sometimes passionate about, to obtain my master's degree.

I would like to thank my graduation committee for the last, almost, nine months. First of all I would like to thank Oded Cats, for the inspiration to find an interesting scope within the subject of air/rail substitution. Also, I'ld like to thank Barth Donners, for introducing me to RoyalHaskoningDHV, sharing interesting topics he was involved in, and ofcourse his weekly contribution by disucssing my progress. Finally, I'ld like to thank Wijnand Veeneman and Alessandro Bombelli, for both there own contribution to my work, and the critical feedback in their specific specializations. Thanks to all of you, for your contributions, and interesting insights through all of our meetings.

D. (Daan) Nijhof
Delft, October 2023

Contents

List of Figures

List of Tables

Nomenclature

AHC	Agglomerative Hierarchical Clustering
FUA	Functional Urban Area
GHG	Greenhouse Gas(ses)
HSR	High-Speed Rail
IenW	Minsiterie van Infrastructuur en Waterstaat (Ministry of Infrastructure)
IMA	Integrale Mobiliteitsanalyse (Integral Mobility Analysis)
OD	Origin-Destination
PC	Principal Component
PCA	Principal Component Analysis
TEN-T	Trans-European Transport Network

1 Introduction

Total passenger transport demand in Europe has increased by 20% between 2000 and 2019. The demand for air traffic specifically has been globally increasing (Statista Research Department 2023) and even increased by 86% in Europe between 2000 and 2019 (European Environment Agency 2022) until the COVID-19 pandemic and recovered quickly towards pre-pandemic levels in 2022 (ICAO 2022). While on one side (air) mobility demand is increasing, on the other hand, many different climate agreements are made to reduce the impact of travel on greenhouse gas (GHG) emissions. The Paris Agreement by the United Nations in 2015 has led to the long-term focused European Green Deal (European Commission 2019), agreeing to reduce GHG emissions by 55% in 2030 compared to 1990. Transportation has a significant role in GHG emissions, and its share even relatively increased by 33% from 2000 to 2019 to 29% of all GHG emissions in Europe (European Environment Agency 2022). To reduce the transport sector's impact, The European Green Deal also states that transport-related emissions should be reduced by 90% in 2050.

1.1. The role of aviation in greenhouse gas-emissions

Aviation accounts for 13,9% of total GHG emissions in European transport and is (after road transport with 43.9%) the second largest GHG emitter (Figure 1.1). Considering the European Commission's plans to reduce transport-related CO_2-emissions, it is desired to reduce the amount of air travel. However, the desire does not directly transfer into a significant change in travel behavior. The choice of travel mode depends highly on factors like pricing, frequencies, level of service, travel times, and travel purposes. Travel purposes and distances vary greatly within air transport; business and leisure passengers travel from a few hundred to a few thousand kilometers per flight. The impact on CO_2-emissions depends on a flight leg's length. According to (Grimme and Jung 2018), short-haul flights cause approximately twice as many emissions per kilometer as intercontinental, long-haul flights, mainly originating from the larger share of the take-off and ascending in the flight. However, long-haul flights' total impact remains larger than short-haul flights. The main difference between long- and short-haul flights lies in the substitution possibilities. For inter-continental travel, flying is the only reasonable option. In smaller ranges, a flight could (potentially) be replaced by other travel modes, such as car and rail transport.

1.2. Substitution of air transport for high-speed rail

As aviation is one of the largest emitters of GHG, it is relevant to investigate what rail transport in Europe can do compared to air travel. Rail transport is widely recognized as a sustainable travel mode and an important topic for national governments and the European Union regarding mobility improvement (European Commission 2021a). Much research has been conducted on the substitution of air travel for rail transport, which shows potential in a range from 5% up to 75% of flight reduction (Reiter, Voltes-Dorta, and Suau-Sanchez 2022; Bergantino and Madio 2020; Zhang, Graham, and Wong 2018; Savelberg and De Lange 2018; Donners, Buuren, and Rijniers 2018). These results strongly depend on the choice of scenarios, parameters, and timeline but show that rail transport can be considered a

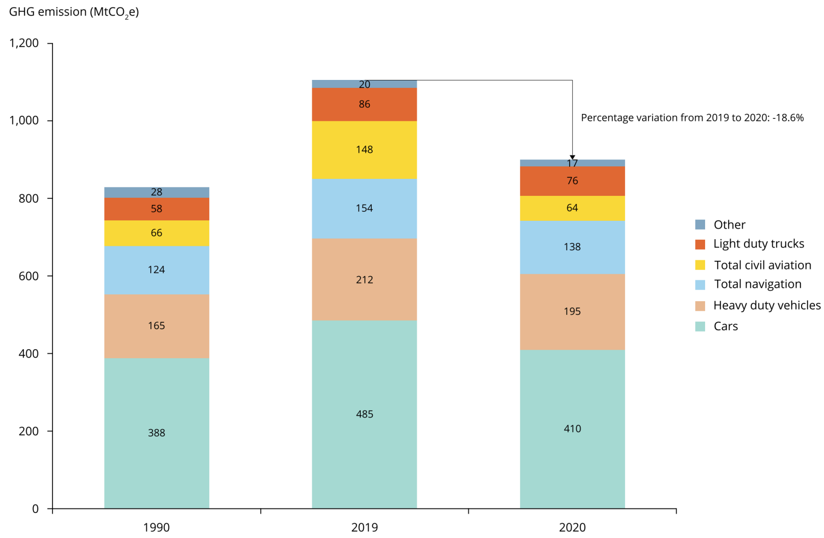


Figure 1.1: Greenhouse Gas emissions per transport mode in Europe (European Environment Agency 2022)

good option to substitute flights for rail travel.

Most research considers direct substitution, focused on origin-destination (OD) travel. Generally, these direct substitutions cover short- and medium-haul flights with a competitivity range of up to 5 hours of train travel time (Kroes and Savelberg 2019). A large share of these competitive flights are performed within a range of 750 kilometers, in Schiphol Amsterdam's case approximately 25% (Donners and Heufke Kantelaar 2019), and are a feasible option for direct substitution. However, long-haul flights have no feasible option for direct substitution (yet). At Schiphol, 19% of the flights in 2019 were performed on long-haul flights, accounting for 29% of the total number of passengers in that year (Schiphol 2019). Furthermore, 36% of Schiphol's passengers are transferring to a new flight. In that case, whether the incoming and outgoing flights are short-, medium- or long-haul flights is unknown. However, a significant share of Schiphol's passengers is transferring. Also, as most long-haul flights from Europe are intercontinental flights, many passengers cannot substitute their full journey with a different mode.

1.3. Transferring passengers at hub airports

Schiphol Airport is an example of an airport in Europe with a major hub function for air traffic. Apart from Schiphol, Frankfurt (54%; Frankfurt Airport 2019), Paris Charles de Gaulle (23%; Groupe ADP 2020) and London Heathrow (34%; Statista 2020) are known for their share of transfer passengers, based on pre-COVID-19 data from 2019. According to Maertens and Grimme (2015), transfer data of airports is calculated in different ways by several data sources. They state that the number of transfer passengers can be split into *beyond*, *behind*, and *bridge* passengers. The first type of passengers, *beyond*, starts their flight at the subject airport and transfers at the next airport(s) in their journey. For the analysis of transfer passengers, they can be added to the group of local (OD) passengers that do not connect at the subject airport. *Behind* passengers fly to the subject airport and transfer to their last leg towards the final destination. The latter type, *bridge* passengers, transfer at the subject airport and will perform at least one additional leg before and after the incoming and outgoing flight leg. Considering this theory, *behind* and *bridge* passengers can be linked to hub airports serving many transfer passengers.

1.4. Intermodality between air and rail travel

Passengers who connect in a hub airport generally transfer between two flight legs. Some have an origin and destination that allow them to substitute their journey for a complete rail trip. However, some

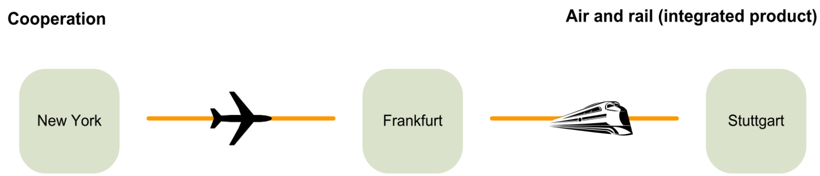


Figure 1.2: Example of intermodal travel with two cooperative modes (Duarte Costa and Abreu Silva 2012)

have an origin or destination not on the European mainland or the UK, connected to the mainland by rail. When one of the legs of their journey is performed on the mainland or in the UK, it would be possible to perform a partial substitution (Figure 1.2). This is where the concept of intermodality plays a role.

Direct substituting short-haul flights would contribute little when considering the decrease of flights for transfer passengers but is mainly effective for OD travelers. Transfer passengers already have chosen a non-direct travel option, which can be a forced choice by the availability of flights between their origin and destination, or it could be a cost consideration. When one of the legs of this trip isn't available because of regulations banning short-haul flights, that leaves a traveler two options. Firstly, smart ticketing platforms can find other travel options via plane via a hub further from the origin or destination. As a result, none of the legs is considered a short-haul flight and thus banned. An example would be a flight from München to JFK, New York, via Frankfurt. Suppose the short-haul leg from München to Frankfurt is restricted. In that case, travelers will still have the option to fly via London Heathrow, Amsterdam Schiphol, or Paris Charles de Gaulle, depending on the threshold distance for the short-haul flight ban. The second option is to improve the attractiveness of other, more sustainable travel options. If intermodal options are promoted, this could fill the gap that originated from the short-haul flight ban. To achieve this, intermodal products should be improved and easily accessible for travelers to be a feasible option.

1.4.1. Definition of intermodality

The European Commission has defined intermodality as the characteristic of a transport system that integrates at least two different transport modes (Brida et al. 2017). The access and egress legs of a journey are excluded from this concept because intermodality is focused on the main legs of a journey. In a partly intercontinental multi-leg journey, the European part could be substituted and thus become an intermodal trip. Their *behind* and *bridge* passengers could be stimulated to substitute a short-leg flight for rail transport at an airport. To do so, airports should focus on their role as intermodal hubs, accommodating air and rail travel and transfer possibilities that are as seamless as possible. Some research explores important factors to improve this intermodal hub function and vary from coordination of services (frequencies) of air and rail to the improvement of ticketing and baggage handling conveniences (Wu and Han 2022; Song, Hess, and Dekker 2018, Donners, Buuren, and Rijniers 2018, Savelberg and De Lange 2018). They also recognize that the same factors apply to choices for intermodal travel as to the substitution of air for rail travel; travel times, frequencies, and connecting times form an important basis for mode choice. These factors can be divided into two themes. One focused on services, from ticketing and baggage conveniences to the offered air and rail services and their frequencies. The second theme focuses on physical characteristics, which mainly determine travel times, e.g., high-speed rail (HSR) infrastructure availability and the geographic situation in the air and rail networks. These variables can be assessed for each (significant) airport in Europe to create insight into the characteristics of these airports.

Integration and synchronization of travel modes cover a broad field of modes and types of integration. Both factors are important aspects of the concept of intermodality. This research considers an intermodal network of rail and air transport in Europe. To assess this network, the intermodality must be clearly defined as its application to the specific air and rail transport modes. When the concept is established, defining factors for the intermodality in the European network can be collected and treated to create a database for the network assessment.

As mentioned, intermodality can be applied to various travel modes. This research assesses the intermodality between air and rail transport in Europe. This section elaborates on the specific intermodality for these travel modes and the factors that influence the intermodality of a network from a transport node's perspective.
There are various interpretations of intermodality in literature. A report by Eurocontrol distinguishes two forms of intermodality (Duarte Costa and Abreu Silva 2012). One in which the ground leg functions as an access leg from an urban area to the nearest airport (type 1). This type of intermodality can be considered as access to the airport, for which public transport is a common option. The other form, type 2, is based on integrating the airport into an extensive ground network. This study focuses on the latter, type 2, as a basis for defining an intermodal network. This type of intermodality can also be divided into short/medium distance service between 100-300 kilometers in which the rail network is considered to be a feeder, and medium/long distance service (300-800 kilometers) in which the rail leg can substitute air travel (Duarte Costa and Abreu Silva 2012). This study reviews intermodality to substitute short-haul flights by improving air/rail intermodality possibilities. Such intermodality would mostly apply to passengers who normally travel by a multi-leg journey. By intermodal options within the transport network, they can transfer from air to train, or vice versa, and substitute a short-haul flight with more sustainable train travel.

The rate of intermodal possibilities and attractiveness in a network is determined by physical integration, service synchronization, and offering additional services (such as baggage handling or integrated ticketing). With a higher integration and synchronization, and higher numbers of services for both modes, this can be defined as a higher performance of intermodality. As the research aims for a network assessment, the effect of physical integration and service synchronization is considered. The offer of the additional services is a factor that is especially important for passenger mode choice. The traveler's' behaviour would be captured in studies such as a stated preferences research. However, it is too arbitrary to model this factor from a network perspective. Therefore the policies focusing on the network are considered. These policies can mainly be made by the European government, national governments, or airports whose roles are considered in the network. In subsection 1.4.2, the factors that define intermodality from the network perspective are elaborated.

1.4.2. Defining factors for air/rail intermodality

According to (Resource Systems Group et al. 2015), there are three main components for the successful integration of air and rail transport:

1. There are superior intermodal services from the network perspective.
2. Physical facilities are available for a seamless transfer.
3. The network is supported by integrated ticketing and extra service provisions.

In several governmental recommendations, the improvement of international rail transport is discussed. Figure 1.3 shows an example of Dutch recommendations for improvements on the different aspects of international rail by Raad voor de Leefomgeving en Infrastructuur (2020). These aspects include infrastructure, traffic regulation systems, recommendations for train operators, and improvements in ticketing and passenger rights. This should be achieved and integrally executed by better international coordination of the international railway network. The network assessment considers the factors determining physical and transport services, leaving the additional services outside the research scope. Therefore, the defining factors are based on components one and two, as mentioned.

To offer a superior intermodal service from the network perspective, both sides of the intermodality should provide high-quality services. One of the most critical factors for multi-modal accessibility is the frequency of services (Wang et al. 2020). From a node's perspective, the total frequencies of air and rail travel influence this accessibility. In this research's scope, a division can also be made in the type of service. On the airside, there's a difference between substitutable and non-substitutable flights. Most multi-legged journeys contain a long-distance, intercontinental leg unsuitable for substitution. An airport offering a relatively high number of these flights will likely function as an air hub. To cover these characteristics, the frequency of intercontinental flights and the ratio of these intercontinental flights compared to the total number of flights are considered to be important factors as well.

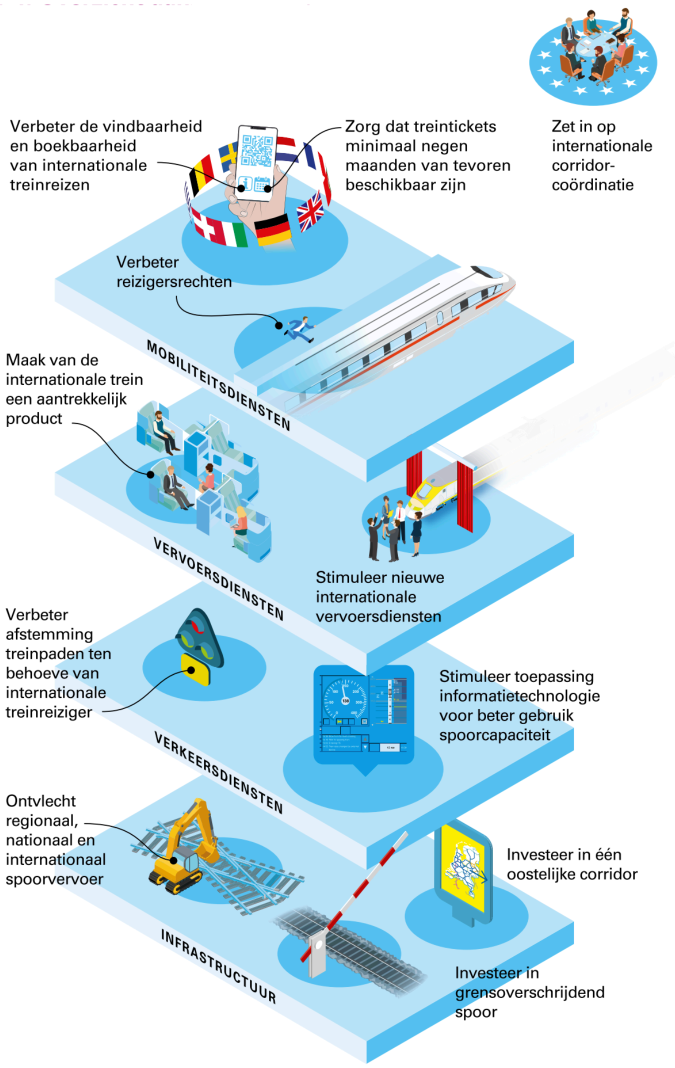


Figure 1.3: Recommendations to improve international rail travel. Obtained from the recommendations by Raad voor de Leefomgeving en Infrastructuur (2020).

Another aspect of mode choice is the offer of a variety of services. When there are more options, it is more likely that a traveler will use the node. This characteristic can be covered by the number of destinations an airport serves.

The service level of the rail side determines the other side of intermodality. In this case, the frequency of services is also one of the most critical factors. In addition, as the rail legs compete with the more common air services, the rail service should be convenient. Travel convenience is based on different aspects but is mainly determined by comfort, travel time, waiting time, and number of transfers. The number of transfers can be reviewed from the network perspective, looking at the shortest paths between nodes. Other travel (in)convenience factors can be included in a characteristic like travel impedance or generalized travel costs.

The interaction between air and rail travel can be reviewed based on the potential competition and substitution per link in OD matrices. Research by Bruno (2022) evaluated these potentials for European Functional Urban Areas. The number of links suitable for competition or substitution originating from a specific area shows the potential of flight substitution from and to this airport.

The second component mentioned by Resource Systems Group et al. considers the physical integration between modes. Examples, such as Frankfurt and Schiphol Amsterdam, show the convenience offered by on-site railway stations. The type of services provided at a station at the airport also indicates the integration into a ground network. An airport station can be connected to an urban metro-like network to access the nearby urban area or directly integrated into the national network. There's also a difference on the national level whether this network offers regular national services or long-distance and high-speed services.

The connection to the ground network can also be assessed by the travel time to the main station in the related urban area. This parameter contains a mixture of the physical distance to the urban area and the service level of the connection. With low travel times to a central station, the integration between the airport and the main rail network can be considered higher.

Finally, the geographical location can be assessed to assess the potential of an airport without considering existing rail infrastructure. Because water bodies, mountain ranges, or other natural barriers can influence travel time over a simple geodesic distance for ground travel, the geographic location can be modeled by car travel times between two locations. The road network is the most advanced ground-level network and can be a good measure of travel distances.

1.5. Different perspectives on intermodality

The current performance of intermodality, defined earlier as a measure of the integration and offer of services, is currently still limited. This is caused by a lack of extensive intermodality opportunities and differences between the two travel modes regarding governance, economic interests, and travelers' experiences. Throughout Europe, there are and have been some initiatives to stimulate intermodality. Examples of successful air/rail service products in Germany, Switzerland, and France include integrated ticketing, timetable coordination, and luggage services (Beeravelli et al. 2022). However, no international integrated products are offered from a European network perspective. This is also partly caused by the lack of coordination between air and rail travel from a governance perspective.

The businesses in the aviation market focus on customers, considering the fares, travel conveniences, and comfort. Rail companies focus more on their position in the national transport market (Raad voor de Leefomgeving en Infrastructuur 2020). To achieve an effective intermodal product, both markets should align their interests. A decision by the European Union on the approach of intermodal services and the roles that air and rail companies play in it can help align their interests to develop an effective intermodality policy.

The international policies on intermodal travel depend on the separate approaches to air and rail travel. In the rail market, it is seen that national governments mainly focus on their national market (Witlox et al. 2022). With this focus, international interests are sometimes neglected. Initiatives like TEN-T (European Commission 2021b) address the problem of international train travel and propose solutions to expand it. However, when creating new corridors and networks, regional and national borders are crossed, which requires coordination of policies between municipalities, local governments, and national governments (Priemus and Zonneveld 2003). These policies consider the transportation sector but also have to allow for economic activities, societal effects, and the environment. The multi-level governance and multi-actor policies differ from studies regarding Chinese infrastructural projects, where such side aspects are considered less, and no national borders have to be traversed.

In air travel governance, there's a trend of banning super-short-haul flights, as is enforced in France (Euronews 2022). In the meantime, the research discusses the effectiveness of such policies (Dobruszkes, Mattioli, and Mathieu 2022) and opts for a focus on longer-haul flights. This offers a dilemma between environmental policies and the satisfaction of airlines, which often possess a strong lobbying position. Substitution of long-haul flights is not feasible, and the more environmental options of electrical or hydrogen-fuelled aircraft currently are only feasible on shorter distances. Therefore, removing flights seems to be the only feasible option when aiming for environmental improvements, which would reduce mobility and limit airlines' businesses.

Banning flights is a difficult topic considering the governance involved in aviation. National governments and the European Union can only stimulate or discourage specific flights and destinations by implying different policies (Dutch Ministry of Infrastructure (IenW), Appendix B). The governments cannot oblige air operators to use or discontinue specific flight legs. These decisions are made by a specific authority that assigns the slot to airlines. A local government can limit the number of flights an airport may facilitate. However, these discussions are politically difficult, as recent examples of the situation at Schiphol Airport show (NOS 2023). This situation also shows the difficulties with the number of stakeholders involved in this subject. Local governments have to deal with European obligations or with diplomatic discussions. Also, airline operators and airports have a strong economic position and have a large influence on the decisions of the governments. Moreover, on the rail side of these difficulties, nationally, the governance approach of railway operators, concessions, and railway managers differ a lot. European coordination would need all involved parties to be aligned, and differences in their governance approaches should be overcome to achieve a joint solution.

Economic factors could also contribute to intermodal cooperation. Governmental institutions are searching for opportunities to limit polluting transport modes. France was one of the first countries to ban short-haul flights that serve OD pairs, which can be reached within two and a half hours by HSR (Euronews 2022). Cooperation between air and rail companies can benefit both when these policies limit the aviation market. Aviation can profit by maintaining the passenger supply to serve their longer-distance flights, while rail companies could reach more customers when offering services to airports. Furthermore, it is shown in a Chinese study (Zhang, Yang, and Wang 2017) that the introduction of parallel HSR services impacts the air transport demand, becoming more elastic. To cover these uncertainties for air travel, cooperation could help identify OD pairs where the most severe impact will occur and where opportunities lie to strengthen connections.

The convenience and reliability of services are considered essential drivers for choosing rail access to airports from the passengers' perspective (Sperry et al. 2012). These are mainly indicators of the network performance. However, the travel experience for passengers plays a vital role as well. A study on intermodal bus and rail travel in Nanjing (Chen, Stathopoulos, and Nie 2022) shows that the choice of a transfer station is based on trip attributes of all involved modes but is mainly determined by the attributes of the slower, less comfortable modes. Also, stations with more amenities are more likely to be chosen as transfer stations. Translating this to air/rail intermodality, it can be expected that the 'slower' access mode of rail travel is an important determinant for air/rail intermodality. Also, the services offered at an airport determine intermodal attractiveness. The travelers' perspective shows that the attributes of individual travel modes and the attractiveness of the hub both play a role in intermodal travel.

1.6. Network perspective on intermodality

As the points above suggest, increasing possibilities for intermodality is mainly a choice by governments, transport companies, and travelers. However, also a technical part plays a role. International train travel is in the process of international technical coordination in Europe regarding track characteristics, safety systems, and power supply. Therefore, the international rail network should also show improvements to facilitate an optimal intermodal network. (Raad voor de Leefomgeving en Infrastructuur 2020). Also, the network's service performance can be considered a main determinant for intermodal possibilities. Offered services, like frequencies or the directness of the service, are important factors for intermodality from a network perspective. The network's performance can be assessed from the level of served links but can also be considered from the node perspective, which focuses on the services offered per airport.

1.6.1. Role of airports in an intermodal network

As every airport has a certain role in the European air network, they would also have a role in a European intermodal network. In air networks, the concept of hubs is already widely known. Most airports function within a hub and spoke network for mostly economic reasons. Some research shows that

airports within a network can be clustered and classified based on the airports' data. Ryerson and Kim (2013) consider the air network in the United States and find 5 clusters of hubs, which can be assigned to different tiers, based on three parameters: airport connectivity, passenger volumes, and number of flights offered by the airport. Over time as these parameters can change over time, a change in clusters to which airports were assigned is shown.
This example shows classification possibilities in purely air-oriented networks. The same concept could roughly be applied to the classification of intermodality in airports. Chen et al. (2022) added the geographical distance between airports and HSR networks as an extra variable to their cluster analysis and concluded a classification of air-rail interaction in China (Figure 1.4). Towards an even more elaborated intermodality classification, a wider variety of variables could be added to the clustering, as mentioned in the previous paragraph. Concerning these classification possibilities, some European airports already offer services, as mentioned in the previous paragraph, which could also be considered intermodal hubs. Also, some airports, such as local airports, do not show any hub function in the airport but would also have no potential to be considered an intermodal hub. Between these cases, a variety of intermodality in European airports could appear. Because intermodality is based on more factors (the characteristics of both air and rail networks and the integration between the two), clustering probably would show a wider variety of differences in metrics. Therefore, a classification also would have a less linear characteristic.

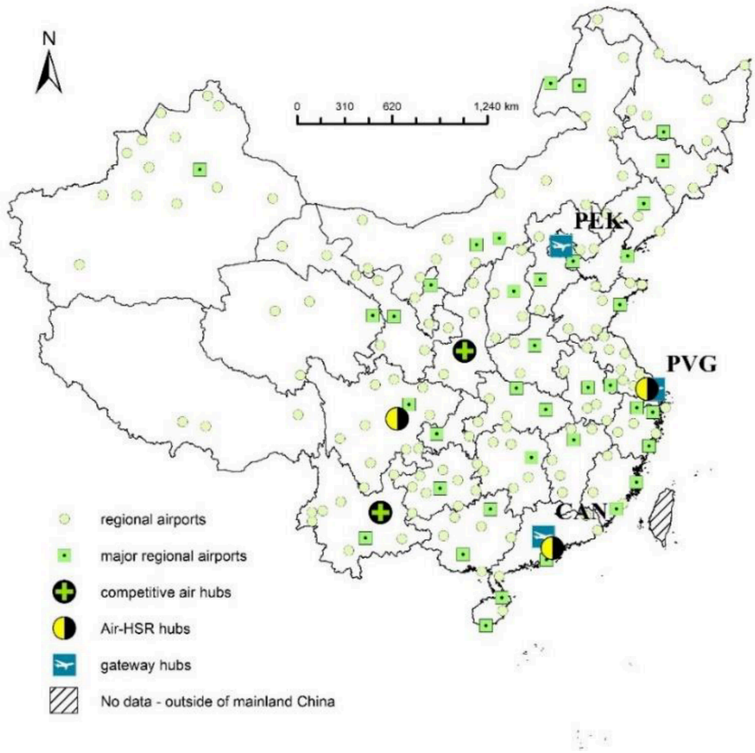


Figure 1.4: Classification example for Chinese air rail interaction from Chen et al. (2022)

The classification of the airports' intermodality could identify the role of specific airports and their potential in an intermodal European network. This is interesting to show the situation in the current network, but it can also be used to assess future possibilities for more extensive intermodal options. The European Union has created plans for a Trans-European Transport Network (TEN-T; European Parliament 2021). They assigned requirements to this network focusing on road, rail, and water infrastructures, but also on the facilitation of intermodality in which they aim to improve the air-rail connection between airports and the TEN-T (Wortel and Kaptein 2023). The underlying land network in these plans is shown in Figure 1.5. The major institutions for air (IATA) and rail (UIC) transport have also signed a Memorandum of Understanding to express both parties' intention to create strong cooperation to improve intermodal possibilities between the two travel modes (IATA 2020). These mostly focus on digital platforms for seamless ticket retailing and to create new business models aimed at intermodal transport solutions.

These plans demonstrate that there will be probable progressions in intermodal travel in the future. Therefore, the earlier mentioned clustering could be applied to scenarios to reveal the potential impacts of these plans. It can show possibilities of improved classifications of specific airports, but more importantly, it can also show effects throughout the entire network. Considering current airport and network loads, the impact of an improved network intermodality can be linked to the number of enabled passengers to travel via an intermodal trip. This would not quantify increased mobility but will quantify the number of passengers that can opt for an intermodal trip. To assess the choices these passengers will eventually make, new research can be constructed to perform a choice experiment considering intermodal travel in these scenarios that may occur in the future.

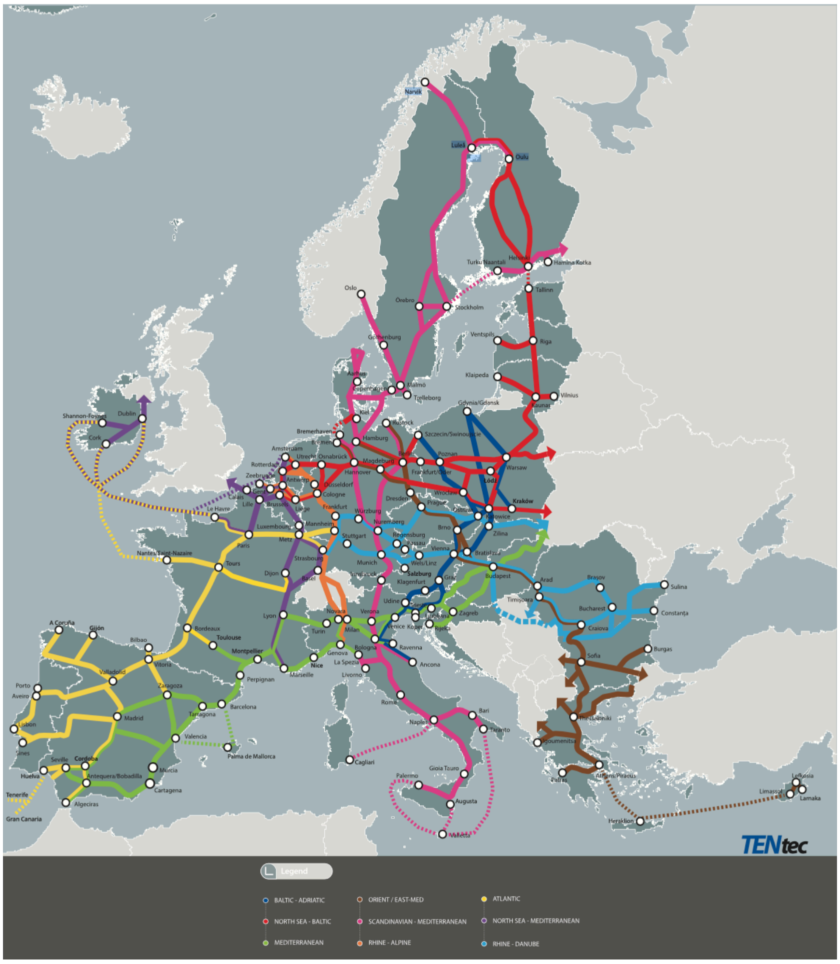


Figure 1.5: Underlying land network for Europe's TEN-T plans (European Commission 2021b)

1.6.2. Possible effects of an improved intermodal network

When the improvement of the intermodal network in Europe can lead to a reduction in short-haul flights, this can be reviewed from different perspectives. At first, reducing the relatively more polluting short-haul flights (per kilometer) by rail will directly reduce GHG emissions because rail transport is a more sustainable alternative than air travel. However, many European airports must deal with limitations caused by infrastructural capacity problems or environmental restrictions. In both cases, airports have a limitation on the number of flight movements per day. When short-haul flights are canceled, this will free the so-called slots at the airport and enables an airport to assign these slots to flights over a longer distance (Ministerie van Infrastructuur en Waterstaat 2020). Considering the mobility perspective, this will improve overall mobility because (more) destinations can be reached on a higher frequency. Whether this mobility improvement will satisfy the existing demand for these destinations or enable a latent demand by increasing travel options is questionable (Givoni and Banister 2006)). Meanwhile, environmentally considered, these long(er)-haul flights have a higher load on GHG emissions (Grimme and Jung 2018), and thus will cause a higher total GHG emission in aviation. When short-haul flights are substituted, a good balance between improving total mobility and reducing GHG emissions by the transport sector should be found. Governments could restrict airports in their available slots if they aim to reduce flights to minimize GHG emissions, as the Dutch government already restricts Schiphol to limit nitrogen emissions.

The balance between increasing mobility and reducing the transport sector's GHG emissions exposes all stakeholders' perspectives. Apart from some environmental skeptics, who are exceptions, all involved stakeholders will acknowledge the need for a more sustainable world. However, the different stakeholders also have their interests. Environmental organizations focus on pushing people and organizations towards lifestyle or business operations that are as sustainable as possible to limit, or even turn around, the negative impact people have (had) on the environment. On the other end of these perspectives, the transport sector (airports, air and rail operators) and its passengers aim to reduce the footprint of the transport they facilitate or use. However, they are restricted by operational costs, aim for profit, or the desire to reach specific destinations. Therefore, the road to more sustainable modes of transport will be slowed down by mixed interests. In between, governments have to make difficult trade-offs between their environmental agreements and the economic benefits of transportation. With the Paris Agreement 2015, governmental institutions have a hard deadline to achieve the set emission goals. Still, they will maintain or increase their country/region's mobility as much as possible. The latter perspective is interesting to consider when interpreting the expected impact of intermodality encouragement.

1.7. Problem statement

Imposed by the European Green Deal, the transport sector must reduce its total CO_2-emission by 90% in 2050. Substitution of short-haul flights for rail transport is a known concept with the aim of GHG reduction and is widely researched (Reiter, Voltes-Dorta, and Suau-Sanchez 2022; Bergantino and Madio 2020; Zhang, Graham, and Wong 2018; Savelberg and De Lange 2018; Donners, Buuren, and Rijniers 2018). These studies mainly focus on direct substitution of OD traffic in short- to mid-range flights. However, it is not feasible to substitute the (entire) trip for long-haul and intercontinental travel. Some of these trips exist of multiple legs, where the possibilities lie in partial substitution within the hub & spoke function. Therefore, intermodality can play a role in improving transport sustainability. Research in this specific part of substitution is less extensive than direct substitution. However, the topic is gaining interest, as is research into intermodality (Duarte Costa and Abreu Silva 2012). To stimulate intermodal travel, airports must focus on their position in the air and rail networks and the integration between the two modes. This integration can be improved by several factors regarding service, like frequencies in both networks or integrated ticketing and luggage services. Also, the integration can be improved regarding the physical network characteristics, like HSR infrastructure availability and geographic situation (Resource Systems Group et al. 2015).

To explore the possibilities of a European intermodal network, it is interesting to discover the current position of airports regarding intermodality. Their performance and characteristics linked to the factors of intermodality can show the air network's status and possibilities of the airports. A similar study has been conducted in China (Chen et al. 2022). A similar network in Europe has some similarities with China. Still, more research on the European case is needed because of the international character and the different administrative cultures in the European Union. Considering the entire European network, classifying the airports' intermodality can provide insights into the networks' performance and the airports' position within. Also, implementing European plans to improve an intermodal network can assess the possibilities and change of position in such a network. This can help identify focus areas in Europe when such policies are implemented.

The questions around the status and possibilities for an intermodal network in Europe and the position of European airports within show a possible outline to research this network. With data for all significant European airports, a clustering of airport typology can be conducted. Subsequently, the clusters can be classified to assess the role of airports within the network. Implementing European intermodality plans in the constructed data set can explore possibilities and changes in classification within the network. This study presents a data-driven approach to identifying intermodal hub potential and the effects of policy implementation in an intermodal European network.

1.8. Research outline

This research to explore European intermodality, its potential, and the role of European airports in the network can be summarized by the following main research question:

Where lies the potential of air/rail intermodal policies within a European intermodal network?

The research entirely focuses on European airports with an international function within the aviation network. The data set will be filtered with a lower bound of air and rail connections in proximity and by total passenger numbers to limit implementation efforts in further research steps. The intermodal network will consist of an air network and an HSR network. The air network contains internal direct European flights and the number of intercontinental flights leaving the airports. The HSR network contains links with potential for substitution, as defined by literature on air/rail substitution. The research aims to assess the (potential) European intermodal network and the role of specific airports as intermodal hubs within this network.

To answer the main question, it is essential to determine the used definition of Air/Rail integration. Also, the variables used to classify airports on their intermodal hub function should be selected carefully to classify the current situation and possibilities after scenario implementation. These combined steps help scope the problem to a data set focusing on network infrastructure and services. Using this data, the airports in the intermodal network can be clustered based on intermodal performance and potential. Subsequently, the implementation of European intermodality plans in the data set will showcase the impact on the intermodal network and possible changes in the role of specific airports in the network. The potential expansion of intermodality will give some passengers an extra option of mode choice, which can show the impact of the implementation of European plans on the network.

The following sub-questions can answer all mentioned aspects:

1. What variables should be used to quantify the degree of air/rail intermodality at European airports?
2. How can the airports be classified in the current European intermodal network?
3. What European plans can affect the intermodal network and how?
4. What is the impact of European intermodality plans on airport classification in the European intermodal network?
5. What effect does change in classification have on the potential of intermodal travel in Europe?

These questions will guide the research to an answer to the main question. The study aims to show the possibilities of a European intermodal network and the effects current plans for intermodality will have on the airports' positions.

1.9. Research contribution

This research aims to create a method to assess intermodality within a network of two different modes, with, in this case, air and rail transport. This has already been conducted in a Chinese case study (Chen et al. 2022). Assessing the case of a European network will extend the analysis to a further focus on HSR infrastructure and services and add an extra case study to airport classification regarding intermodality. This approach first identifies a benchmark situation on the European network. This is used to assess the current situation of intermodality in the network and shows the comparability of airports based on the offered services of the airport and the corresponding area.

The novelty of this approach mainly lies in the scenario implementation. Scenarios of possible policy implementations considering air/rail intermodality are constructed. These are used to estimate changes in the data and assess a new classification of airports in the European intermodal network. The differences in classification can help identify the effects of a policy implementation to see what airports could become important nodes in an intermodal network. This could help governmental institutions to identify focus areas where improvements in rail infrastructure and intermodal hub facilities are needed.

This study uses different data than the Chinese classification study (Chen et al. 2022). The Chinese example assesses service and physical aspects for the air side. However, Chen et al. only assessed physical distances between airports and HSR stations, limiting the role of HSR in their study. The air and rail services should be equally reflected in the data to assess intermodality fully. In contrast, this study will consider the specific characteristics of air and rail services. Also, the physical location of the airports will represent the location within the network, while the location compared to HSR stations will be used as a variable of in performance intermodal opportunities.

The wider variety of data also enables various changes due to policy implementations. Therefore, policies focusing on air service, rail service, or physical integration can be reflected in the data. Still, implementing these policies in the data set depends on assumptions about their effects. Thus, the method does not supply an exact approach to assess policies but a technique to evaluate well-argued assumptions about these policies.

This method can also be used on other networks when different variables are used. For example, an urban network can also use this approach to assess the intermodality between urban transport modes, like buses and trams, and a national/regional train network. Also, the interaction between continental and intercontinental air travel could be assessed, concluding that the method is suitable for assessing two-level networks on different scales.

Royal HaskoningDHV (RHDHV), the company involved in this study, can benefit from the research results by applying them and the resulting tool to identify the potential of specific air/rail combinations in Europe. They can use this information in projects regarding European international train travel and for HSR integration in airport projects. The advisory position of a consultancy firm, such as RHDHV, links them to governments that have to decide on air/rail integration projects. RHDHV can use their scientific knowledge and apply it to the method from this study to help governments gain insight into the impact of plans for intermodality on a network level. With some adaptations, the method could also be applied to lower-level transport systems. Besides consultancy firms, the operational companies involved in air/rail travel can also benefit from the study. The findings from a network analysis can help them decide on strategic choices. If a specific airport can gain a more central position in a European intermodal network, it can use this knowledge to respond to expected policies and enter related business opportunities.

1.10. Structure of the report

This report contains a step-by-step process for classifying airports in an intermodal network. In chapter 2, the construction of the method for the network analysis is elaborated. Also, the defining factors for intermodality are converted to variables that can be used in the clustering method to assess the role of airports in the intermodal network. A benchmark scenario is set in chapter 3. The benchmark section first presents descriptive statistics of all used variables. Then, the methodology is followed to cluster the airports in the benchmark situation. This cluster analysis uses 2019 data to assess the current status of intermodality in Europe. Here, all theoretical steps in finding the correct clusters are explained, which results in a clustering of the benchmark and an explanation of the mean variables and the geographical situation of each cluster. The clusters are also labeled in different classifications to combine the findings from descriptive statistics with the describing variable data from each cluster. The classification helps understand the role of the airports in each cluster in the benchmark scenario. After setting the benchmark, chapter 4 construct scenarios to assess the effects of the clustering when policy scenarios are implemented. These scenarios are based on visionary documents by (inter)national governments and expert interviews with policy officers of the Dutch Ministry of Infrastructure. The effects of these scenarios are then implemented in the data set to cluster the new situation affected by the scenarios. To assess these effects, the new data statistics and the clustering results are compared with the benchmark scenario. For each scenario, the classification from the benchmark scenario is reassessed. These changes in classification are connected to the passenger data of the European airports to see how many people are enabled for an intermodal trip now and in different scenarios. Finally, the analysis results are discussed in chapter 5 and linked to the factors regarding intermodality, as acknowledged in chapter 1. Also, a discussion on the link between mathematical models and the implementation of their results is presented, together with recommendations for further research and the use of this method.

2 Methodology

Identifying the potential effects of air/rail integration can be achieved by studying different aspects as discussed in subsection 1.4.2. This study investigates the network perspective of intermodality in Europe. Therefore, a method to assess the intermodal network should be constructed. The analysis aims to identify the location of the intermodal potential and focuses on the position of the individual airports in the larger network. This raises questions about the airports to be assessed, the important factors to evaluate airports' positions and the scientific methods to distinguish the selected airports based on the chosen variables. To answer these questions, a method is constructed to transfer the factors of intermodality into insights into the European intermodal network.

To interpret the role of airports in the network, a shift from quantitative network data towards a classification of airports' positions in the network should be made. A constructed model treats the multidimensional data obtained from the determining intermodality factors and performs clustering and classification techniques based on the available data. The classification helps the interpretation of the current situation of intermodality in Europe. Moreover, the addition of scenarios considering European plans for future policies helps assess the effects that specific changes can have on airports within an intermodal network. The total model is built by several steps of data treatment and sequent modeling techniques, shown in Figure 2.1 and listed as follows:

1. **Data treatment & calculations**
   Construction of a data set representing the key variables to classify node intermodality in the network.
2. **Principal Component Analysis (PCA)**
   Construction of principal components (PCs) based on the key variables of intermodality to reduce dimensionality and to statistically weigh variables.
3. **Agglomerative Hierarchical Clustering (AHC)**
   Clustering airports in data set based on the obtained PCs.
4. **Classification**
   Interpretation of the constructed clusters and identification of airports' roles in the intermodal network of Europe.
5. **Scenario implementation**
   Implementation of European plans for intermodality to the current network. The clustering process will be redone to assess the differences in clustering and classification.
6. **Intermodal Possibilities Evaluation**
   Interpretation of the changes in clustering and classification. For a quantitative analysis of the effects, the clusters will be linked to the passenger data of their airports.

2.1. Airport data on intermodal airport characteristics

The introduced factors for intermodality assessment are based on data from air transport, rail transport, and manual information retrieval on airport information. Given that the intermodal network and airports'

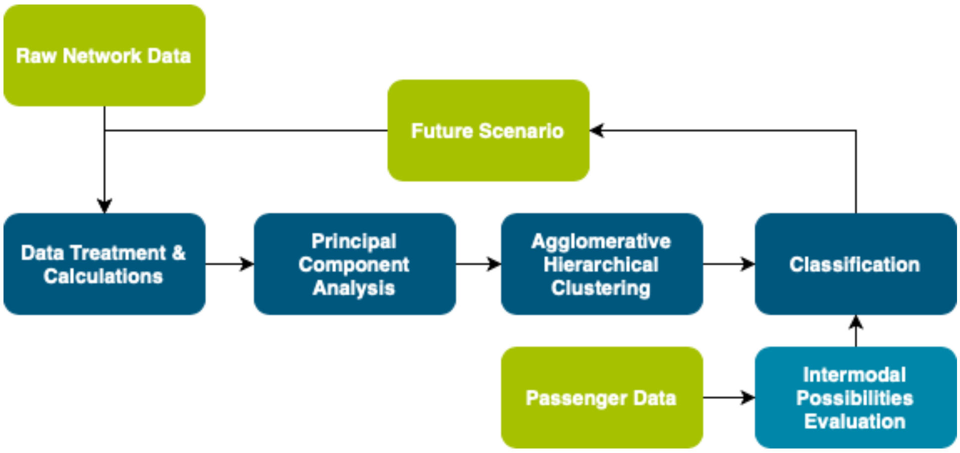


Figure 2.1: Modeling structure

role within this network are assessed, the data is considered from an airport perspective. Regarding air transport data, the connection with airports is very clear. For rail transport data, however, the connections around the airport are important, and access times to the airport are also observed to consider the connection between the airport and the rail network.

2.1.1. Criteria for airport selection

The air traffic data is retrieved from the Eurostat Browser (European Commission 2022). A detailed database per country with all significant connections from the countries' main airports shows the number of passengers, flights, and seats on these connections. This data is available for 35 European countries, from which four countries are excluded, marked red in Figure 2.2: Iceland, Ireland, Malta, and Turkey. The other countries, which are included, are marked green. Because the airport's role in the intermodal network is assessed, only airports connected to the rail network are included in the data set. Airports that are part of an included country, but cannot be connected to the rail network, are individually excluded, represented by the red markers in Figure 2.2. Besides, some French overseas islands and Svalbard (Norway) are excluded, but not shown on the map. Lastly, some inconsistencies may occur with the current (2023) situation, because older data was used. The airport Berlin Brandenburg has already been finished, but is not considered in the data used in this study. The study includes Berlin Schönefeld (both ICAO-code EDDB), the previous airport on the same location. Other discrepancies are not known, but not ruled out.
The selection of airports is then evaluated on annual passenger data taken from the pre-COVID year of 2019 to narrow down the final selection. Two thresholds are introduced to only include airports of significance in this study. The first is based on the total number of passengers. Therefore the annual passenger numbers for each connection are summed for each origin airport. A lower bound of one million passengers is set, as shown in Equation 2.1. Also, every individual connection is reviewed on the number of passengers. A reasonable boundary for substitution would be when all passengers almost fill one large high-speed train. A Eurostar, or a double Thalys, train has a capacity of 750 people. With a load factor of 90%, a daily service would transport approximately 250,000 people. Therefore, an airport will only be excluded if at least one connection transports over 250,000 people. The conditions are shown in Equation 2.1 and Equation 2.2,

$$\sum_{j \in S^D} pax_{ij} \geq pax^{min\,tot} \quad \forall \quad i \in S^O \tag{2.1}$$

$$pax_{ij} \geq pax^{min\,link} \quad \forall \quad i \in S_{origins},\ j \in S_{dest} \tag{2.2}$$

where:

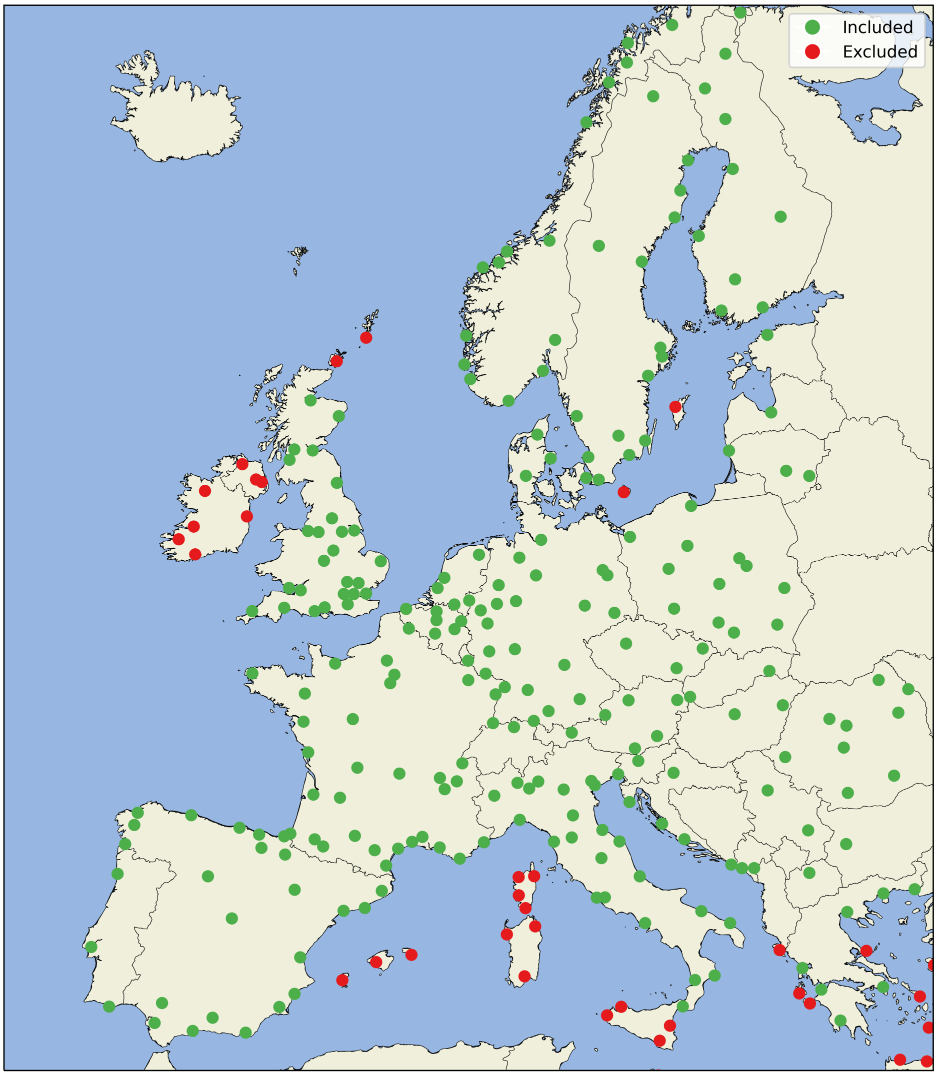


Figure 2.2: Selection of included countries in air data and individually excluded airports.

S^O = Set of origin airports in Europe,
S^D = List of all unique destinations from origin airports in Europe,
pax_{ij} = Annual number of passengers from airport i to airport j,
$pax^{min\,tot}$ = Minimum value threshold of total annual passenger numbers, set to 1 million passengers,
$pax^{min\,link}$ = Minimum value threshold of annual passenger numbers on unique OD pair, set to 250.000 passengers.

For the preliminary selection of airports, rail data is collected from (Bruno 2022). This research studied the connectivity of the air and rail network in Europe. Data in this study was based on Functional Urban Areas (FUAs) and their demographic and economic statistics. The set of 125 FUAs partly overlaps with the airports from this study's pre-selection. To connect the airports to the rail data from the Bruno study, a manual allocation of airports to the urban areas is performed. A set of 22 airports could not directly be linked to any of the FUAs, which are shown in Figure 2.3. Due to the lack of data for these locations, concessions are made about the treatment of these airports. The individual decisions are explained in Table 2.1. Most exclusions are supported by their lack of connection to the main rail network and a lack of significant data on locations in proximity. Helsinki is the only example with no direct connection to the European network, only via Russia. The included airports are linked to a significant FUA nearby that has rail data available. The access time of these airports would be relatively high compared to airports linked to their own FUA to show their distance to the rail network. The final selection of airports contains 113 airports in Europe that are accessible by land. The full selection can be found in Appendix A.

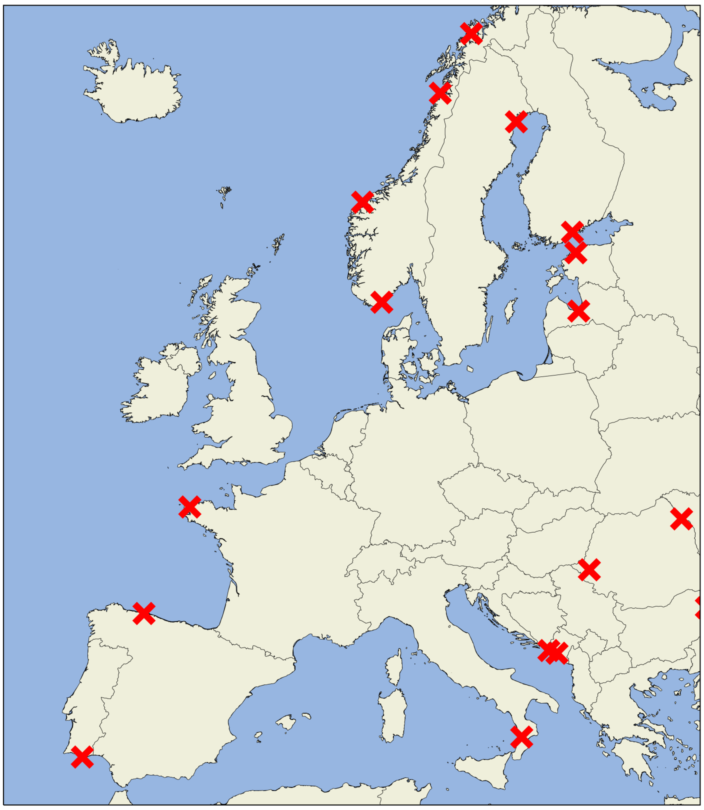


Figure 2.3: Airports excluded due to lack of available rail data

2.1.2. Variables used in classification process

From the definition of defining factors in network intermodality, the frequency and type of services, the geographical location, and the physical integration of travel modes in a node are the main aspects of mode integration. The combination of these factors supports the chosen variables used in this study. All used variables are explained in Table 2.2.

Table 2.1: Airports with no direct FUA and their treatment in data

	ICAO	Location	Treatment
Excluded	EFHK	Helsinki (Finland)	No direct connection to European rail network
	ESPA	Luleå (Sweden)	Remote location in Europe
	ENTC	Tromsø (Norway)	
	ENBO	Bodø (Norway)	
	ENAL	Ålesund (Norway)	
	EETN	Tallinn (Estonia)	No central role in railway network
	EVRA	Riga (Latvia)	
	ENCN	Kristiansand (Norway)	
	LFRB	Brest (France)	
	LEAS	Asturias (Spain)	
	LPFR	Faro (Portugal)	
	LICA	Lamezia Terme (Italy)	
	LDDU	Dubrovnik (Croatia)	
	LYTV	Tivat (Montenegro)	
	LBWN	Varna (Bulgaria)	
	LRIA	Iași (Romania)	
	LRTR	Timișoara (Romania)	
Included	LEXJ	Santander (Spain)	Coupled to FUA of Bilbao
	LEGR	Granada (Spain)	Coupled to FUA of Málaga
	LIRP	Pisa (Italy)	Coupled to FUA of Florence
	LIPH	Treviso (Italy)	Coupled to FUA of Venice
	LIBR	Brindisi (Italy)	Coupled to FUA of Bari

Table 2.2: Overview of used variables

	Variable		Description
Air services	TF air	Total frequency	Total number of flights departing from the airport, representing the average frequency of air services.
	IF air	Intercontinental frequency	Number of flights to intercontinental destination, representing non-replaceable flights.
	IR air	Intercontinental ratio	Share of intercontinental flights, compared to the total flight frequency. Differentiates airports based on a focus on short- or long-haul services.
	NC Air	No. of destinations	Number of unique significant destinations served from the airport by air, representing the variety of services from the airport.
Rail services	TF rail	Total frequency	Total number of rail services departing from the Functional Urban Area linked to the airport, representing the average frequency of rail services.
	SP rail	Average shortest path	The average shortest path within the rail network studied by (Bruno 2022). Representing the possibilities for direct and convenient train journeys to the airport.
	CN rail	Rail connectivity	The node connectivity calculated for the rail network studied in (Bruno 2022). Representing the relation between connection intensity and travel impedance at the airport.
Physical characteristics	STA	Station availability	Dummy variable considering the availability of a (HSR) station. Valued zero when unavailable, one for a conventional station and two for an HSR station.
	SAC	Station accessibility	Access times from the airport to the main railway station in the linked FUA, retrieved using a Google Maps API. Representing the travel time to the main railway connections.
	AR rail	Airports in range	Number of airports on the European mainland within potential HSR range from the airport. Based on a 6-hour travel time by car, retrieved from the Open Source Routing Machine API. Representing the geographical location of the airports within the network.

2.1.3. Air side variables

Regarding the air services, four variables are constructed. These are mainly observing the frequency and type of service as one of the main aspects of intermodality. All frequencies included in the data are the total number of flights per year to create an image of the average services. For example, seasonal differences in the case of touristic destinations are not considered by this simplification. The first variable considers the frequency of all air services at the airport. Therefore, the commercial flight data per country from European Commission is sorted for each origin airport. Subsequently, the number of annual flights per connection is summed to determine the total annual flights at each main airport in that country. This variable represents the magnitude of the airside of the airport. According to Duarte Costa and Abreu Silva (2012), the frequency is, in this case, more important than, for example, a total passenger load, given that a higher frequency will reduce waiting times and improve intermodal possibilities. The variable is defined in Equation 2.3,

$$f_i^{air} = \sum_j n_{ij}^{flights} \quad \forall \quad i \in S^O,\, j \in S^D \tag{2.3}$$

where:

S^O = Set of origin airports in Europe,
S^D = List of all unique destinations from origin airports in Europe,
$n_{ij}^{flights}$ = Annual number of flights from airport i to airport j.

The total frequency is also filtered for the frequency to intercontinental destinations. To determine the set of airports considered intercontinental, the list of airports on the European mainland is subtracted from the list of all unique destinations. This list of intercontinental airports represents the airports that would not be reachable by any rail connection from the European mainland. The same data is used for this variable as for the total frequency. However, the data is sorted for each origin airport only to contain destinations selection included in the list of intercontinental destinations. The frequency of intercontinental destinations represents the annual number of flights suitable for partial substitution. Those specific legs can not be replaced; however, potential access or egress flights for those intercontinental trips can be replaced. The definition of this variable is shown in Equation 2.4,

$$f_i^{IC\,air} = \sum_j n_{ij}^{flights} \quad \forall \quad i \in S^O,\, j \in S^{IC} \tag{2.4}$$

where:

S^{IC} = List of all unique intercontinental destinations from origin airports in Europe.

While it is interesting to differentiate airports on their focus of service, a ratio of intercontinental flights compared to the total number of flights. This will highlight smaller airports, regarding the total number of services, with a relatively high number of intercontinental flights. The ratio is defined in Equation 2.5.

$$r_i^{IC\,air} = \frac{f_i^{IC\,air}}{f_i^{air}} \quad \forall \quad i \in S^O \tag{2.5}$$

Finally, the number of unique destinations with a significant load is considered. This number represents the variety of services that are offered at the airport. A threshold of three weekly flights is assumed to filter the regular connections,

$$n_i^{SC} = \sum_{j \in S^D} X_{ij} \quad \forall \quad i \in S^O \tag{2.6}$$

where:

n_i^{SC} = Number of significant connections,
$X_{ij} = \begin{cases} 1, & \text{if } f_{ij} \geq 3 \text{ per week} \\ 0, & \text{otherwise.} \end{cases}$

2.1.4. Rail side variables

The rail services are based on the data available from (Bruno 2022). This data is calculated for a 125x125 OD matrix with a 125 Functional Urban Areas selection. Using this data, it is assumed that this selection of FUAs forms the European rail network. With the linkage between the selection of airports in this study and the FUAs from Bruno, some FUAs are not used, and some are used double as an origin. However, the summation of different variables is performed on the origins within the selection of airports to all 125 destinations, considered the entire rail network.

The first variable to consider rail services is calculated from the total number of services. This is based on the weekly schedule of long-distance trains in an average week. Because the data is standardized before performing the PCA and AHC, there is no need to convert the data to annual numbers. The frequency of services is retrieved in an OD matrix. The sum of services to all destinations is taken to obtain the total number of services for the FUAs assigned to the selected airports, as shown in Equation 2.7,

$$f_i^{rail} = \sum_{j \in S^{FUA}} n_{kj}^{rail} \quad \forall \quad i \in S^O \tag{2.7}$$

where:

S^{FUA} = Set of all 125 FUAs in the observed rail network,
n_{ij}^{rail} = Weekly number of rail services from FUA k to FUA j,
k = FUA k assigned to airport i.

Because it is observed whether a single flight leg can be replaced, it is desired that the replacing rail leg is as direct as possible. The average of the shortest paths in the rail network is calculated to measure the directness of the services to an airport's FUA. The shortest path represents the lowest number of services, i.e., the number of transfers, plus one, needed to travel from A to B. To obtain an indicator for one single origin, the average value of all shortest paths from the origin is taken, as defined in Equation 2.8,

$$SP_i = \frac{\sum_{j \in S^{FUA}} sp_{kj}}{N^{FUA}} \quad \forall \quad i \in S^O \tag{2.8}$$

where:

S^{FUA} = Set of all 125 FUAs in the observed rail network,
N^{FUA} = Number of FUAs in set S_{FUA},
sp_{kj} = Shortest path between FUA k and FUA j,
k = FUA k assigned to airport i.

Another indicator representing travel convenience is the node connectivity calculated in (Bruno 2022). The node connectivity is calculated as the sum of the link connectivity on all links from one origin (Equation 2.9). The link connectivity represents an indicator that combines the effective frequency on the link with the travel impedance experienced (Equation 2.10). The travel impedance is a weighted travel time, including waiting time and a transfer penalty. The node connectivity quantifies the average attractiveness of links departing from that node. This indicator shows some overlap with the average shortest path. However, both are included in the data set to identify airports with relatively long travel times due to their distance from the network but a good connection to the railway network. The connectivity indices are defined in Equation 2.9 and Equation 2.10,

$$NC_i = \frac{\sum_{j \in S^{FUA}} LC_{ij}}{N^{FUA}} \quad \forall \quad i \in S^O \tag{2.9}$$

$$LC_{ij} = \frac{\theta^f f_{ij}}{\theta^{IV} tt_{ij}^{IV} + \theta^{OV}(tt_{ij}^A + tt_{ij}^W + tt_{ij}^E) + \theta^T tt_{ij}^T} \quad \forall \quad i, j \in S^{FUA} \tag{2.10}$$

where:

LC_{ij} = Link connectivity between FUAs i and j,
θ = Weighing factor for time and frequency numbers,
tt_{ij}^x = Travel time for activity x, between FUAs i and j,
IV = In-vehicle,
OV = Out-of-vehicle,
A = Access,
W = Waiting,
E = Egress,
T = Transfer.

2.1.5. Physical integration variables

Besides the services and their frequencies, the physical integration of air and rail travel is measured. The availability of a railway station at the airport first observes this. It is considered an on-site station when walking from the arrival hall is less than 10 minutes. Physical integration of the station and airport removes the need for a people mover between the two. Also, the status of the station is important for its position in the rail network. In this observation, a difference is made between conventional railways and high-speed railways. The station is considered an HSR station if a service departing from the station will drive over a high-speed line at some point. Therefore a high-speed track doesn't need to be accessible directly from the airport's station, but it shows that there's no need to transfer to a main railway station to use the high-speed rail network. The construction of this variable is shown in Equation 2.11.

$$SA_i = \begin{cases} 0, & \text{if no railway station is available} \\ 1, & \text{if a conventional railway station is available} \\ 2, & \text{if a high-speed railway station is available} \end{cases} \quad \forall \quad i \in S^O \tag{2.11}$$

Given that no airport in Europe is connected to all HSR connections of its FUA, the public transport travel time between the airport and the central railway station of the airport's related city is still important. This is calculated using the Google Maps Directions API, which calculates the route between two locations by a given travel mode. The data is retrieved from this API for all airports in the final selection. To reduce manual labor, the origin is set as '*ICAO-code* airport' and the destination as '*Airport's municipality* (central) station,' where 'central' is included first, but removed when no location is found. The resulting travel time represents the public transport accessibility of the airport from the railway network and vice versa.

Finally, the physical location of the airport is considered. This is represented by the number of airports that can be found within a competitive range. This geographical indicator also shows the potential of the airports' location when competitive rail connections are available. To avoid neglecting physical barriers, such as water or mountain ranges, the travel time by car determines the range. As the road network in Europe is highly evolved, it is a more representative indicator than the current rail network. The travel time by car is calculated using the Open Source Routing Machine (OSRM) API based on the coordinates of all airports in the selection. However, the determination of the significant range is based on a travel time comparison between car and train routes between different destinations in the Netherlands, Belgium, France, England, Germany, and Spain. This comparison resulted in a 0.67 multiplication factor to estimate rail travel times from car travel time. In a range between 4 and 6 hours, high-speed rail is considered competitive with air transport (Yuan, Dong, and Ou 2023). To view the negative effect of transfer times, the lower bound is chosen to assess intermodality. Therefore, the indicator represents the number of airports HSR can reach within 4 hours.

2.2. Statistical data treatment: Principal Component Analysis

Before any clustering can be performed, the data needs to be treated further to find variables that can be used for clustering purposes. Common in transport modeling is to asses a mode choice via calculating a utility for that mode choice. Utilities are calculated by the principal shown in Equation 2.12. This shows that the utility for alternative i is determined by the sum of all attributes j for alternative i multiplied by a parameter β_j that indicates the weight of the attribute j.

$$U_i = \sum \beta_j * X_{ij} \tag{2.12}$$

Parameter β is often determined by stated preference research and focused on mode choices. However, for this research's modeling purposes, the KPIs of a system are considered instead of alternative attributes. They can be combined into a single or a few performance indicators by the same principle. However, in this case, that would be a very arbitrary method due to the lack of argumentation by, e.g., stated preference research.

A PCA can be executed to avoid an arbitrary weighting of performance variables, such as the β-parameters in utility functions. PCA is a method of multivariate statistics to reduce the dimensions of all observations (Maćkiewicz and Ratajczak 1993) while preserving as much information as possible and finding the variability between, in this case, the airports. This method is use to extract new variables with significant distinction from a multi-dimensional data set. The creation of new components reduces the number of variables, which reduces computational efforts during clustering. Also, it extracts the statistically most important information from the data set, which can be used in the clustering to cluster based on the most important characteristics.

2.2.1. Methodology of a PCA

A PC is a new variable that combines the initial data. The constructed PCs aim to contain as much information as possible in the first constructed components without forming any correlation. By reducing dimensionality, visualization is easier to achieve. However, it is harder to directly interpret the results because of the combination of multiple variables. A PCA repeats the process of finding a component with maximum information until the number of PCs matches the dimension of the variable set.

The process of the PCA is structured as follows (Zakaria Jaadi 2023):

1. Standardization of continuous variables.
2. Computation of covariance matrix to identify correlations.
3. Computation of eigenvectors and eigenvalues to identify PCs.
4. Vector creation to assess PCs.
5. Project data on PCs.

The first step standardizes the data to avoid differences in data weight by differences in their units and order of magnitude. Standardization is done by dividing the difference between the attribute's value and mean by the attribute's standard deviation. After performing the standardization, all variables will be scaled similarly. The standardization is defined in Equation 2.13,

$$z_{ij} = \frac{X_{ij} - \mu_j}{\sigma_j} \tag{2.13}$$

where:

z_{ij} = Standardized value of sample i for variable j,
X_{ij} = Value of sample i for variable j,
μ_j = Mean value of variable j,
σ_j = Standard deviation of variable j.

After standardization, a covariance matrix will be constructed to identify relationships between the different variables. If high correlations occur, there is probably an overlap in information between the variables. In this way, the reduction of dimensions can be argued. The covariance matrix is used in the next step to compute its eigenvectors and eigenvalues, which are used to determine the PCs. They indicate the direction of the axes with the most variance, thus the most information, and the amount of variance explained by the PC, thus a coefficient. Their eigenvalue can then rank the PCs to indicate their significance.

There are multiple available methods to determine the optimal number of PCs. One is the graphical approach of identifying an elbow by the scree test. This scree test creates a plot based on the cumulative explained variance, calculated by dividing each eigenvalue by the sum of the eigenvalues. Also, stopping conditions of an 80% or 90% minimum explained variance are commonly used. However, this method is considered relatively arbitrary and mainly an easy-to-use method (Cangelosi and Goriely 2007). A more widely used method is the Kaiser-Guttman criterion, which uses the average of the PCs' eigenvalues as a threshold for the retention of the component. Jolliffe (2002) discussed this method on several examples and added a 70% cut-off for this threshold to avoid including too few components.

To finally apply dimensionality reduction, a Feature vector is created, which contains the eigenvectors of the PCs that are kept in the PCA. The decision of which PCs are kept is based on the total explained variance by the chosen PCs. In this research, reducing dimensionality is the main goal of the PCA, facilitating a clearer clustering. Therefore, it can be chosen to leave less significant PCs out of the Feature Vector.

The final step of the PCA links the Feature Vector to the original data by multiplying the Feature Vector by the standardized original data, as shown in Equation 2.14.

$$FinalData = FeatureVector^T * StandardizedOriginalData^T \tag{2.14}$$

A correlation circle can be plotted for two PCs to obtain insights into the structure of the found PCs. Also, a chart of the correlations between PCs and underlying variables can be constructed for two or more PCs. This can be used to evaluate the clustering step to support the choice of classification based on underlying data.

2.3. Clustering on intermodal characteristics: Agglomerative Hierarchical Clustering

The next step of the research process contains the clustering of the airports. This analysis uses the PCs determined in the previous step as variables. The AHC creates clusters from the PCs' values. AHC is an agglomerative algorithm structured as follows (Maklin 2018):

1. Start with n clusters
2. Join the two clusters with the lowest distance in between.
3. Clusters will be joined until the data set has one cluster.
4. Visualize the clustering process by a dendrogram.
5. Determine the optimal number of clusters.

2.3.1. Methodology of an AHC

In the AHC method, clusters are constructed with a bottom-up approach, starting with as many small clusters as possible and merging them step-by-step until one large cluster is created. If AHC were performed on a data set with n samples equal to the number of airports in this study, the initial number of clusters would be n. Then, based on the distances between the clusters, they are joined to form larger clusters. The distance can be calculated in different ways. Mainly, Euclidean distances are used, which are independent of the dimension of the data set and can be calculated in a multi-dimensional space, Equation 2.15,

$$d_{pq} = \sqrt{\sum_{i=1}^{N}(p_i - q_i)^2} \tag{2.15}$$

where:

d_{pq} = Distance between samples p and q,
N = Number of dimensions (PCs),
p_i,q_i = Coordinate of point p or q in dimension i.

The AHC technique, contains several options for different linkage methods. This analysis uses the *ward*-linkage, which minimizes the increase in the sum of squared Euclidean distances within each cluster when two clusters are merged. This method particularly copes well with noise between clusters, which is useful in this case due to variety in the data set and its variables, that can still be stored in the PCs. An AHC starts the linkage process with a number of clusters equal to the size of the data set. The two clusters that cause the lowest increase in cluster variance when merging, are joined to create a new larger cluster. To assess this lowest increase in variance, a matrix can be constructed, with the increase of variance for each merge of clusters. The calculation of these entries is mathematically described by Equation 2.16,

$$D(W, C_i \cup C_j) = \frac{|C_i| * |C_j|}{|C_i| + |C_j|} * d(m_i, m_j)^2 \tag{2.16}$$

where:

$D(W, C_i \cup C_j)$ = Increase in cluster variance when clusters C_i and C_j merge into cluster W,
$|C_i|$ = Number of data points in cluster C_i,
$d(m_i, m_j)$ = Euclidean distance between m_i and m_j (Equation 2.15),
m_i = Centroid of cluster C_i, calculated by the mean of all dimensions in that cluster.

From this matrix, the lowest value is picked, and these clusters are merged. Then, the entries involving the new cluster are recalculated, and the process is repeated step-by-step, merging the closest clusters until the entire data set consists of one cluster. The clustering steps are memorized and can be plotted in a dendrogram when the process is finished. An example dendrogram is shown in Figure 2.4. The dendrogram visualizes the bottom-up merging of the clusters. The length of the vertical lines represents the distance from the cluster to the new cluster. This distance is also the basis of the method, uniquely for AHC, to determine the optimal number of clusters (Dutta Baruah 2020). Horizontal lines can be drawn that show the number of clusters at that point of the clustering process by the number of vertical lines it intersects. The largest vertical distance in which no new clusters are formed indicates the optimal number of clusters. The example in Figure 2.4 has an optimal number of 4 clusters.

Another method to calculate the optimal number of clusters is the Silhouette Coefficient method, as shown in Equation 2.17 (Dutta Baruah 2020),

$$S_i = \frac{b_i - a_i}{\max\{a_i, b_i\}} \tag{2.17}$$

where:

S_i = Silhouette Coefficient for point i, s.t. $-1 \leq S_i \leq 1$,
b_i = Smallest average distance from point i to all points in any other cluster,
a_i = Average distance from point i to all other points in its cluster.

This method evaluates whether individual points are assigned to the correct cluster. The value of S_i can vary from -1 to 1, with positive values that indicate a correct assignment, negative values an incorrect assignment, and values close to 0 indicating indifference. The average of each point's Silhouette Coefficient shows the Silhouette Score for the entire data set based on Equation 2.17 with n samples in the data set. The number of clusters that result in the Silhouette Score's maximum value is considered optimal. The definition of the Silhouette Score is given in Equation 2.18.

$$S = \frac{\sum_{i=1}^{n} \frac{b_i - a_i}{\max\{a_i, b_i\}}}{n} \tag{2.18}$$

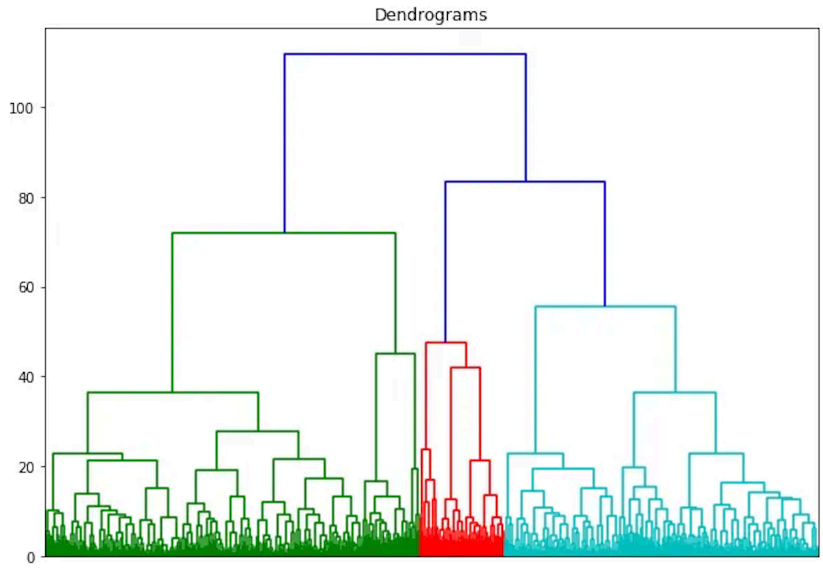


Figure 2.4: Example of dendrogram (Dutta Baruah 2020)

Both validation methods are used to check the variability in the number of clusters. When the correct number of clusters is determined, the clustering can be visualized in various ways. If two or three PCs are used in classification, a simple two- or three-dimensional plot can be used to show the distribution of clusters. When more PCs are used to cluster the data, or if it is desired to plot all original data, pairwise plots or different plotting techniques will be used. These visualizations are based on the data's values. Also, the location of the airports and their cluster can be plotted to show the network layout and the role of different airports, as shown in the example of Figure 1.4.

2.4. Implementation of policy scenarios on airport data

As the classification assessed the current situation, it will be repeated after iterations for future scenarios. The constructed scenarios aim to depict realistic possibilities for a future network. This concerns governmental considerations, network possibilities, and air and railway company behavior. These scenarios are translated to a change in the underlying variables and can be estimated based on example calculations and assumptions on the effect on the entire network. After the iteration of this data, the clustering process can be repeated with the PCA and AHC. Afterward, all results from the different clustering iterations are considered in the next classification step.

2.5. Classification of clustered airports

The final step of the network analysis is related to the labeling of the clusters that are found. The results of the AHC are translated to tables with mean values per cluster for each variable. Together with a geographic visualization of the clusters, the characteristics of each cluster can be classified. This classification is performed on characteristics such as the performance of service variables or the ratio between several variables indicating a 'balanced' or 'unbalanced' situation. Also, the geographical location of airports within the study area can affect classification or explain the differences between found clusters.

The classification is performed on all constructed scenarios. Changes in clustering or values of specific variables in these clusters ask for a reassessment of the clustering. For each scenario, it is assessed whether the clusters have the same characteristics and if the label assigned to the clusters is still applicable. It is possible that the number of clusters changes or that the structure of these clusters is different, and other variables determine the character of these clusters. After assessing these changes, the effect of the scenario implementation on the classification can be summarized and interpreted. This can help conclude if scenarios can fully restructure an intermodal network or if only the roles of individual airports in the network are affected by these changes.

2.6. Evaluating intermodal network potential

When the overall classifications within the intermodal network show changes, they can be interpreted to see the effects of the scenario on the network. If more airports are assigned to a cluster that can indicate a high performance in intermodal opportunities, the availability of intermodal travel in the network is increased. Also, a full classification change could change the view on intermodality in the network and, therefore, the potential to enable people to travel intermodal. The changes in the airport classification can be linked to the airports' passenger numbers. Calculating the total number of passengers per cluster can estimate the effects of the scenario implementation. This approach indicates the effect on the integral European intermodal network. From the changes in classification and shifts in clustering by individual airports, specific airports of interest can be identified. Depending on the scenarios and the changes that are caused, this can help to prioritize the execution of policies or to identify the airports where they can be applied ideally.

3
Classification of a benchmark scenario in Europe

With the constructed method for network analysis, a benchmark can be set based on the current network variables. This benchmark sets a base case for the constructed tool, enabling comparison with results obtained after scenario implementation. The current situation is based on European airport and railway data from 2019, just before COVID-19 interrupted the developments in the transport sector. Using this situation as the benchmark situation responds to the expectation that the transport sector will recover to its old levels and continue its growth in a similar way as before the pandemic.

In this chapter, the collected and calculated data, as determined in chapter 2, is assessed by the constructed method, subsequently performing data standardization, a PCA, and an AHC. The numerical results of this process are described in this chapter. This contains statistical information, the determination of the used PCs, and a discussion on the optimal number of clusters. This chapter aims to describe the numerical outcomes and to describe the interpretations of these results by comparing them with the outcomes of the analysis of scenarios.

3.1. Descriptive statistics of the airport data

For the selection of airports, the variables as described in Table 2.2 are calculated. In Figure 3.4, histograms show the distribution and values of the different variables. The air variables in Figure 3.2 identify the locations of the airports with a high number of air services. More services are mainly found in the western and central parts of Europe. The ratio of intercontinental flights is distributed more evenly, location-wise. The more peripheral countries have a relatively high ratio of intercontinental flights, which can be explained by their situation near the edge of Europe. Therefore, flights like Athens-Ankara, for example, are intercontinental, while they are relatively shot-ranged. Furthermore, the airports with higher numbers of air connections correlate with the total number of flights. However, the distribution of the number of destinations, as shown in Figure 3.4d, is a bit broader compared to the total flight distribution from Figure 3.4a.

The geographical distribution of the higher numbers of rail services, as shown in Figure 3.3, indicates a strong correlation with geographical centrality. Germany and the UK especially offer a high frequency of rail services. More peripheral countries, like the Scandinavian, Iberian, and Eastern countries, show a very small number of rail connections, explainable by the limited directions in which a connection to the rail network can be made. The travel time between airports and the main railway network, captured by the variable for station accessibility, is distributed very evenly geographically. The outliers with a high travel time are indicated by small markers in Figure 3.3c, geographically distributed in a fairly random manner. The availability of the different types of railway stations at the airport follows the same main corridor as the total rail connections, as shown in Figure 3.3d. However, the UK has not transitioned to HSR services at its airports. Also, multiple peripheral cities lack an on-site connection to the railway network. The same correlation is shown in the average shortest path for each airport, as seen in Figure 3.3b.

To perform the clustering and classification in a later phase, the data, as described in section 2.1, is standardized to scaled values, so all variables have a mean $\mu = 0$ and a standard deviation $\sigma = 1$. The distribution of all scaled variables is visualized and can be compared by the boxplot in Figure 3.1. This figure shows the differences in the variables' structures. *TF air*, *IF air*, *NC air*, and *SAC* show a narrow variability between the 1st and 3rd quartile, with a broad range of outliers. Also, almost all medians have a value lower than the mean of 0, which shows that overall, there's a large share of lower numbers on each variable, with all outliers being higher values. In the clustering and classification phase, this could indicate that the outliers can distinguish clusters. The distribution in the data also implies the expectation that clusters will be structured from specific combinations of variable values. Therefore, each variable is of significant value to assess during the classification process.

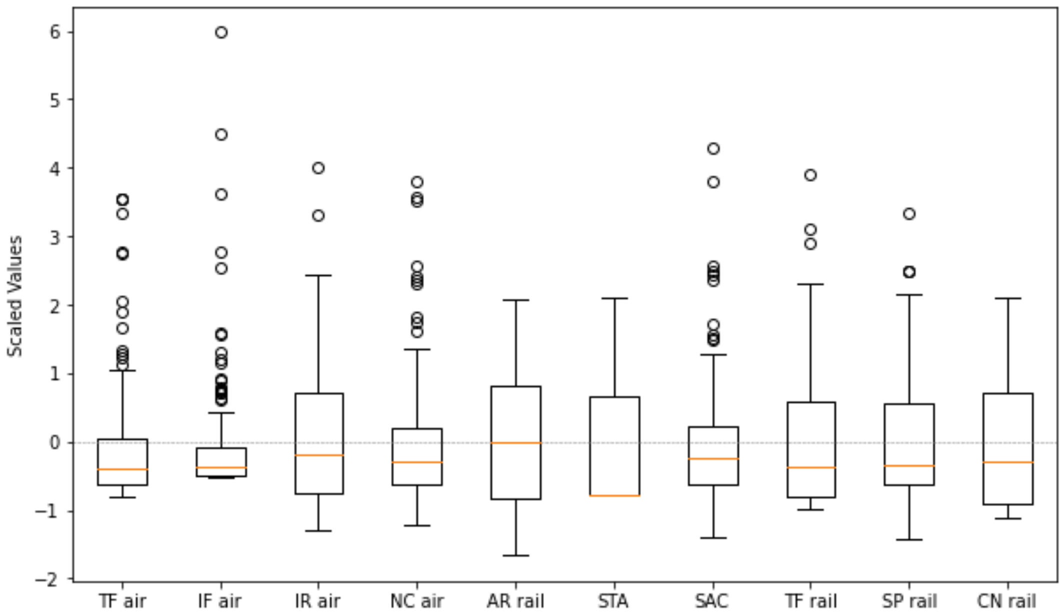
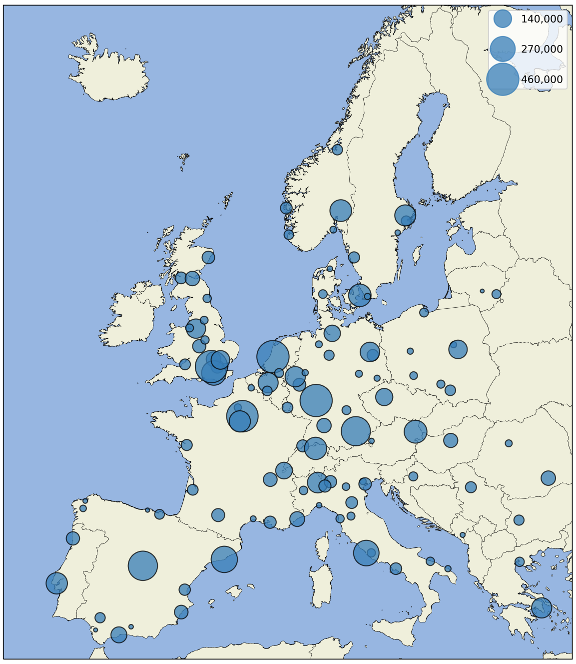
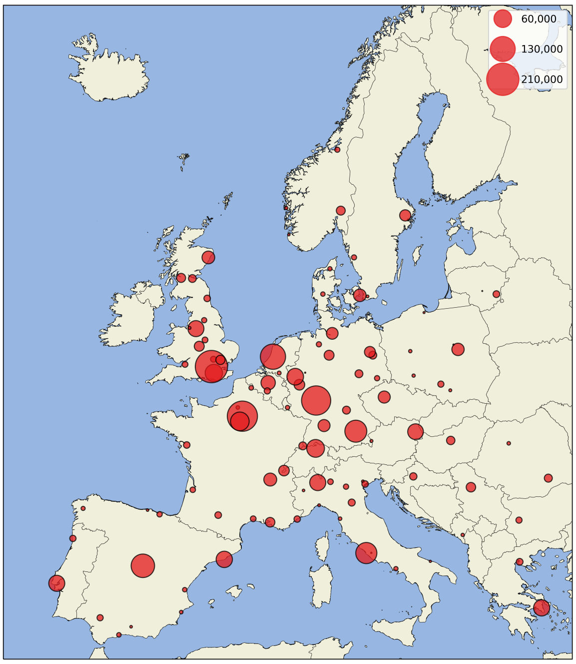


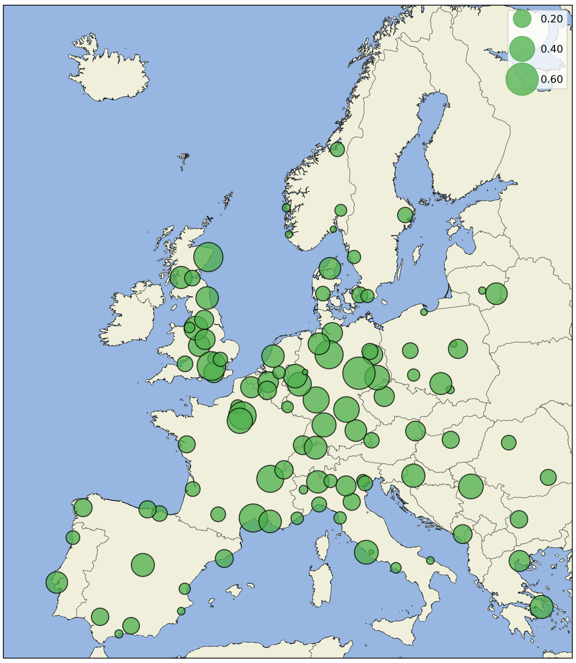
Figure 3.1: Boxplot describing the scaled data



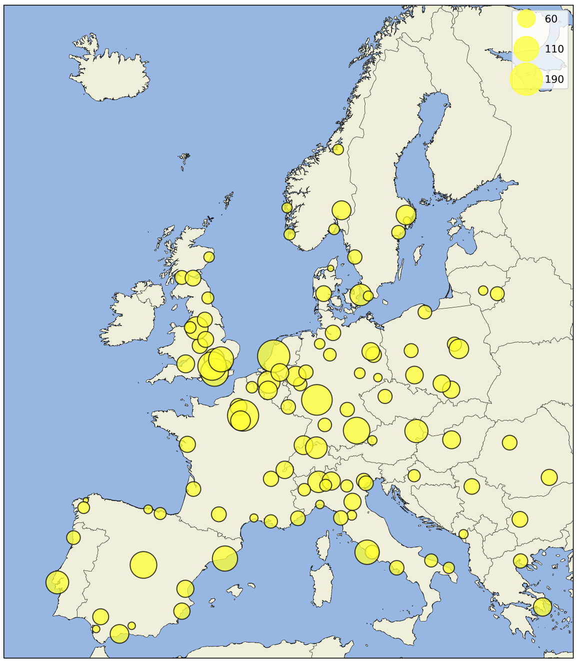
(a) Total annual air frequency



(b) Annual air frequency for intercontinental flights

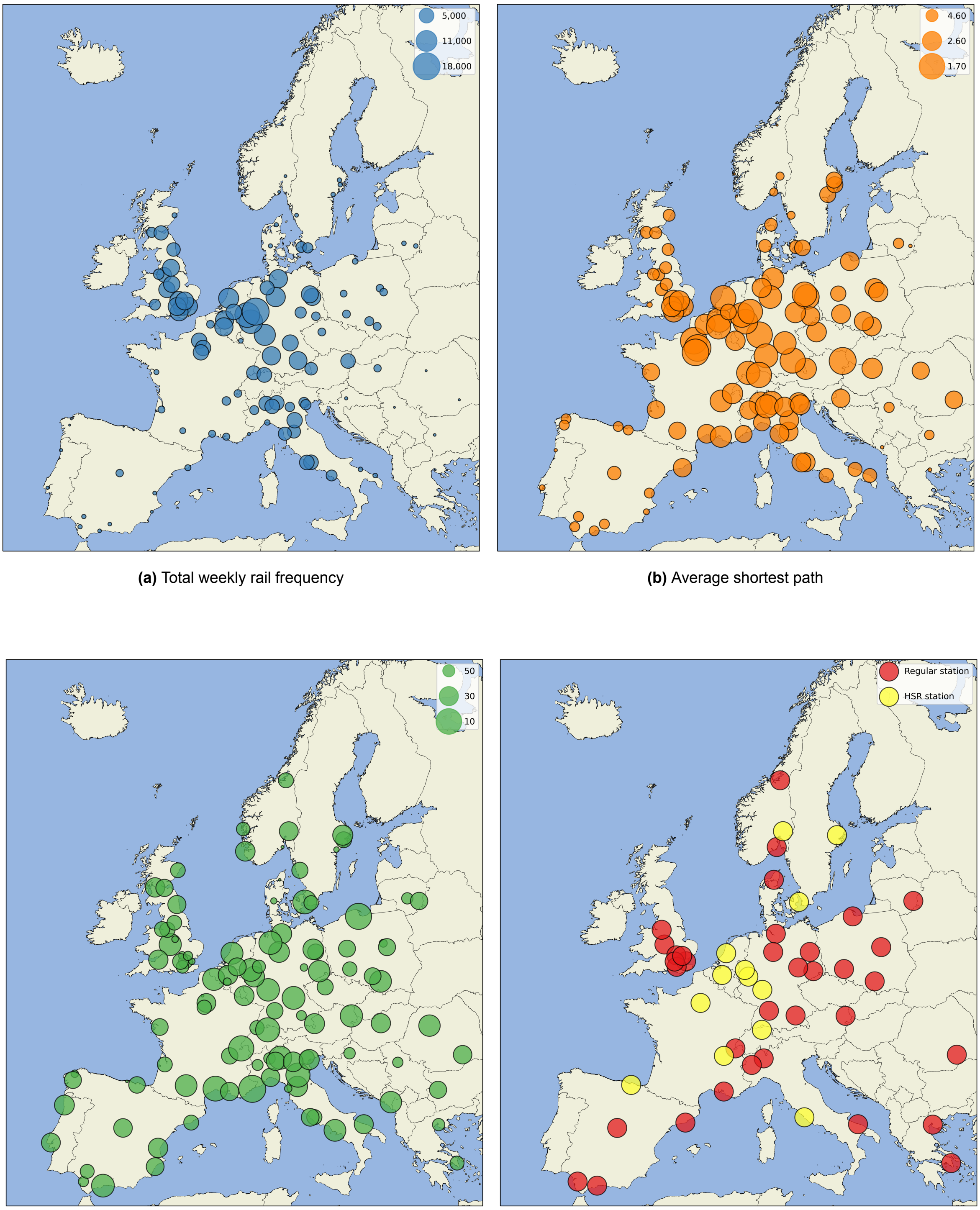


(c) Ratio of intercontinental flights



(d) Number of air connections for airport locations

Figure 3.2: Air services for airport locations



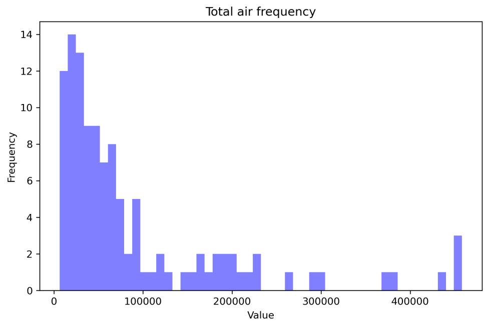
(a) Total weekly rail frequency

(b) Average shortest path

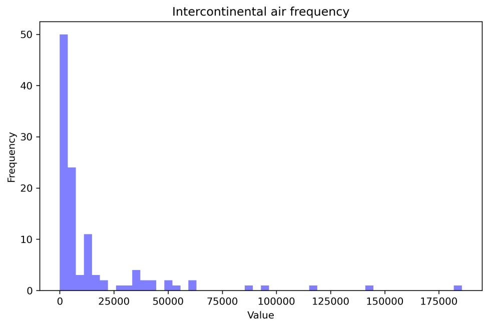
(c) Station accessibility by access time

(d) Station type availability

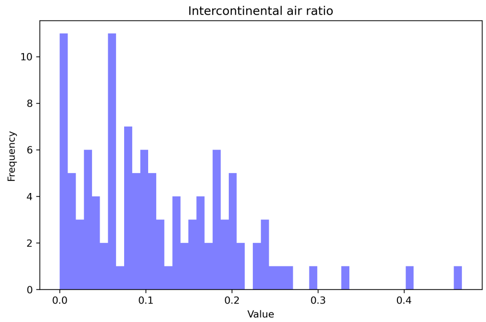
Figure 3.3: Rail services for airport locations



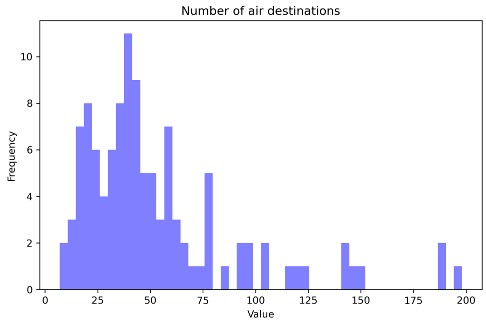
(a) Total air frequency



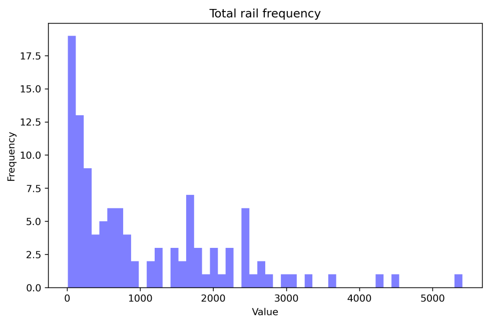
(b) Intercontinental air frequency



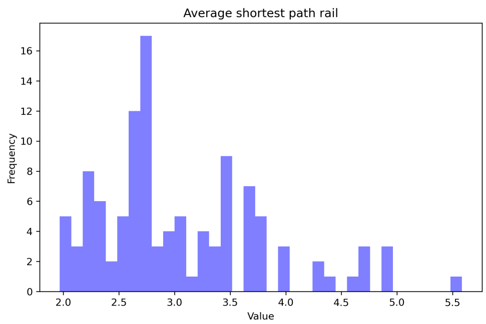
(c) Intercontinental air ratio



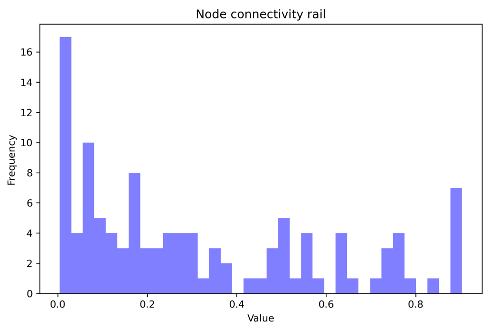
(d) Number of air destinations



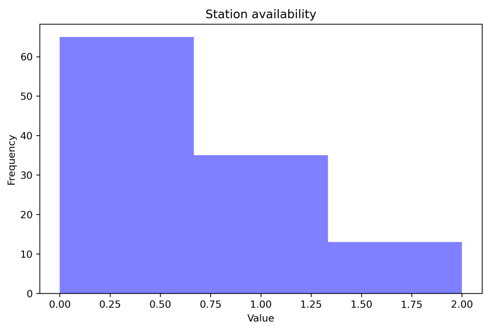
(e) Total rail frequency



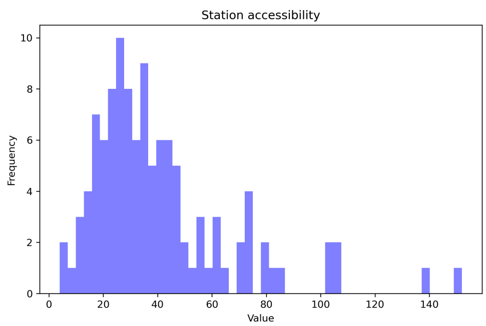
(f) Average shortest path rail



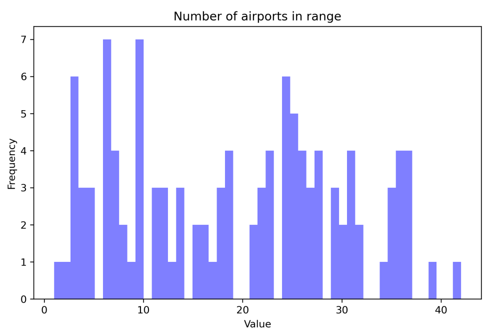
(g) Node connectivity rail



(h) Station availability



(i) Station accessibility



(j) Number of airports in range

Figure 3.4: Descriptive statistics of all variables

3.2. Reduction of dimensionality in PCs

To reduce the dimensionality of the data set, a PCA is performed. A PC is a unit vector fitting the data best while orthogonal to the previously constructed PCs. Table 3.1 indicates the Load Factors for each PC, which indicate what variables correlate with the specific PC. The value is a covariance that lies between -1 and 1. The higher the absolute value, the higher the correlation between the variable and the PC. In Table 3.1 the values with a relatively high correlation ($|X| > \mu_{PC} + \sigma_{PC}$) are marked green, while relatively low correlations ($|X| < \mu_{PC} - \sigma_{PC}$) are marked red. This shows an average correlation on the first PC for almost all variables. The further along the table, the PCs become more specific and especially load on one or two variables. The first PC shows a factor loading, which is fairly evenly correlating with almost all variables. Only station accessibility (SAC) has a significantly lower correlation with the first component. The second component is loading more on the number of airports in a potential HSR range (AR rail). Also, a negative correlation with the average shortest path (SP rail) is found in this PC. Considering a medium correlation with the other rail-related variables, this component seems to be correlated with the geographical location of the airports, as supported by the conclusions from Figure 3.3. The third, and final, included PC highly correlates on station accessibility, showing that including more PCs would add more specific characteristics. The choice for using three PCs is explained in further detail using several assessment techniques.

Table 3.1: Load Factors for each PC

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail
PC 1	0,37	0,37	0,29	0,38	0,30	0,26	**-0,12**	0,35	-0,29	0,35
PC 2	-0,34	-0,33	**-0,22**	-0,29	**0,44**	**-0,23**	**0,07**	0,34	**-0,38**	0,36
PC 3	**0,05**	**0,14**	**0,00**	**0,07**	**-0,13**	**-0,04**	**0,96**	**0,08**	**0,04**	**0,14**
PC 4	**-0,07**	**0,20**	**0,63**	**-0,05**	**0,11**	**-0,73**	**-0,03**	**-0,11**	**0,04**	**-0,01**
PC 5	-0,30	**-0,14**	**0,61**	-0,35	**0,15**	**0,58**	**0,13**	**-0,05**	**-0,03**	**-0,14**
PC 6	**-0,08**	**-0,06**	**0,10**	**-0,11**	**-0,16**	**0,02**	**-0,13**	**0,51**	**0,73**	0,36
PC 7	**-0,02**	**-0,08**	**0,17**	**-0,12**	**-0,78**	**-0,05**	**-0,10**	0,30	**-0,48**	**0,06**
PC 8	**0,04**	**0,70**	**-0,22**	**-0,67**	**-0,01**	**0,05**	**-0,06**	**-0,05**	**-0,03**	**0,09**
PC 9	**0,08**	**0,09**	**-0,05**	**-0,04**	**0,15**	**-0,08**	**0,06**	**0,62**	**0,00**	**-0,75**
PC 10	**0,79**	**-0,41**	**0,09**	**-0,42**	**0,07**	**-0,06**	**0,05**	**-0,05**	**0,02**	**0,03**

For each row, the Factor Loadings with $|X| > \mu + \sigma$ and $|X| < \mu - \sigma$ are marked.

Each PC has a specific eigenvalue that indicates the amount of information it contains. This is used in the determination of the optimal number of components. Multiple methods to obtain this optimal number are explained in section 2.2. The more arbitrary 'explained variance method' is shown in Figure 3.5. The benchmark of a cumulated explained variance of 80% or 90% states that 2 to 4 PCs are considered optimal within this data set.

The Kaiser-Guttmann criterion is used as another perspective for the optimum determination. This method states that PCs with a higher eigenvalue than average should be included. As the eigenvalues are based on the covariances of the variables, the average of all eigenvalues equals 1. The Jolliffe modification to this method uses a reduction to 70% of this threshold to avoid underestimation. However, as mentioned earlier, the literature also warns of overestimation by the Jolliffe modification. In Table 3.2 it is shown that two PCs would be chosen for the Kaiser-Guttmann method, and the Jolliffe modification results in four PCs.

Table 3.2: Eigenvalues of constructed principle components

k^{th} **PC**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Eigenvalue	4.83	2.12	0.98	0.74	0.54	0.44	0.19	0.13	0.09	0.04

Finally, a scree test can be performed to find an optimum by identifying the elbow in a plot of all individual variables, as shown in Figure 3.6. This plot indicates that the line flattens out just after the third PC. All verification methods show similar results in this case. Finally, three PCs are used in the further

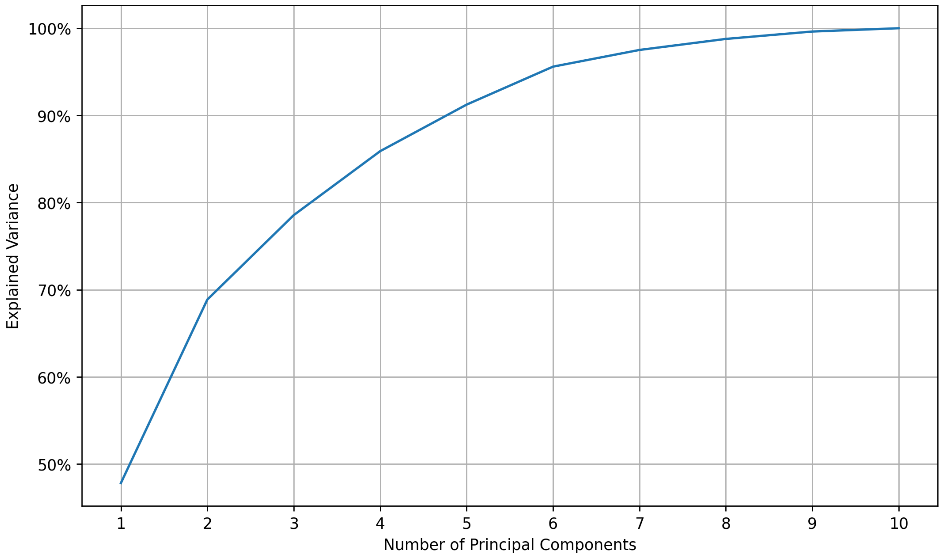


Figure 3.5: Cumulative Explained Variance per principle component

analysis of the data.

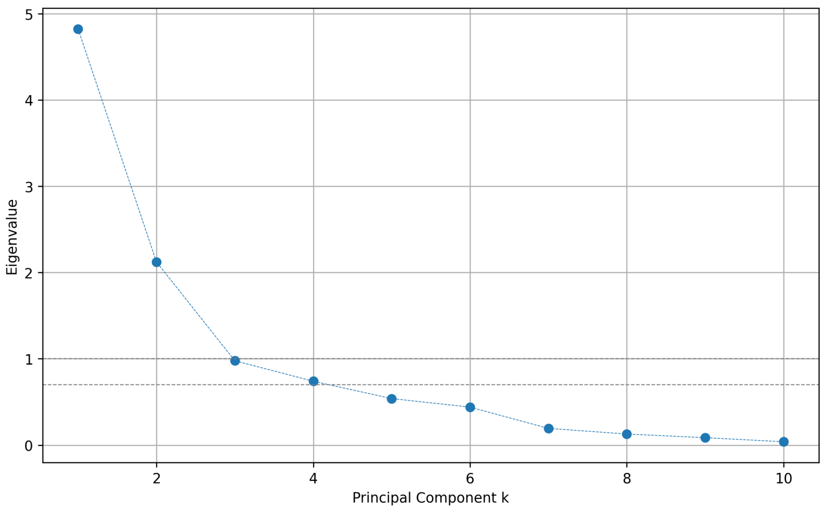


Figure 3.6: Principle components' eigenvalues

3.3. AHC results: benchmark scenario

Based on the dimensionality reduction to three PCS, the AHC is performed by step-by-step merging the closest data points based on a generalized distance matrix. The dendrogram in Figure 3.7 shows this process, with the bottom being the starting point where all data points form their own cluster. The process continues until all data points merge into one large cluster. The optimal number of clusters can be estimated visually by looking at the dendrogram or determined more exactly by calculating the Silhouette Score.

The first method is based on the similarities between the clusters. The larger the vertical distance before new clusters are merged, the smaller the similarity. Therefore, the optimal number of clusters can be found where the vertical distance between two merges is the largest. The number of clusters would equal the number of vertical lines crossed by a horizontal line at that point. In Figure 3.7, this point can be found at 5 or 7 clusters. The colors correspond to the colors assigned to the clusters in Figure 3.9.

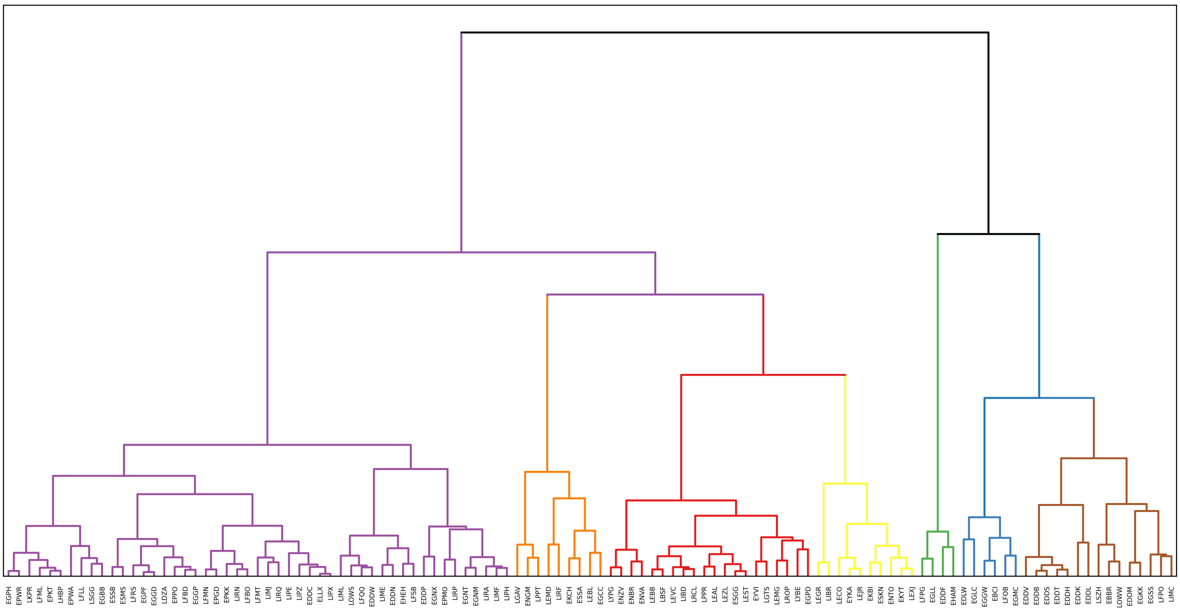


Figure 3.7: Dendrogram of the AHC on the treated data, using colors to indicate clustering for 3, 5 and 7 clusters.

The Silhouette Score is a more exact approach, assessing the correctness of each clustering assignment based on the average distance to the cluster it is assigned to and the other clusters. These scores are calculated for 2 to 10 clusters and are listed in Table 3.3 and visualized in Figure 3.8. Theoretically, the clustering shows the most correct assignment for two or three clusters. However, it is debatable whether this gives the desired insights when there are few classification possibilities. The Silhouette Score shows a small drop when four clusters are set but shows a steady increment for five to seven clusters.

Table 3.3: Silhouette Score for different numbers of clusters in the AHC

No. of clusters	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Silhouette Score	0.413	0.377	0.304	0.339	0.337	0.340	0.291	0.285	0.284

The combination of conclusions from the silhouette score and the dendrogram shows the possibility of choosing 5 or 7 clusters. Both choices are calculated, and the mean values of their variables are displayed per cluster in Table 3.4, 3.5, and 3.6. The variable names match the abbreviations in Table 2.2. As seen in Figure 3.7, the difference between the 5- and 7-cluster results is made by merging clusters 3a and 3b and clusters 5a and 5b. By the nature of AHC, the other clusters remain the same. Also, leading from the silhouette scores, a 3-cluster AHC is included to assess the consequences of a further reduction of clusters.

3.3.1. Numerical Results of the AHC

All tables show the average values for each variable. The best and worst values are marked to consider the different meanings of a high or low value. For most variables, a higher value means a better contribution to the offer of intermodal possibilities. For example, a higher frequency of air services would contribute to a better assessment of the intermodal offers' performance. However, for the travel time to the main railway station (Station Accessibility, SAC) and the average shortest path (SP rail), a lower value indicates a higher contribution to network intermodality. Therefore, Tables 3.4, 3.5, and 3.6 mark the best values in green, the second-best in blue, and the worst in red.

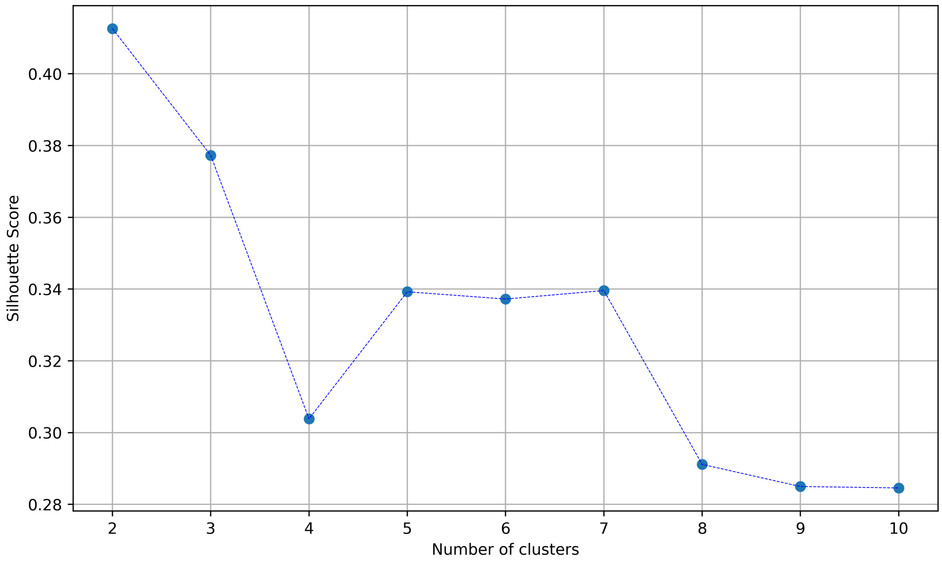


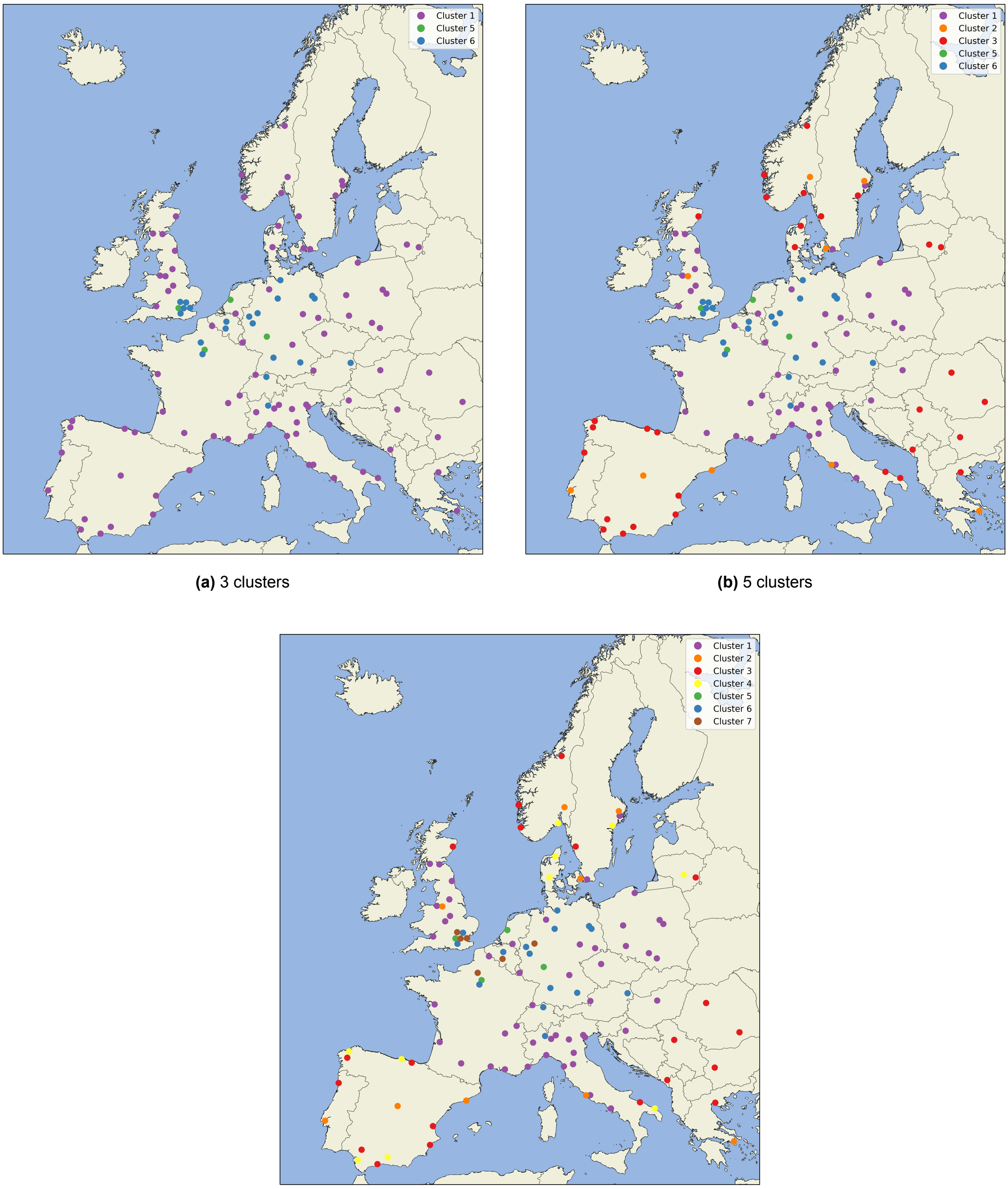
Figure 3.8: Silhouette Score for each number of cluster

For all number of clusters that are used in the analyses, a clear distinction can be found between the clusters. The more specific 7-cluster analysis shows several combinations of high- and low-scoring variables. A logical correlation between the air variables can be identified, and also for the rail variables this can be identified. The meaning of these differences is discussed in the classification of the clusters.

3.3.2. Visualization of the clustering

The cluster corresponding to the Tables 3.4, 3.5, and 3.6 are plotted on the maps in Figure 3.9 to show the differences in the choice of cluster numbers. The difference between the 3- and 5-cluster results shows that three clusters fail to divide airports that score low on the rail variables. The five clusters show that within that cluster, a big difference can be seen in the other, especially air-side, variables, as clusters 2 and 3 show the second-best and the worst values. The maps of the AHC also show a clear geographical influence on the clustering. The step from the 3- to 5-cluster analysis distinguishes centrally located airports from airports situated on the edges of the study area. This division is mainly caused by the low number of airports within an HSR range (AR), as more centrally located airports have more connections to the European mainland airports.

The extra distinction found by adding two more clusters, from five to seven, is less valuable. This division is mainly based on the details in the values of the variables and divides the geographically separated clusters into sub-categories. The clusters 3 and 4 and clusters 6 and 7 score similarly on the rail-side variables. The air-side variables do differ. These variables are still low for clusters 3 and 4 compared to the other clusters. A larger difference between the two sub-clusters lies in the accessibility of the airport. For clusters 6 and 7, a significant difference is found in the air-side variables and the airport accessibility. As both clusters show a distinctive difference in specific variables, it is considered to be useful to assess the 7-cluster results.



(a) 3 clusters

(b) 5 clusters

(c) 7 clusters

Figure 3.9: Maps with different numbers of clusters

Table 3.4: Mean variable values for 3 clusters

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail	n
1	**64.606**	**8.197**	**0,10**	**43**	**16**	**0,40**	**40**	**619**	**3,28**	**0,20**	88
5	**452.354**	**135.437**	**0,30**	**181**	**32**	**1,75**	**26**	**2852**	**2,20**	**0,79**	4
6	**133.253**	**21.538**	**0,14**	**70**	**30**	**0,90**	**45**	**2719**	**2,28**	**0,72**	21

Table 3.5: Mean variable values for 5 clusters

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail	n
1	48.456	4.913	0,09	40	23	**0,24**	33	925	2,86	0,30	49
2	**237.474**	**39.747**	**0,16**	**100**	10	**1,33**	**31**	617	3,61	0,16	9
3	**39.125**	**4.095**	**0,08**	**32**	**7**	0,37	**55**	**120**	**3,86**	**0,04**	30
5	**452.354**	**135.437**	**0,30**	**181**	**32**	**1,75**	**26**	**2852**	**2,20**	**0,79**	4
6	133.253	21.538	0,14	70	**30**	0,90	45	**2719**	**2,28**	**0,72**	21

Table 3.6: Mean variable values for 7 clusters

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail	n
1	48.456	4.913	0,09	40	23	**0,24**	33	925	2,86	0,30	49
2	**237.474**	**39.747**	0,16	**100**	10	**1,33**	**31**	617	3,61	0,16	9
3	51.800	5.908	0,11	38	**7**	0,40	34	**102**	**4,02**	**0,04**	20
4	**13.775**	**469**	**0,03**	**19**	**7**	0,30	**96**	157	3,56	0,06	10
5	**452.354**	**135.437**	**0,30**	**181**	**32**	**1,75**	**26**	**2852**	**2,20**	**0,79**	4
6	43.983	2.624	0,06	46	**30**	0,33	71	**2827**	2,36	**0,81**	6
7	168.961	29.103	**0,17**	80	**30**	1,13	34	2676	**2,25**	0,69	15

*For each variable the **best value**, **2nd best value**, and **worst value** are marked.*

3.4. Classification of the benchmark scenario

The clusters are classified to assess the meaning of the different clusters found and to see the effect of the changed clusters on the role of some airports. This classification adds a label to each cluster, showing the clusters' qualitative characteristics and assessing each cluster's added value more qualitatively compared to the quantitative analysis.

The variable values in Table 3.6 roughly show four categories of clusters, combined by high or low performance on air and rail services. Other variables, such as the station availability or the accessibility of the rail network, subsequently split these categories further. The colored cells in the table indicate the best (green), second-best (blue), and worst (red) scoring variables regarding their contribution to intermodality.

The constant performance of the core four airports is notable. Cluster 5 scores the highest in (almost) all variables and is built up from the same four airports for all numbers of clusters that are analyzed. The already noted airports beforehand, London Heathrow, Paris Charles de Gaulle, Amsterdam Schiphol Airport, and Frankfurt Airport, are clustered with high-scoring values on all aspects. These core airports are labeled as *Main Intermodal Hubs*, highly performing on their rail and air services, with a good station availability and connection to the main railway network. Also, these airports focus relatively on intercontinental flights, forming a good example of serving transfer passengers who could access the airport by rail. Especially considering total and intercontinental flights numbers, this cluster can be identified from the descriptive statistics in Figure 3.2.

Clusters 2 and 7 also perform fairly well regarding the offered frequency of air services. They are distinguished by their offered services on the rail side. Cluster 2 has an average to low number of rail services in the FUA it is linked with. On the other hand, Cluster 2 does perform well on the availability of a railway station and its accessibility to the main railway network. Some examples of airports included in this cluster are Lisboa, Rome, Athens, and Oslo. These airports are relatively large airports, with available infrastructure for rail services. However, due to their location on the edge of Europe, they lack centrality, thus connection to the European rail network. Therefore, it can be classified as *Peripheral International Hubs*. Cluster 7 contains airports with similar air services but a higher rail transport service. The higher node connectivity index in the railway network confirms the more central location of the airports. Containing airports such as Berlin, München, Brussels Zaventem, and London Gatwick, this cluster is the central counterpart of Cluster 2 and, therefore, can be classified as *Central International Hubs*. The division in Peripheral and Central hubs, is mainly driven by the total frequency in rail travel. In Figure 3.3 this geographic distribution is identified as well.

The remaining clusters consist of airports with a lower level of air services. Cluster 6 is a cluster of airports that also have a relatively low number of annual flights. However, they are situated near FUAs with a good position within the rail network. The accessibility of these smaller airports is fairly low, as already identified in Figure 3.3c, with high travel times to the main railway network and limited availability of on-site railway stations. The cluster contains secondary airports such as London Luton, London City, Southend (near London), Paris Beauvais, Brussels Charleroi, and Dortmund, of which the latter is linked to the Ruhr area in Germany. These airports can be classified as *Central Secondary Airports*.

Cluster 3 and Cluster 4 are characterized by an extremely low level of rail services. Also, the number of airports within a potential substitution range is very low for both clusters. Their relation with the geographic situation, as shown in Figure 3.3, explains this low number, with all airports situated on the edges of Europe. The eastern airports in these clusters, such as Vilnius, Sofia, and Thessaloniki, could be placed in this cluster by the boundaries of this study. Rail connections to the Asian continent are not considered and could have influenced the clustering outcome. However, due to the really low number of direct connections between Europe and Asia, the effect is expected to be low. Nevertheless, the number of flights these airports serve is also very low. The distinctive factor between these two clusters lies in the station accessibility, i.e., the connection to the railway network. Cluster 4's travel time to the studied railway network is extremely large. Again, this could be affected by modeling choices because the airports are linked to an FUA that is situated relatively far away. Still, their FUAs were unavailable in the obtained data set because the areas are relatively small and play a smaller role in the European railway network. Therefore, the airports from Cluster 4 are classified as *Disconnected Airports*, as they have low services in every aspect of this study. Cluster 3 contains small airports on the edge of the European network. They can be classified as *Peripheral Minor Airports*.

Lastly, Cluster 1 is the largest calculated cluster. This cluster contains airports with low levels of services but a very central location in Europe. The size of this cluster can be explained by the distribution of air services from Figure 3.4a. Also, with the low availability of on-site railway stations, as pictured in Figure 3.3d, and low air and rail services, these airports can be classified as *Central Minor Airports*.

An overview of the classification and the airports contained in each class is shown in Table 3.7. The airports are indicated by their IATA codes. All airport information can be found in the table of the airport selection in Appendix A. The classification is also visualized in the map in Figure 3.10 to see the geographical distribution of each class.

Table 3.7: Classifications of the benchmark network

C#	Class	n	Airports
C5	**Main Intermodal Hubs**	4	*CDG, FRA, AMS, LHR*
C7	**Central International Hubs**	15	*VIE, BRU, ORY, SXF, HAM, CGN, DUS, MUC, STR, TXL, HAJ, MXP, ZRH, LGW, STN*
C2	**Peripheral International Hubs**	9	*CPH, ATH, FCO, OSL, LIS, BCN, MAD, ARN, MAN*
C6	**Central Secondary Airports**	6	*BVA, LCY, SEN, CRL, DTM, LTN*
C1	**Central Minor Airports**	49	*Remaining airports*
C3	**Peripheral Minor Airports**	20	*BIO, VLC, SVQ, CLJ, SOF, SKG, BRI, VNO, TGD, BGO, TRD, SVG, OPO, OTP, BEG, ALC, AGP, SCQ, GOT, ABZ*
C4	**Disconnected Airports**	10	*LCG, GRX, XRY, SDR, BDS, KUN, TRF, NYO, BLL, AAL*

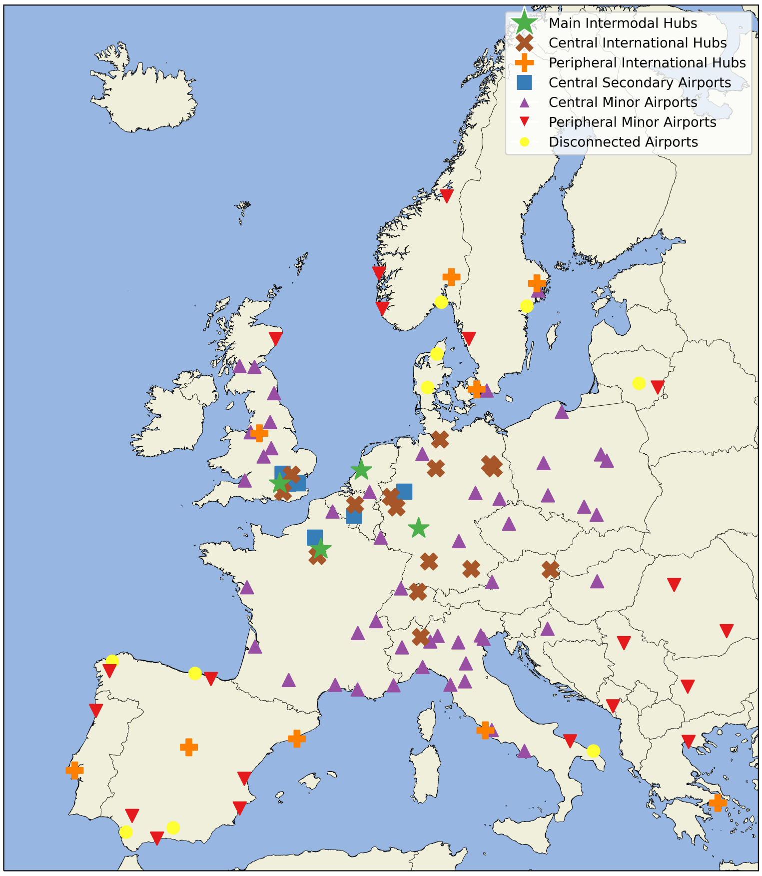


Figure 3.10: Benchmark network in seven classes

4 Air/rail policy effects on classification

The assessment of the intermodal network using 2019 data sets a benchmark for further analysis. The clusters and classification found in chapter 3 show a distinction caused by unique combinations of variable values. To identify the potential of this intermodal network, it is interesting to assess the effects that possible policies can have on the airports and their role in the network. Ideally, this study can assess realistic plans of national and international governments. However, national governments do not always share interests, and European policies depend on the implementation of national governments, which complicates the execution of such plans. To construct scenarios with a balance between having a significant impact, being affordable, and being a possible direction of decisions that will be made, the position from IenW is combined with visions for European international travel. These scenarios are translated to a change in the data used for the benchmark situation. The adapted data can be used in the constructed analysis tool to construct new clustering for each scenario and assess the contribution to the network intermodality for each cluster and its airports by the mean values of the stated variables of intermodality.

4.1. Expected policies on air/rail intermodality

Integrating air and rail products is a popular topic in government policies. However, as discussed in section 1.5, the influence of governments on specific flight and rail paths is limited. Concessions and independent organizations mainly determine these. The air and rail companies can be influenced in their choice of services by governments using tax policies or other incentives.

Because of the indirect influence on executed transport policies, it is difficult to translate elaborate transport visions to complete execution of the plans. To explore the extent of possibilities in implementing policies, interviews with a division of IenW are conducted (Appendix B). This division is responsible for the facilitation of air/rail products. The interviews show the trade-offs between larger-scale visions, such as the TEN-T, and feasible solutions within the foreseeable future. Dividing the interviews into the two subjects of air and rail travel provides a basis for two basis scenarios: one focused on rail policies, while the other focused on policies considering air. As the study scoped to a network perspective on intermodality, the scenarios do not consider integrated ticketing or baggage handling services.

4.1.1. Focus areas of government in air/rail travel

As discussed, (inter)national governments have to deal with different views on transport problems from other governments, air and rail operators, and public opinions. Therefore, not all visions of international travel can be executed as planned and consider different perspectives. Several visions in advisory documents are created considering the topic of air/rail integration or substitution. The *Actieagenda Trein en Luchtvaart* (Action Agenda for Rail and Aviation) Ministerie van Infrastructuur en Waterstaat et al. 2020 analyses the situation in a collaboration of IenW, and the Dutch air and railway operators. This results in a list of actions the organizations intend to take to improve air and rail travel and their integration.

They constructed the joint goal to 'further improve international rail travel as an attractive alternative for air travel, focused on six destinations: Brussels, Paris, London, Düsseldorf, Frankfurt and Berlin.' This agenda points out actions such as decreasing air frequencies on specific OD pairs, improving train connections, and improving the performance of high-speed infrastructure. Also, integrated air/rail products are part of the vision to improve the integration. In the context of the scope of this study, the research in changing destination stations closer to airports is an interesting vision for scenarios. However, with the used data and variables in the study and the large scope of assessed airports, it is difficult to implement such a specific measure in relevant airports.

The cooperation between the government and air and rail operators shows the importance of collaboration. The societal vision mostly originated from political organizations operating in the context of environmental regulation and economic growth. The operators that have to execute the plans are also focused on the commercial aspects of their company. Also, in some countries, such as the Netherlands, the railway operators depend on winning concessions to be allowed to operate on each corridor of the railway network. Therefore, they have to balance their economic interests and the satisfaction of the government's wishes. The cooperation between instances helps align the plans of all organizations and increases the chance of a vision becoming a reality. Moreover, the European Union aims to increase competition between operators on the international railway market. This increases insecurities in the future for railway operators, which raises the importance of collaboration for railway operators.

On the rail side, there are large infrastructural visions that should solve the technical differences in railway systems and allow operators to facilitate a successful international train service. TEN-T is a planned network for land transport to connect the members of the European Union (European Parliament 2021). This network's goal is to improve the intermodal possibilities over long distances. Extensive infrastructural plans like this are costly and time-consuming to achieve the final goal. Therefore, optimizing the use of current infrastructure is a topic to improve the rail network service. This can be achieved by focusing on rolling stock and minimal alignment of the international safety systems and other technicalities. This policy prevents the need for all tracks and uses the existing corridors.

Governmental institutions often have to deal with environmental regulations on the air side of the intermodality. Air pollution and noise pollution in areas that become even more densely populated cause the desire to reduce the number of flights in airports while the demand for air travel keeps increasing. Governments are also focusing on integrated air/rail products that offer integrated ticket or baggage handling to deal with these difficulties. Limiting the number of flights is complicated by the balance of increasing mobility, economic interests, and environmental limitations. The ban on short-haul flights, as imposed by France, also has questionable effects because of the limited influence national governments have on slot allocation. Domestically, this decision can be made; however, to arrange such policies internationally, the European Union should integrally decide to take this step. Not all members will be happy with similar policies. Island states, such as Malta, or more remote countries like Portugal, depend heavily on aviation and, therefore, have a different policy perspective than more centrally located countries. Although it is questionable whether such policies will be implemented in the foreseeable future, it is interesting to see their effects on the intermodality in Europe.

The governmental visions on air/rail intermodality are applied to two scenarios. The first focuses on extending rail services, with minimal intervention in the railway infrastructure, by performing small improvements and maximizing the capacity of the current infrastructure. The second scenario investigates the impact of improving short-haul flight substitution. When this becomes a more feasible option in Europe, countries could become more positive about banning short-haul flights, and a reduction of medium-haul flights can be achieved. A third scenario combines both scenarios to see what the combination of policies can contribute to the intermodality in Europe's network.

4.2. Scenario 1: Expanding rail services

The Dutch railway manager ProRail composed a report, *Integrale Mobiliteitsanalyse* (IMA), on the Dutch railway system (ProRail 2021. This report investigates the potential mobility for all land-based travel modes, with a sub-report on rail mobility. Infrastructural improvements are expensive, as discussed with IenW (Appendix B), and the IMA shows that not all rail infrastructure is used to the fullest. Table 4.1 shows the potential daily timetable for border crossing services from the Netherlands. In this table, the regional trains (R) are included to show the total capacity of the border crossings. In further analysis, these services are neglected because they would most likely represent local OD traffic and are not feasible for intermodal travel in most cases. The InterCity (IC) and International (INT) services represent medium- and long-distance trains.

These improvements in international timetabling are based on minor infrastructural investments, such as safety systems or energy supply. No new railway tracks are needed to facilitate the number of services presented in Table 4.1. However, the available rolling stock is currently a limiting factor in offering this number of services. During the COVID-19 pandemic, the international railway market had to deal with insecurities considering the future demand, which caused a lower demand for rolling stock. Therefore, a new impulse in rolling stock availability should be given to facilitate this increment of services.

Table 4.1: Timetable potential 2030-2050 in the Netherlands using current infrastructure, adapted from IMA (ProRail 2021).

Border	**Daily timetable 2018**	**Daily timetable 2030-2050**	**IC improvements**
Bad Nieuweschans – Weener	18x R	18x R, 9x IC	+ 9 IC
Coevorden – Emlichheim	-	18x R	-
Oldenzaal – Bad Bentheim	16x R, 7x IC	18x R, 18x IC	+ 11 IC
Enschede – Gronau	33x R	33x R	
Zevenaar – Emmerich	18x R, 7x INT	18x R, 9x IC, 14x INT	+ 9 IC; + 14 INT
Venlo – Kaldenkirchen	18x R	18x R, 18x IC	+ 18 IC
Eygelshoven – Herzogenrath	18x R	36x R	
Eijsden – Visé	18x R	36x R	
Breda – Noorderkempen	16x IC, 15x INT	36x IC, 32x INT	+ 20 IC; + 17 INT
Roosendaal – Essen	17x R	18x R, 18x IC	+ 18 IC
IC / INT division	23x IC; 22x INT	108x IC; 46x INT	+ 85 IC (4x); + 24 INT (2x)
Total IC services	45	154	+ 109 (3x)

R: Regional train | IC: Dutch InterCity Train | INT: International Focused Long-Distance Trains

The increase in international services shown in Table 4.1 is an example from the Dutch railway manager. To verify the scalability of the assumptions made from this example, documents from different countries help interpret the effects throughout Europe. In Germany, cooperation between rail providers investigates the possibilities in the German network: Deutschlandtakt (Deutschlandtakt 2023). The cooperation aims to serve an expected doubling in rail travelers by 2030 (Railtech 2020) by structuring the network's timetables throughout the country. According to the classification of the benchmark scenario in chapter 3, this network mostly represents the central airports in Europe, with a high number of rail services. In Poland, the new airport near Warsaw (CPK) expects a tripling of rail travelers when the airport is finished (Centralny Port Komunikacyjny 2023). This is based on increased rail demand and improvements in air/rail integration.

More remote areas, such as Romania, Norway, and Spain, all have a focus on the improvement of rail services as well. In Romania, plans exist to implement the first high-speed rail infrastructure (Railtech 2022) to cope with the expected doubling of rail demand. In the meantime, Norway, a country with a low focus on rail travel, also discusses plans to increase the supply of rail travel (Jernbanedirektoratet 2020). Due to Norway's geographical characteristics, most connections between cities are mainly served by air travel. Oslo-Bergen, for example, serves 26 daily flights versus 4 daily trains. This also causes a modal split of 60%/25%/15% for air, car, and rail travel, respectively, highly favoring aviation. Still, Norway's vision in 2050 is to improve rail travel, considering a doubling in long-distance rail travel. Besides, Norway aims for a stronger international connection with Sweden and the rest of Europe. Furthermore, Spain is known for their strong railway services. However, the peripheral situation shows fewer rail services in the clustering than the more centrally located airports. The future vision of Spain also considers a doubling of rail travelers in 2030 (Equipo Barcelona 2021). The example of Barcelona also states the weak link between Catalunya and Perpignan and between the Basque Country and Bordeaux.

The examples of different countries all consider a doubling of rail passengers, mostly in 2030. The expectation of a tripling of the passenger market volume by 2050 in Europe (The European Rail Research Advisory Council 2011), can confirm the expectations from the Dutch mobility analysis (ProRail 2021). Combining this increase in travel volume, with the positive effect of increasing frequencies, a multiplier of four will cope with the increase in demand and slightly increase the service frequency. The international examples also pointed out that most international connections are still quite weak. When the vision of further international improvement is considered, an increase in these services can be expected. However, these long-distance travels would most likely need greater infrastructural investments, which are excluded in this scenario. Therefore, the increase in long-distance travel is considered lower than the medium-distance services. Implementing the findings from these examples into the entire data set, the assumption is made that the status and potential of the Dutch railway network are representative throughout Europe. Furthermore, IC trains that cross borders are assumed to be train trips between areas with a travel time between 1.5 and 3.5 hours. The lower bound is based on the example of Amsterdam - Brussels of just over 1,5 hours. International trains are considered long-distance trips of over 3 hours. This is based on a lower-bound example of Amsterdam - Paris, which takes just over 3 hours to reach one of the airports. The multipliers of 4 times the medium-range services and twice the long-distance services can, therefore, be implemented to the specific services from each FUA in the data set.

Figure 4.1 summarizes the scenario of rail service expansion. The existing track infrastructure is used, while safety and communication systems are improved to facilitate increased rail services, based on the Dutch IMA (ProRail 2021), and extrapolated for Europe. This will show a relatively higher focus on international and longer-distance train travel than national short-distance trains.

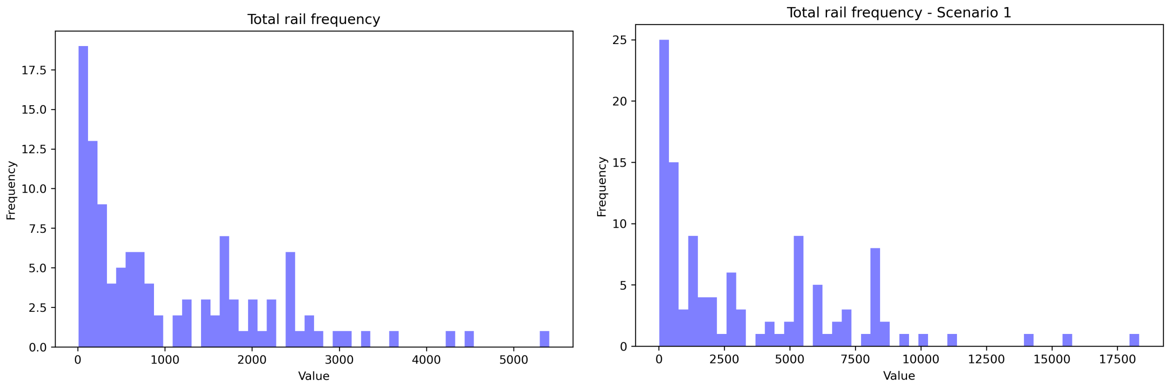
Figure 4.1: Scenario 1: Expand rail services

The execution of this scenario in practice depends on European coordination on infrastructural improvements, the willingness of train operators to implement all extra services, and the possibility of retrieving that many extra trains. This scenario shows a relatively affordable option that could be implemented soon. Compared to the more elaborated land network from the TEN-T, this scenario focuses on the current infrastructure and mainly on improving services, which increases the chances of success. Still, however, the dependency of different stakeholders is a topic needing attention in executing these plans.

4.2.1. Data modifications

The assumed multipliers from Table 4.1 can be applied per range on the OD frequency matrix for rail services used for the benchmark scenario. As discussed, the IC services are assumed to be from 1.5 to 3.5 hours, while the international services have a travel time of over 3.5 hours. The data set for rail services offers a travel time matrix, which can divide the rail services into different categories. The set multipliers are applied to the distinctive categories.

In Figure 4.2 and Figure 4.3, the change in the distribution of the total rail frequencies is shown. This indicates a broader base of the lower to medium frequencies, while the higher outliers more or less remain equal relative to the other frequencies. Figure 4.4 shows the changes for the individual cases and their distribution. The percentages of the changes vary from +100% to +300%, equaling the different multipliers for international (2x) and long-distance (4x) trips. The airports with an increase of 100% are shown in Figure 4.5. The extremely peripheral location of these airports, suggests that these airports do offer medium-distance services in the considered study area. The other extremities, are indicated in Figure 4.4 as Eindhoven (the Netherlands) and Kaunas and Vilnius (Lithuania). The high change of +300% suggests that almost all rail services offered near these airports belong to the category of medium-distance services. No direct long-distance services are served in these FUAs.



(a) Benchmark scenario **(b)** Scenario 1

Figure 4.2: Comparison of total rail frequencies' distributions - Scenario 1

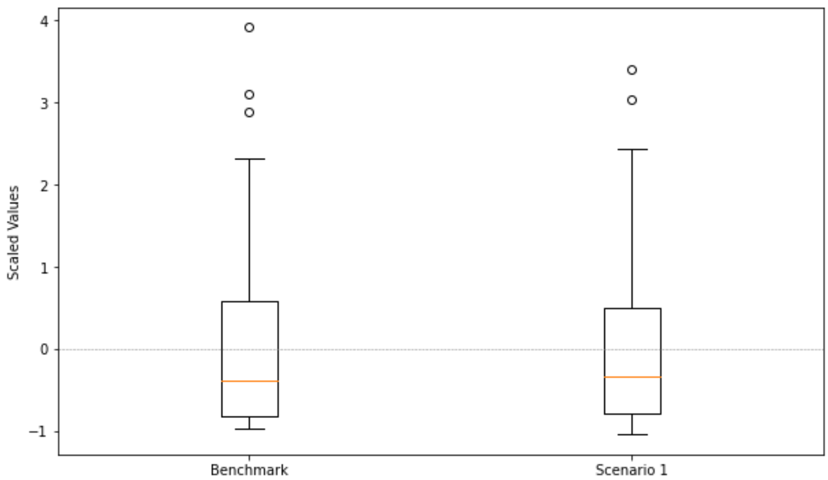


Figure 4.3: Comparison of the scaled value distribution - Scenario 1

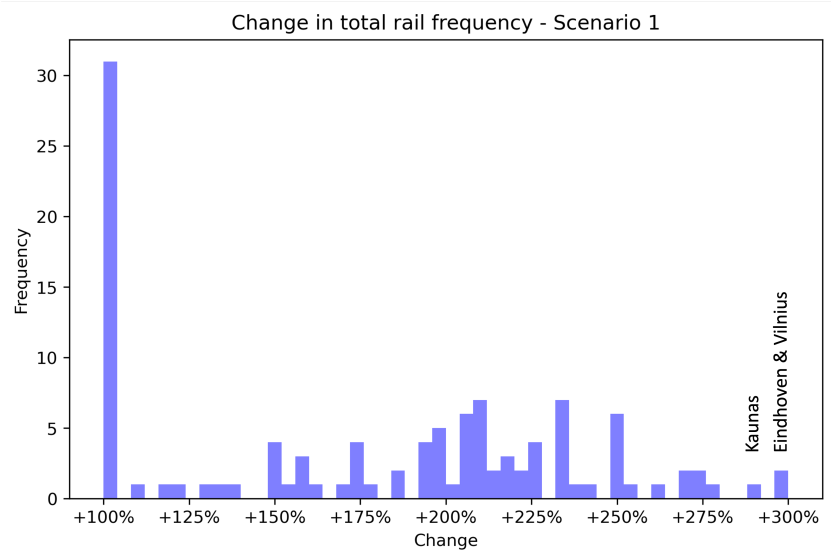


Figure 4.4: Distribution of changes in total rail frequency - Scenario 1

Figure 4.5: Airports with 100% increase in rail services

4.2.2. Numerical results

After the implementation of the data, the process of standardization, PCA and AHC clustering are performed. As a consequence of the change in the variable of total rail frequency, the correlation between variables can be changed. Therefore a new assessment of the number of PCs is needed. Jolliffe's modification of the Kaiser-Guttmann method was mainly used to determine the optimal number of components used in the PCA. The eigenvalues of the PCs after the implementation of scenario 1 are indicated in Table 4.2. Compared to Table 3.2, only minor differences can be found. Therefore, still a number of 3 PCs is used in further analysis of this scenario.

Table 4.2: Eigenvalues of constructed principle components after the implementation of scenario 1

k^{th} **PC**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Eigenvalue	4.72	2.22	0.98	0.70	0.60	0.45	0.19	0.12	0.09	0.03

Using the three PCs, the AHC is performed. The dendrogram in Figure 4.6 still indicates a preference for 5-7 clusters. Also, the Silhouette Scores as shown in Table 4.3 and Figure 4.7 indicate a similar result as the results found in the benchmark situation calculated in section 3.3. However, due to the shift in railway data, some clusters may be less distinctive for classification.

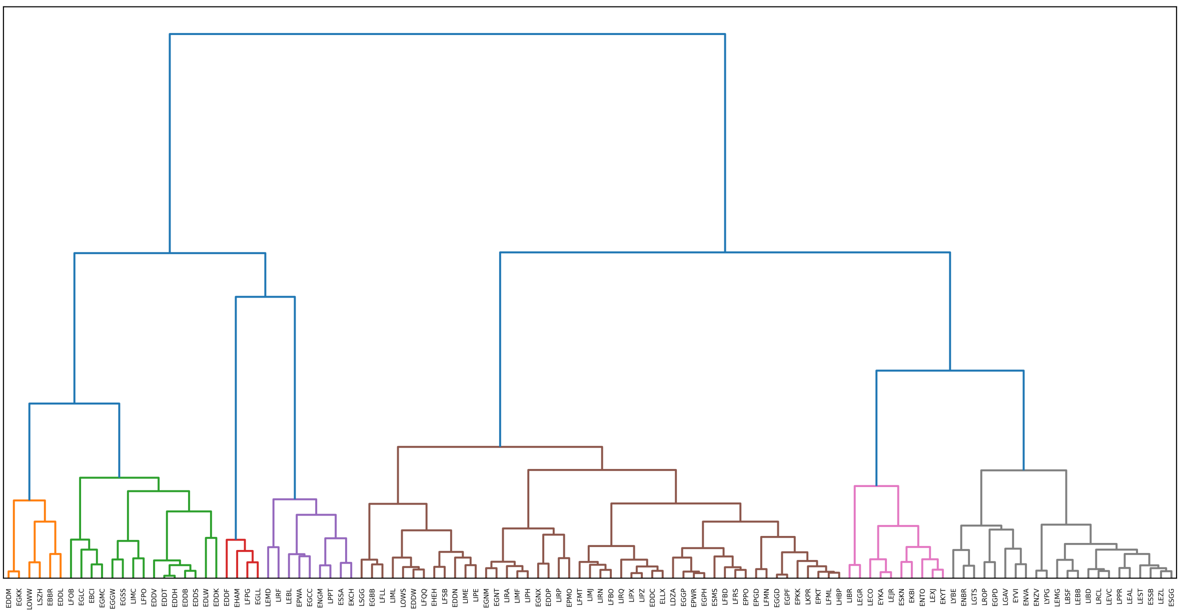


Figure 4.6: Dendrogram of the AHC in scenario 1, using random colors to distinguish 7 clusters.

Table 4.3: Silhouette Score for different numbers of clusters in the AHC of scenario 1

No. of clusters	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Silhouette Score	0.388	0.276	0.309	0.341	0.352	0.347	0.295	0.286	0.278

In section 3.3 a three cluster AHC was found to supply to little distinction for classification possibilities. For this scenario 5 and 7 clusters are considered. The mean values for each variable are shown in Tables 4.4 and 4.5. The comparison between the 5- and 7-cluster AHC shows the same results as the benchmark situation. Comparing the unchanged variables even show that the changes in clustering are minimal. For a 5-clustering AHC, the results are completely equal to the 5-clusters results from the benchmark situation. The extended clustering of 7 clusters shows a small shift between clusters 6 and 7. The distinction between the two clusters is large enough to see a difference in classification.

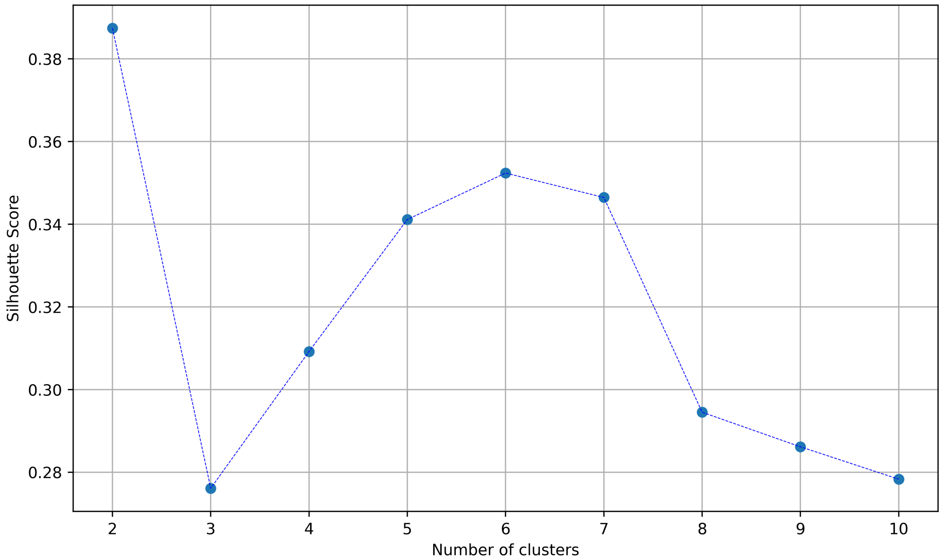


Figure 4.7: Silhouette Score for each number of clusters in scenario 1

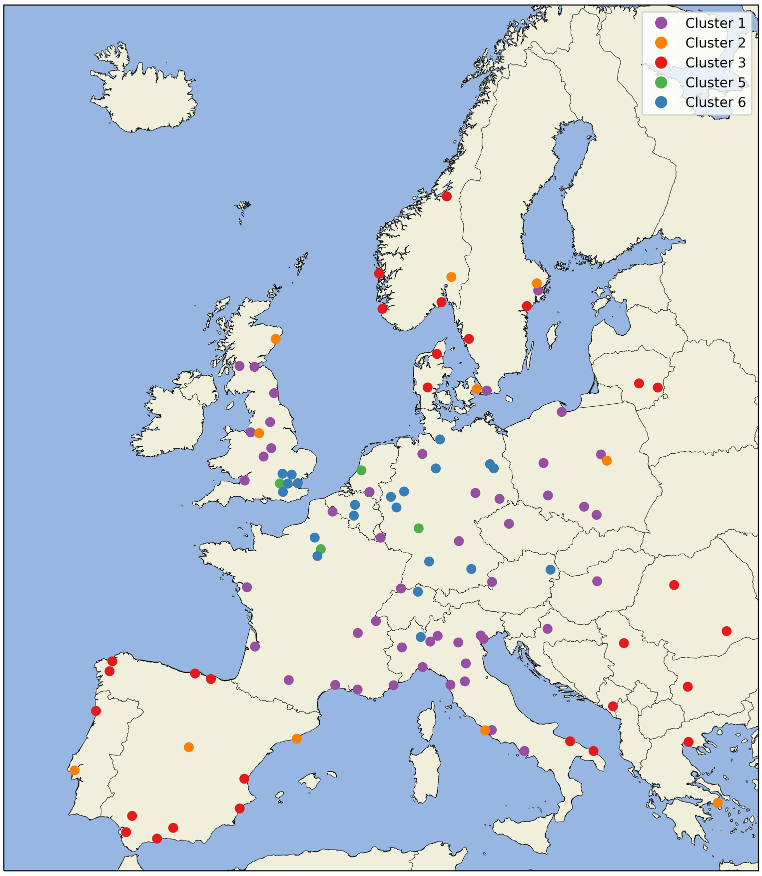
Table 4.4: Mean variable values for 5 clusters in scenario 1

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail	n
1	46.331	4.575	0,09	39	23	0,23	33	2.922	2,86	0,30	48
2	214.264	37.381	0,19	91	10	1,18	33	1.745	3,51	0,16	11
3	38.088	3.121	0,07	32	7	0,38	55	259	3,88	0,04	29
5	452.354	135.437	0,30	181	32	1,75	26	8.861	2,20	0,79	4
6	133.253	21.538	0,14	70	30	0,90	45	8.677	2,28	0,72	21

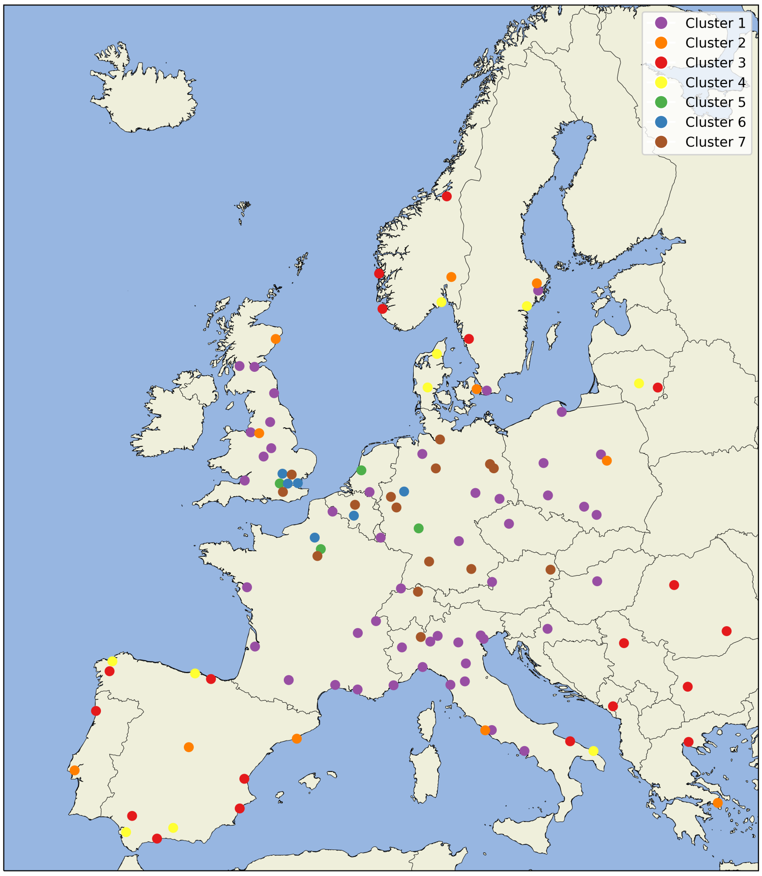
Table 4.5: Mean variable values for 7 clusters in scenario 1

	TF air	IF air	IR air	NC air	AR air	STA	SAC	TF rail	SP rail	CN rail	n
1	46.331	4.575	0,09	39	23	0,23	33	2.922	2,86	0,30	48
2	214.264	37.381	0,19	91	10	1,18	33	1.745	3,51	0,16	11
3	50.884	4.517	0,09	39	7	0,42	34	207	4,04	0,03	19
4	13.775	469	0,03	19	7	0,30	96	358	3,56	0,06	10
5	452.354	135.437	0,30	181	32	1,75	26	8.861	2,20	0,79	4
6	43.983	2.624	0,06	46	30	0,33	71	9.498	2,36	0,81	6
7	168.961	29.103	0,17	80	30	1,13	34	8.349	2,25	0,69	15

*For each variable the **best value**, **2nd best value**, and **worst value** are marked.*



(a) 5 clusters



(b) 7 clusters

Figure 4.8: Maps of clustering for scenario 1

4.2.3. Interpretation of clustering differences

The data from Table 4.5 shows some shifts in clusters when looking at the number of airports in each cluster. However, the variable of air frequencies also shows that the structure of Cluster 2 has changed, despite being the same size as in the benchmark scenario. Table 4.6 shows the movement of airports through the clusters in scenario 1 compared to the benchmark network. This table indicates that a small shift has occurred from clusters 1 and 3 to cluster 2. These airports (Warsaw and Aberdeen) both are airports from the large group that only experienced an increase in rail frequencies of 100%.

Table 4.6: Moved airports in scenario 1

	Added	**Removed**
Cluster 1		*WAW*
Cluster 2	*WAW, ABZ*	
Cluster 3		*ABZ*
Cluster 4		
Cluster 5		
Cluster 6		
Cluster 7		

The effect of the movements of airports is highlighted in Table 4.7. This shows that the individual shifts between clusters 1, 2, and 3 are not causing a change in the interpretation of the network position of these clusters, as only small changes to the mean variables can be seen. Also, the position regarding the total rail frequency is equal to the benchmark situation, although the values have changed with the implementation of scenario 1.

Table 4.7: Differences highlighted for changed clusters between benchmark and scenario 1

	TF air	**IF air**	**IR air**	**NC air**	**AR air**	**STA**	**SAC**	**TF rail**	**SP rail**	**CN rail**	**n**
B1	48.456	4.913	0,09	40	23	**0,24**	33	925	2,86	0,30	49
S1	46.331	4.575	0,09	39	23	**0,23**	**33**	2.922	2,86	0,30	48
B2	**237.474**	**39.747**	0,16	**100**	10	**1,33**	**31**	617	3,61	0,16	9
S2	**214.264**	**37.381**	**0,19**	**91**	10	**1,18**	**33**	1.745	3,51	0,16	11
B3	51.800	5.908	0,11	38	**7**	0,40	34	**102**	**4,02**	**0,04**	20
S3	50.884	4.517	0,09	39	**7**	0,42	34	**207**	**4,04**	**0,03**	19

4.2.4. Classification

As there has been no significant changes in the clustering, there is no need for a revision of the cluster interpretation and classification. Therefore, the classification remains the same. The complete structure for all classes is shown in Table 4.8. Its visualization is shown on the map in Figure 4.9.

Table 4.8: Classifications after implementing scenario 1

C#	Class	n	Airports
C5	Main Intermodal Hubs	4	*CDG, FRA, AMS, LHR*
C7	Central International Hubs	15	*VIE, BRU, ORY, SXF, HAM, CGN, DUS, MUC, STR, TXL, HAJ, MXP, ZRH, LGW, STN*
C2	Peripheral International Hubs	9	*CPH, ATH, FCO, OSL, WAW, LIS, BCN, MAD, ARN, MAN, ABZ*
C6	Central Secondary Airports	6	*BVA, LCY, SEN, CRL, DTM, LTN*
C1	Central Minor Airports	49	*Remaining airports*
C3	Peripheral Minor Airports	20	*BIO, VLC, SVQ, CLJ, SOF, SKG, BRI, VNO, TGD, BGO, TRD, SVG, OPO, OTP, BEG, ALC, AGP, SCQ, GOT*
C4	Disconnected Airports	10	*LCG, GRX, XRY, SDR, BDS, KUN, TRF, NYO, BLL, AAL*

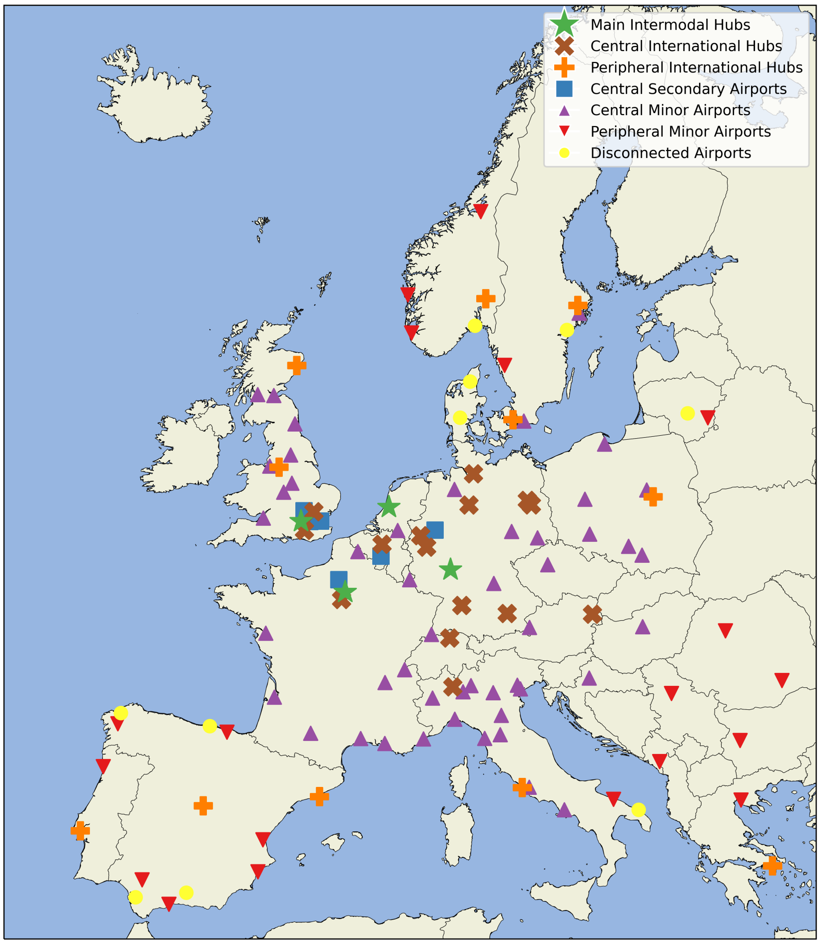


Figure 4.9: Scenario 1 in seven classes

4.3. Scenario 2: Short-haul flight reduction

Many airports suffer from capacity problems. Some are caused by the lack of space for expansion opportunities, but environmental factors also play a role. In the meantime, governments are aiming for a reduction in short-haul flights. France already banned flights that can be replaced by a train journey of less than 2,5 hours from 2023. Losing these flights, some of the allocated slots are released to use for intercontinental flights.

The shift from short-haul flights to train journeys can help airports offer more services to longer-distance destinations, even when airport expansion is impossible. This also gives airports and airlines an incentive to cooperate with railway companies. This case enables airports to offer different services while rail companies gain extra travelers.

The governmental influence on the division of slots is indirect Appendix B. In the Netherlands, the government can only stimulate desired or discourage undesired behavior by several policy choices. However, the division of the slots is influenced by the demand of airlines and airports and the decision of the slot allocation party. Also, it is expected that the hub characteristics of airports from the air perspective will not change significantly. Therefore, it can be assumed that smaller airports will not deviate too much from their current division of destination ranges.

The scenario consists of actions that consider flight reduction and the reaction of airports and airlines, shown in Figure 4.10. At first, the example of the super-short-haul flight ban from France is extrapolated throughout Europe. All flights below 500 kilometers will be banned in this scenario. In reality, this will need the mentioned change in authority for the European Union and the slots allocator. For now, France could only ban domestic flights because the international example of Brussels-Paris falls outside of France's authority. The assumption in this scenario is made as if the European Union holds the power to ban international flights as well.
Secondly, short- to medium-haul flights are heavily reduced by rail possibilities. Flights between airports within 500-1000 kilometer distance are reduced by 50%. It is assumed that improved railway services cause a shift in mode choice, but no ban on these flights is performed, which maintains half of the flights in this range.
Lastly, the slots released by reducing short- and medium-haul flights are filled by offering higher frequencies on the remaining flights. It is assumed that the airports will provide the same ratio of medium- and long-haul continental flights compared to intercontinental flights. Schiphol Amsterdam and Manchester Airport examples are shown in Table 4.9.

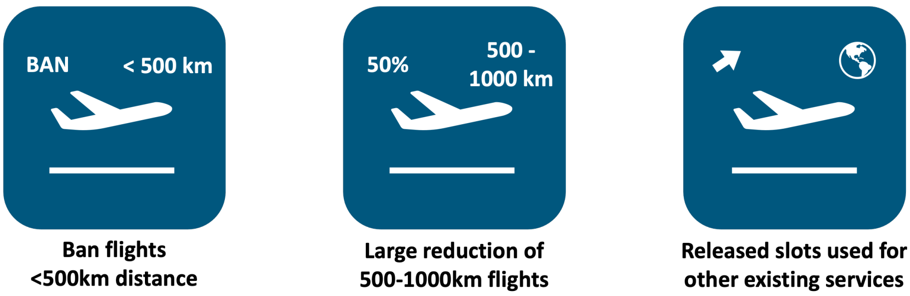


Figure 4.10: Scenario 2: Shift in slot allocation

The scenario responds to the expectation of airports not willing to shrink in services. However, it facilitates the growth of mobility, while the desire to limit the growth of airports is satisfied. A redistribution of the slots per airport can change the focus of offered air services and, thus, the position of an airport in the intermodal network.

Amsterdam	**SH**	**MH**	**Rest**	**Total**
Benchmark	33.267	43.198	381.602	458.067
Scenario	0	21.599	436.468	458.067
Change	-100%	-50%	+14,4%	

Manchester	**SH**	**MH**	**Rest**	**Total**
Benchmark	5.574	4.357	153.733	163.664
Scenario	0	2.179	161.486	163.664
Change	-100%	-50%	+5,0%	

Table 4.9: Examples of scenario implementation in Amsterdam and Manchester

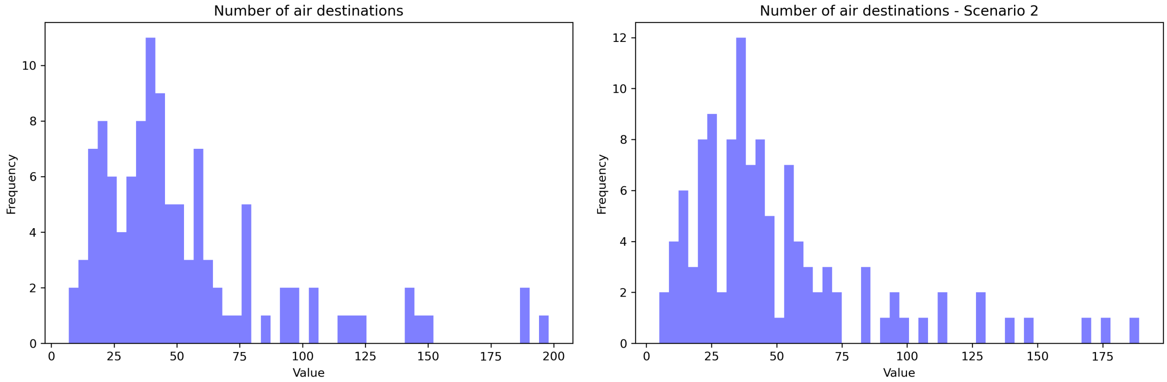
4.3.1. Data modification

The reduction of flights is applied to two categories: flights within a 500-kilometer range and flights in a range of 500-1000 kilometers. The geodesic distances between the airports in the used selection and all destination airports calculate these ranges. The frequencies between destinations in the short-haul range are removed from the frequency matrix, and the frequencies in the medium-haul range are reduced by 50%. The total number of removed flights is then distributed according to the ratio over the longer-distance flights. The examples in Table 4.9 show the calculation for the percentage increase of long-distance flights. This means that, in the case of Amsterdam Airport, a long-haul destination with 100 annual flights, will offer 114 flights after the implementation of this scenario.

The scenario affects a few variables. Firstly, the number of destinations shrinks by canceling flights of less than 500 kilometers. The distribution of this variable is shown in Figure 4.11, showing some small shifts but a similar distribution as in the benchmark scenario. The distribution of the relative changes at each airport from Figure 4.12 mainly shows a decrease in the number of destinations of up to 10%. Also, some large reductions are seen, but a few increases are identified. This is caused by the rise in some longer-distance destinations, after redistributing the slots freed by the short-haul flight reduction. For some of these destinations the frequencies could be pushed just over the set threshold for being a significant connection. These airports where this occurs are relatively small. The average of annual flights at the indicated airports in Figure 4.12 is approximately 25,000. These smaller airports are more sensitive for changes around the significant connection threshold.

The distribution of the performed air services is different as well. This changes the number of intercontinental flights and the ratio compared to the total number of flights. Figures 4.13 and 4.14 show minor differences in the distribution of these variables. The histogram in Figure 4.15 shows that most airports have an increase of 0% to 150%. Only some outliers experience a larger increase or decrease in intercontinental flights. The reduction is seen at Alicante (Spain) and Otopeni (Romania) airports. These cities are located at the edges of the study area and have possible short-distance destinations on other continents, which causes a decrease in intercontinental flights. On the other hand, the extreme increases happen at the airports of Bergen, Stavanger (Norway), and Aalborg (Denmark), which have many short-distance destinations in Scandinavia. Airports with a relatively high number of short-distance destinations can show an artificially high increase in intercontinental flights because many flights are removed and redistributed over a few intercontinental destinations.

The new distributions of the scaled variable values are shown in Figure 4.16. For all variables, the general distribution has not changed significantly. However, from an individual airport's perspective, some significant changes were found in Figures 4.12 and 4.15, as mentioned earlier.



(a) Benchmark scenario **(b)** Scenario 2

Figure 4.11: Comparison of distribution in the number of destinations - Scenario 2

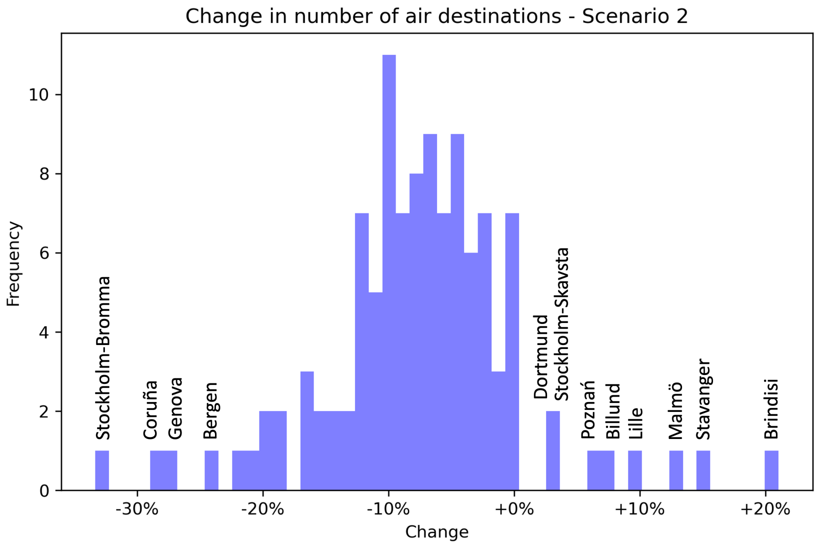
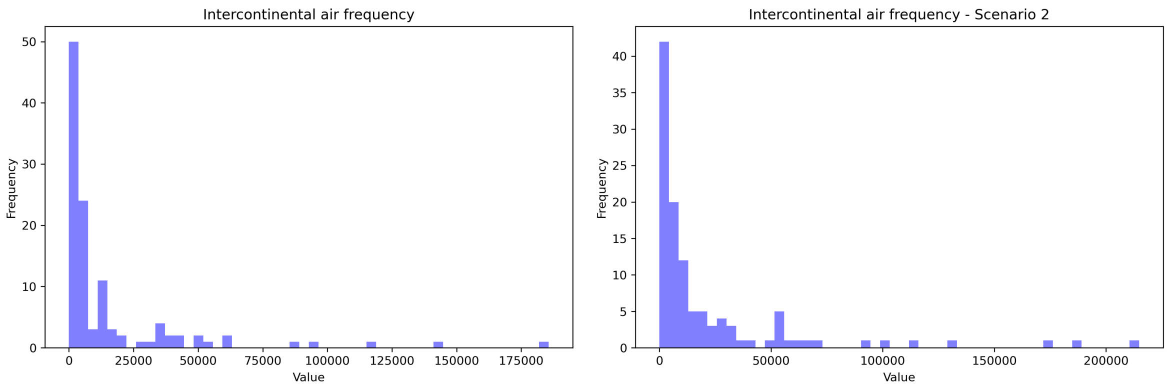
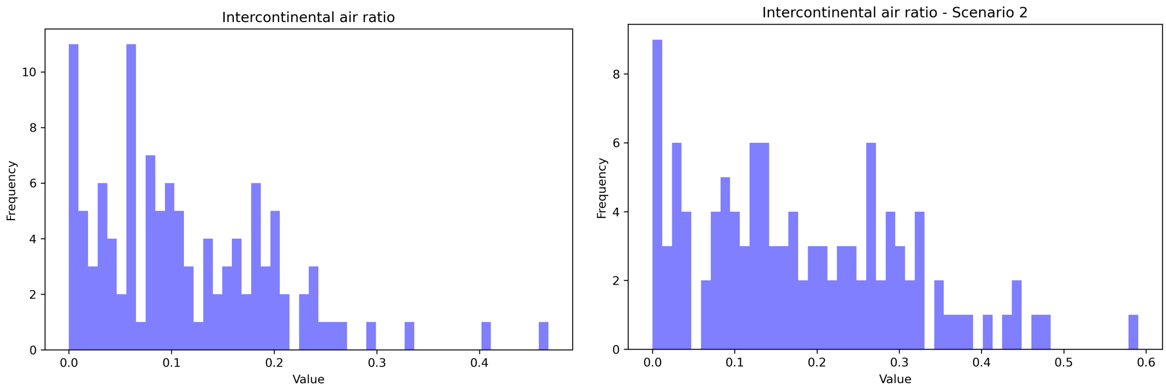


Figure 4.12: Distribution of changes in the number of air destinations - Scenario 2



(a) Benchmark scenario **(b)** Scenario 2

Figure 4.13: Comparison of intercontinental air frequencies' distributions - Scenario 2



(a) Benchmark scenario

(b) Scenario 2

Figure 4.14: Comparison of intercontinental air ratios' distributions - Scenario 2

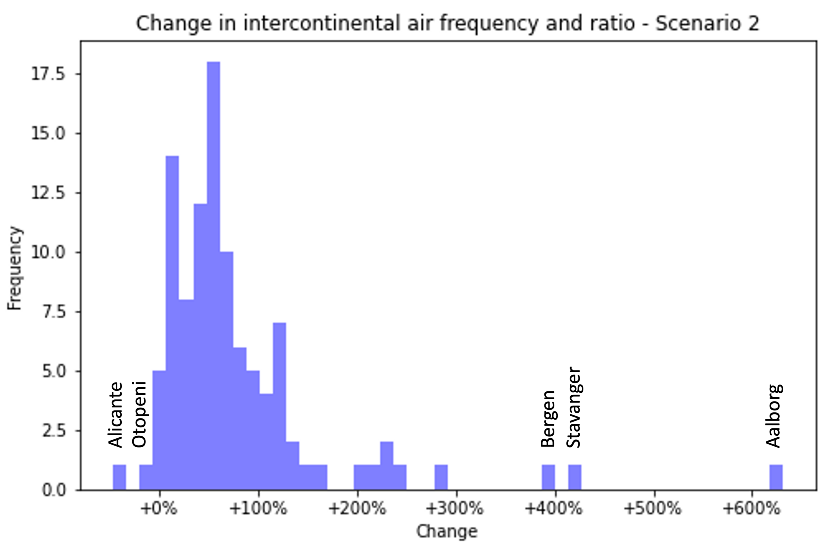


Figure 4.15: Distribution of changes in intercontinental frequencies and its ratio - Scenario 2

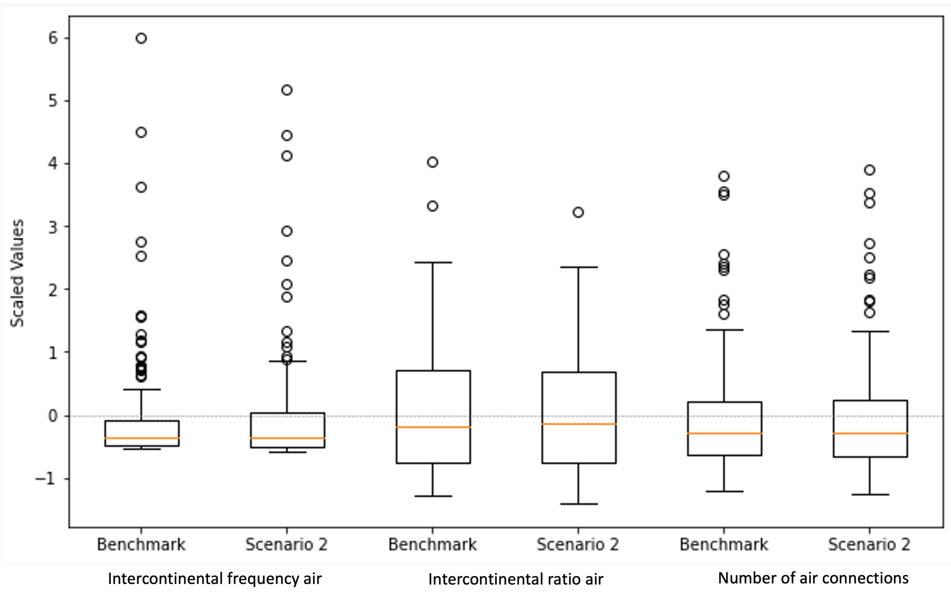


Figure 4.16: Comparison of the scaled value distributions - Scenario 2

4.3.2. Numerical results

The distribution in the individual variables again seems to be not significantly different, and also, the correlation in the data set has not changed noticeably. Table 4.10 shows the eigenvalues of the PCs for the new data set. In this case, still 3 PCs are sufficient to describe the data in the same way as the previous situations.

Table 4.10: Eigenvalues of constructed principle components after the implementation of scenario 2

k^{th} **PC**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Eigenvalue	4.80	2.04	0.99	0.78	0.62	0.44	0.18	0.11	0.09	0.04

The dendrogram in Figure 4.17 has similar characteristics as the previously created dendrograms. The silhouette scores in Table 4.11 and Figure 4.18 also show a short reduction at four clusters, followed by some higher scores. Again a theoretically high preference for a two or three-cluster division is found. In previous cases, there was no clear difference from 5 to 7 clusters. This scenario shows a stronger theoretical preference for 7 clusters. To see possible differences in build-up of the clusters, the 5 to 7 cluster analyses are considered.

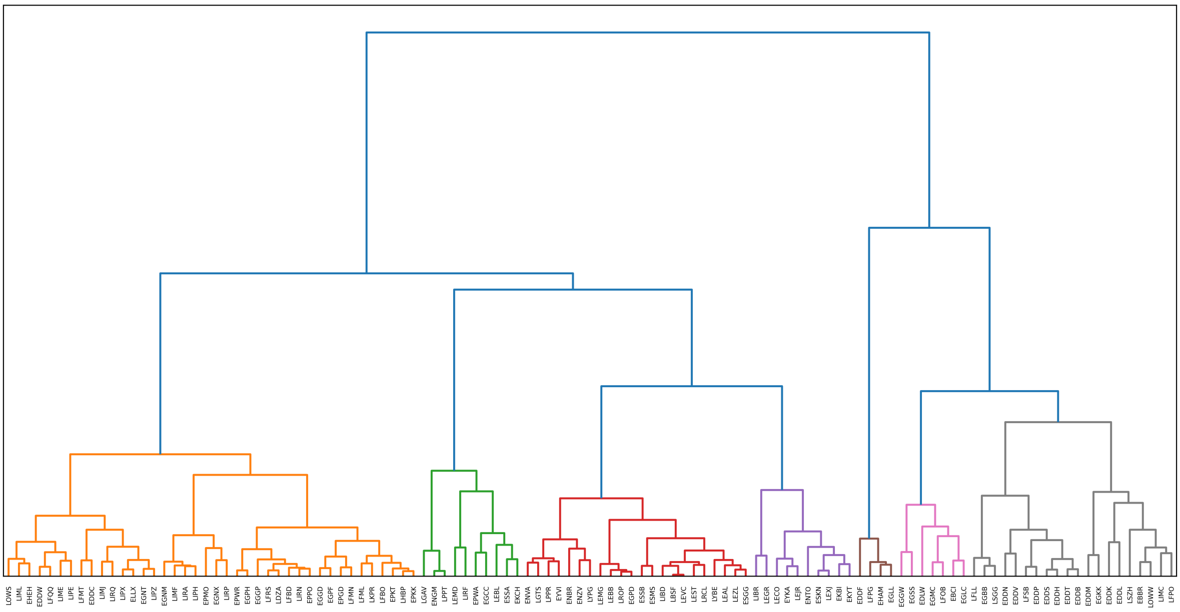


Figure 4.17: Dendrogram of the AHC in scenario 2, using random colors to distinguish 7 clusters.

Table 4.11: Silhouette Score for different numbers of clusters in the AHC of scenario 2

No. of clusters	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Silhouette Score	0.382	0.362	0.275	0.331	0.332	0.348	0.330	0.312	0.312

The values for 5 and 7 clusters are shown in Table 4.12. Still all seven clusters are distinctive and can be useful in the classification. Therefore, also in this scenario a 7-cluster analysis is used. In the tables, it can be identified that the clusters have similar distinctions as in the other situations. The cluster numbers in the 7-cluster analysis are set to equal the benchmark situation as much as possible. Figure 4.19 shows the differences between the 5 cluster AHC and the 7 cluster AHC. Cluster 4 has merged into cluster 3, and cluster 7 has merged into cluster 6 from the 7- to the 5-cluster analysis.

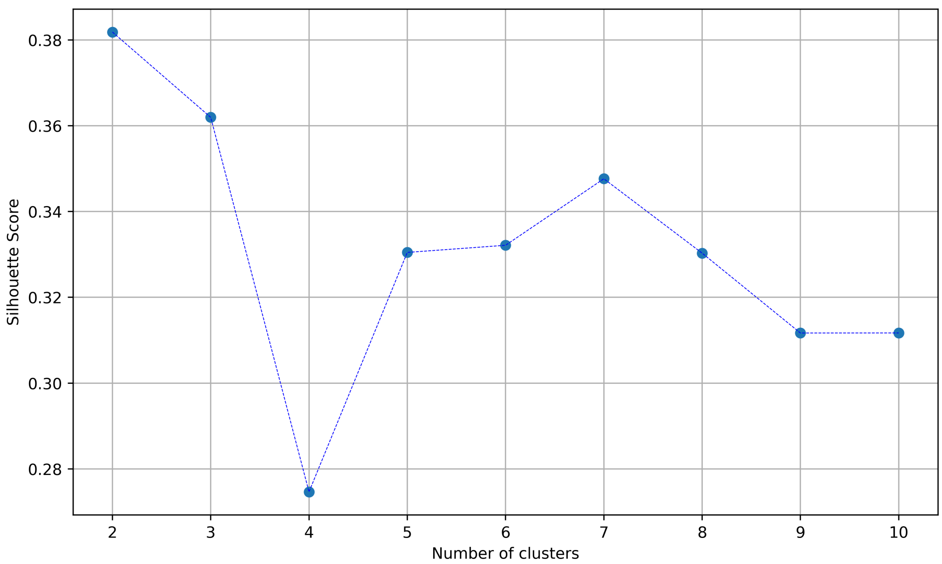


Figure 4.18: Silhouette Score for each number of clusters in scenario 2

Table 4.12: Mean variable values for different numbers of clusters in scenario 2

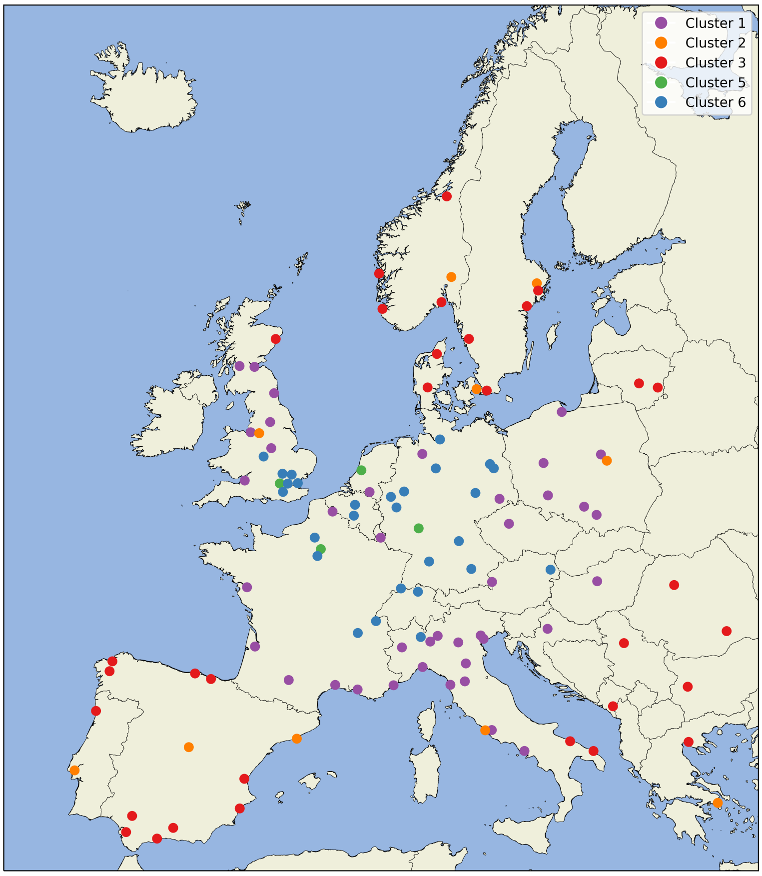
(a) 5 clusters

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail	n
1	43.946	6.407	0,15	36	22	**0,15**	**32**	894	2,89	0,30	40
2	**228.771**	**52.704**	0,23	**89**	10	**1,30**	**32**	627	3,52	0,17	10
3	**38.557**	**5.041**	**0,12**	**29**	**7**	0,34	**54**	**144**	**3,81**	**0,05**	32
5	**452.354**	**177.619**	**0,39**	**171**	**32**	**1,75**	**26**	**2.852**	**2,20**	**0,79**	4
6	118.608	29.430	**0,24**	61	**30**	0,89	42	**2.404**	**2,36**	**0,65**	27

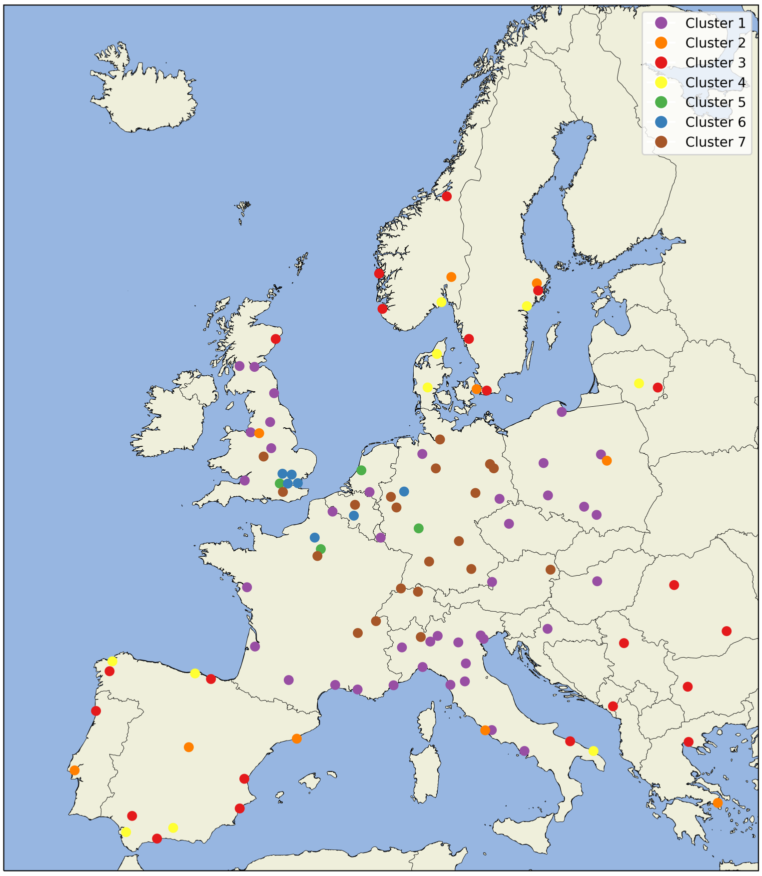
(b) 7 clusters

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail	n
1	43.946	6.407	0,15	36	22	**0,15**	**32**	894	2,89	0,30	40
2	**228.771**	**52.704**	0,23	**89**	10	**1,30**	**32**	627	3,52	0,17	10
3	49.821	6.829	0,14	33	**7**	0,36	35	**139**	**3,93**	**0,04**	22
4	**13.775**	**1.105**	**0,08**	**19**	**7**	0,30	**96**	157	3,56	0,06	10
5	**452.354**	**177.619**	**0,39**	**171**	**32**	**1,75**	**26**	**2.852**	**2,20**	**0,79**	4
6	59.501	5.083	0,07	55	30	0,43	70	**2.770**	2,39	**0,83**	7
7	139.295	37.951	**0,30**	63	**31**	1,05	**32**	2.276	**2,34**	0,59	20

*For each variable the **best value**, **2nd best value**, and **worst value** are marked.*



(a) 5 clusters



(b) 7 clusters

Figure 4.19: Maps of clustering for scenario 2

4.3.3. Interpretation of clustering differences

Implementing the second scenario has affected all airside variables except the total frequency. Table 4.12b showed all mean values for the new cluster distribution. Cluster 1 has become smaller after the implementation, while Cluster 7 has grown. Table 4.13 shows that the scenario mainly caused a redistribution of several airports from Cluster 1. Clusters 4 and 5, also in this scenario, remain unchanged.

Table 4.13: Moved airports in scenario 2

	Added	Removed
Cluster 1		*BSL, BHX, GVA, LEJ, BMA, LYS, WAW, MMX, NUE*
Cluster 2	*WAW*	
Cluster 3	*BMA, MMX*	
Cluster 4		
Cluster 5		
Cluster 6	*STN*	
Cluster 7	*BSL, BHX, GVA, LEJ, LYS, NUE*	*STN*

With several airports leaving Cluster 1, the cluster seems to score lower than the benchmark scenario, as shown in the comparison of the clusters in Table 4.14. With the same shift of airports moving from Cluster 1, Cluster 7 is also decreasing its score slightly. This is mainly caused by the fact that the moved airports are 'high-performers' in the first cluster, while they are low-performing airports in Cluster 7. An example is Lyon Airport (LYS), with a total air frequency of 84.000 flights, of which 21% is intercontinental, and 1.800 rail departures. The change in air variables again intensifies the sensitivity of the same variables for airports. Warsaw airport, on the other hand, is moving from Cluster 1 to 2. This airport, with 150.000 flights, 26% intercontinental, was an outlier on that factor in Cluster 1 but changed to the peripheral international hubs class by the higher emphasis on intercontinental flights. Furthermore, Stockholm-Bromma (BMA) and Malmö (MMX) are moving from cluster 1 to cluster 3. Both airports are peripheral situated airports, that also are one of the airports with an increasing number of air destinations, as indicated in Figure 4.12.

Table 4.14: Differences highlighted for changed clusters between benchmark and scenario 2

	TF air	IF air	IR air	NC air	AR air	STA	SAC	TF rail	SP rail	CN rail	n
B1	48.456	4.913	0,09	40	23	**0,24**	33	925	2,86	0,30	49
S1	43.946	6.407	0,15	36	22	**0,15**	**32**	894	2,89	0,30	40
B2	**237.474**	**39.747**	0,16	**100**	10	**1,33**	**31**	617	3,61	0,16	9
S2	**228.771**	**52.704**	0,23	**89**	10	**1,30**	**32**	627	3,52	0,17	10
B3	51.800	5.908	0,11	38	**7**	0,40	34	**102**	**4,02**	**0,04**	20
S3	49.821	6.829	0,14	33	**7**	0,36	35	**139**	**3,93**	**0,04**	22
B6	43.983	2.624	0,06	46	**30**	0,33	71	**2.827**	2,36	**0,81**	6
S6	59.501	5.083	0,07	55	30	0,43	70	**2.770**	2,39	**0,83**	7
B7	168.961	29.103	**0,17**	80	**30**	1,13	34	2.676	**2,25**	0,69	15
S7	139.295	37.951	**0,30**	63	**31**	1,05	**32**	2.276	**2,34**	0,59	20

4.3.4. Classification

Despite the several changes in airports' clustering, again, no difference in cluster interpretation is found. This causes a cluster division as found in Table 4.15, which is plotted on the map in Figure 4.20.

Table 4.15: Classifications after implementing scenario 2

C#	Class	n	Airports
C5	**Main Intermodal Hubs**	4	*CDG, FRA, AMS, LHR*
C7	**Central International Hubs**	15	*VIE, BRU, LYS, ORY, BSL, SXF, HAM, CGN, DUS, MUC, NUE, LEJ, STR, TXL, HAJ, MXP, GVA, ZRH, BHX, LGW*
C2	**Peripheral International Hubs**	9	*CPH, ATH, FCO, OSL, WAW, LIS, BCN, MAD, ARN, MAN*
C6	**Central Secondary Airports**	6	*BVA, LCY, SEN, CRL, DTM, LTN, STN*
C1	**Central Minor Airports**	49	*Remaining airports*
C3	**Peripheral Minor Airports**	20	*BIO, VLC, SVQ, CLJ, BMA, SOF, SKG, BRI, VNO, TGD, BGO, TRD, SVG, OPO, OTP, BEG, ALC, AGP, SCQ, GOT, MMX, ABZ*
C4	**Disconnected Airports**	10	*LCG, GRX, XRY, SDR, BDS, KUN, TRF, NYO, BLL, AAL*

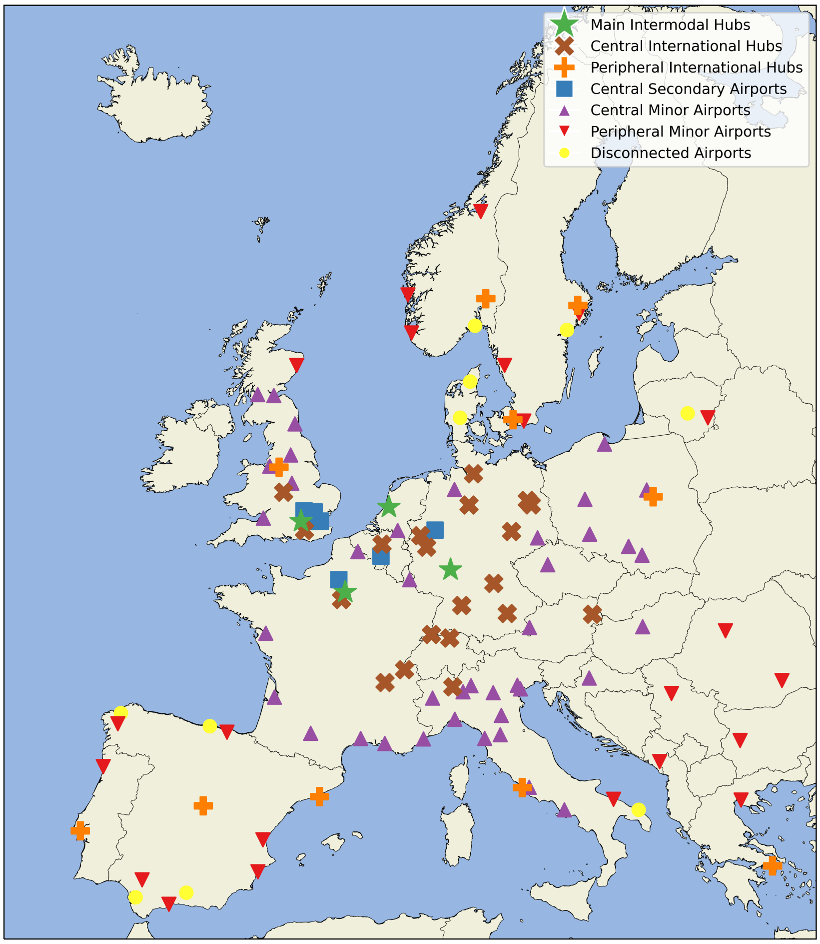


Figure 4.20: Scenario 2 in seven classes

4.4. Scenario 3: Air/Rail policy combination

The previously addressed scenarios show the two sides of the air/rail integration. However, to achieve a higher contribution to the network intermodality, a combination of the two scenarios could cause another shift in the intermodal network. Combining adding extra services with the imposed reduction of short-haul flights creates a different scenario. This scenario involves an even more complex coordination of international stakeholders, including railway managers and operators, airports, airlines, and governments.

4.4.1. Data modification

As this scenario combines scenario 1 and scenario 2, and both scenarios affect different variables in the data set, no extra efforts for the data implementation are needed. The data from scenario 1 and scenario 2 are combined into a new data set, containing increased rail services and a shift in the offered services on the air side. This means that the data comparisons in Figures 4.3 and 4.16 also apply to the third combination scenario.

4.4.2. Numerical results

Combining two scenarios and the subsequent change in multiple variables creates a more complex data set to cluster. Nevertheless, Table 4.16 shows that no extra PCs are needed according to the interpretation of the Kaiser-Guttmann method in the previous situation.

Table 4.16: Eigenvalues of constructed principle components after the implementation of scenario 3

k^{th} **PC**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Eigenvalue	4.84	2.06	0.99	0.74	0.60	0.46	0.17	0.11	0.09	0.03

The dendrogram in Figure 4.21 has a distinctive build-up compared to the previous situations. The linkage distance for the two-cluster solution is much smaller, which, for the first time, shows a clear theoretical preference for multiple clusters. This is confirmed by the silhouette scores, displayed in Table 4.17 and Figure 4.22. The graph again shows a preference for 5-7 clusters, while in this case, the 6-cluster solution has a small advantage.

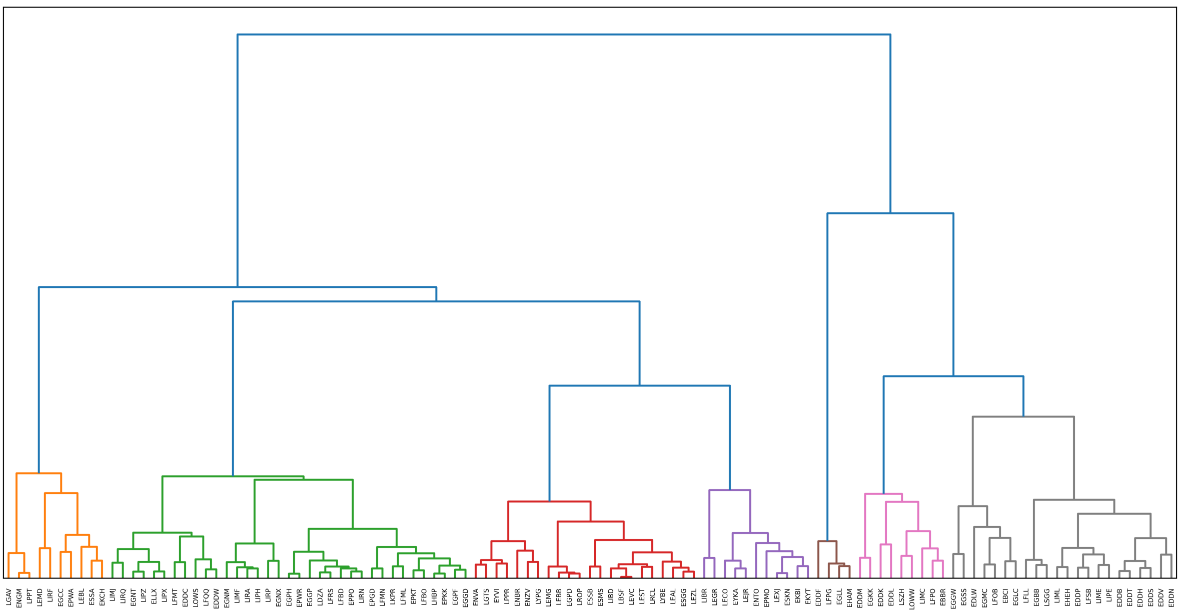


Figure 4.21: Dendrogram of the AHC in scenario 3, using random colors to distinguish 7 clusters.

Table 4.17: Silhouette Score for different numbers of clusters in the AHC of scenario 2

No. of clusters	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Silhouette Score	0.264	0.307	0.298	0.346	0.361	0.349	0.321	0.312	0.315

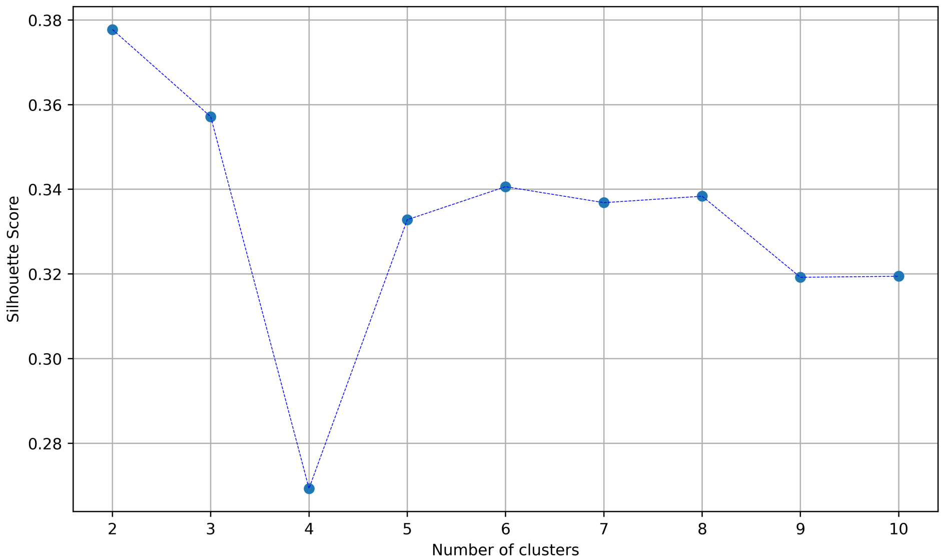


Figure 4.22: Silhouette Score for each number of clusters in scenario 3

The higher silhouette scores for the 5-7 cluster solutions confirm the choice to compare these solutions in the previous cases. The 6-cluster division is also highlighted in this scenario because of the apparent peak in theoretical preference for a 6-cluster solution. The mean variable values are shown in Table 4.18. The maps corresponding with these clusterings are shown in Figure 4.23.

Table 4.18: Mean variable values for different numbers of clusters in scenario 3

(a) 5 clusters

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail	n
1	43.193	6.740	0,16	35	22	**0,17**	**32**	2.515	2,93	0,27	35
2	**228.771**	**52.704**	**0,23**	**89**	10	**1,30**	**32**	1.855	3,52	0,17	10
3	**37.898**	**4.900**	**0,12**	**29**	**7**	0,33	**55**	**386**	**3,78**	**0,05**	33
5	**452.354**	**177.619**	**0,39**	**171**	**32**	**1,75**	**26**	**8.861**	**2,20**	**0,79**	4
6	110.698	26.276	0,22	60	**30**	0,77	40	**7.414**	**2,39**	**0,63**	31

(b) 6 clusters

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail	n
1	43.193	6.740	0,16	35	22	**0,17**	**32**	2.515	2,93	0,27	35
2	**228.771**	52.704	0,23	**89**	10	1,30	**32**	1.855	3,52	0,17	10
3	**37.898**	**4.900**	**0,12**	**29**	**7**	0,33	**55**	**386**	**3,78**	**0,05**	33
5	**452.354**	**177.619**	**0,39**	**171**	**32**	**1,75**	**26**	**8.861**	**2,20**	**0,79**	4
6	69.539	13.104	0,19	48	30	0,55	42	6.986	2,47	0,62	22
7	211.308	**58.475**	**0,28**	**89**	**31**	**1,33**	34	**8.462**	**2,19**	**0,67**	9

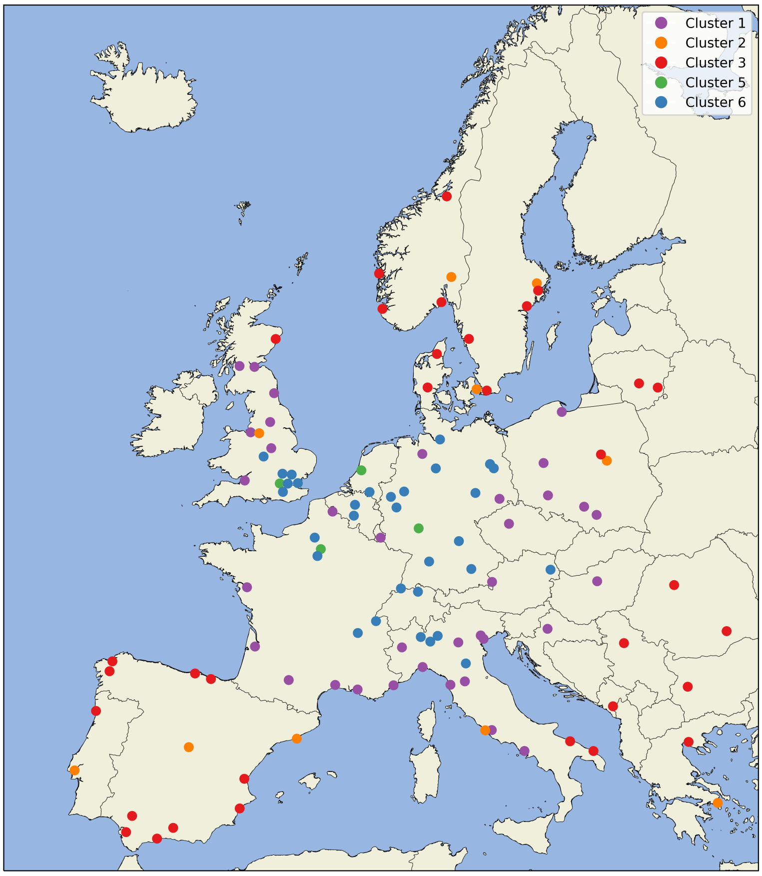
(c) 7 clusters

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail	n
1	43.193	6.740	0,16	35	22	**0,17**	**32**	2.515	2,93	0,27	35
2	**228.771**	52.704	0,23	**89**	10	1,30	**32**	1.855	3,52	0,17	10
3	49.821	6.829	0,14	33	**7**	0,36	35	**351**	**3,93**	**0,04**	22
4	**14.053**	**1.041**	**0,07**	**20**	**7**	0,27	**94**	456	3,48	0,07	11
5	**452.354**	**177.619**	**0,39**	**171**	**32**	**1,75**	**26**	**8.861**	**2,20**	**0,79**	4
6	69.539	13.104	0,19	48	30	0,55	42	6.986	2,47	0,62	22
7	211.308	**58.475**	**0,28**	**89**	**31**	**1,33**	34	**8.462**	**2,19**	**0,67**	9

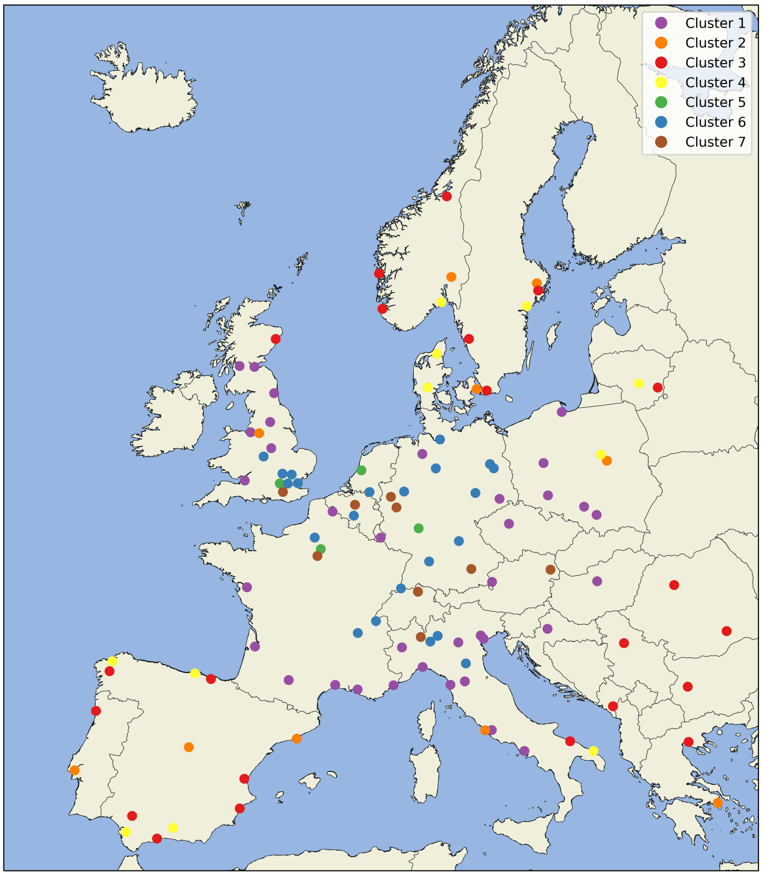
(d) 8 clusters

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail	n
1	43.193	6.740	0,16	35	22	**0,17**	32	2.515	2,93	0,27	35
2	**228.771**	52.704	0,23	**89**	10	1,30	32	1.855	3,52	0,17	10
3	49.821	6.829	0,14	33	**7**	0,36	35	**351**	**3,93**	**0,04**	22
4	**14.053**	**1.041**	**0,07**	**20**	**7**	0,27	**94**	456	3,48	0,07	11
5	**452.354**	**177.619**	**0,39**	**171**	**32**	**1,75**	**26**	**8.861**	**2,20**	**0,79**	4
6a	74.224	16.847	0,25	44	30	0,60	**29**	5.907	2,50	0,52	15
6b	59.501	8.683	0,16	55	30	0,43	70	**9.297**	2,39	**0,83**	7
7	211.308	**58.475**	**0,28**	**89**	**31**	**1,33**	34	8.462	**2,19**	0,67	9

*For each variable the **best value**, **2nd best value**, and **worst value** are marked.*



(a) 5 clusters



(b) 7 clusters

Figure 4.23: Maps of clustering for scenario 3

4.4.3. Interpretation of clustering differences

Scenario 3 combines both previous scenarios, considering changes in the air and rail factors. Also, at first glance, the effects on the clustering look to be a combination of both scenarios and the shifts that occurred in their implementation. Compared to the benchmark situation, Cluster 1 has become smaller, as happened after the implementation of scenario 2. However, these airports are redistributed over Cluster 6 instead of Cluster 7. As most airports join Cluster 6, this cluster can be interpreted differently than the *Central Secondary Airports* from the previous scenarios. The airports that have changed clusters are shown in Table 4.19.

Table 4.19: Moved airports in scenario 3

	Added	**Removed**
Cluster 1		*NUE, WAW, GVA, LYS, LEJ, BGY, LIN, BSL, MMX, BHX, EIN, BLQ, WMI, BMA*
Cluster 2	*WAW*	
Cluster 3	*BMA, MMX*	
Cluster 4	*WMI*	
Cluster 5		
Cluster 6	*NUE, GVA, STN, LYS, LEJ, BGY, TXL, LIN, BSL, BHX, EIN, BLQ, HAJ, SXF, STR, HAM*	
Cluster 7		*STN, TXL, HAJ, SXF, STR, HAM*

In this scenario, only Cluster 5, with the four main hubs, remains unchanged compared to the benchmark network. The changes within the variables for each cluster are shown in Table 4.20. It is also noted that all airports leaving Cluster 7 join Cluster 6. London Stansted is one of the airports and matches the intercontinental flight variables (23.000, 15%) better with Cluster 6. The other shifts of clustering are not having a big impact on the variables' values.

Table 4.20: Differences highlighted for changed clusters between benchmark and scenario 3

	TF air	**IF air**	**IR air**	**NC air**	**AR air**	**STA**	**SAC**	**TF rail**	**SP rail**	**CN rail**	**n**
B1	48.456	4.913	0,09	40	23	**0,24**	33	925	2,86	0,30	49
S1	43.193	6.740	0,16	35	22	**0,17**	**32**	2.515	2,93	0,27	35
B2	**237.474**	**39.747**	0,16	**100**	10	**1,33**	**31**	617	3,61	0,16	9
S2	**228.771**	52.704	0,23	**89**	10	1,30	**32**	1.855	3,52	0,17	10
B3	51.800	5.908	0,11	38	**7**	0,40	34	**102**	**4,02**	**0,04**	20
S3	49.821	6.829	0,14	33	**7**	0,36	35	**351**	**3,93**	**0,04**	22
B4	**13.775**	**469**	**0,03**	**19**	**7**	0,30	**96**	157	3,56	0,06	10
S4	**14.053**	**1.041**	**0,07**	**20**	**7**	0,27	**94**	456	3,48	0,07	11
B6	43.983	2.624	0,06	46	**30**	0,33	71	**2827**	2,36	**0,81**	6
S6	69.539	13.104	0,19	48	30	0,55	42	6.986	2,47	0,62	22
B7	168.961	29.103	**0,17**	80	**30**	1,13	34	2676	**2,25**	0,69	15
S7	211.308	**58.475**	**0,28**	**89**	**31**	**1,33**	34	**8.462**	**2,19**	**0,67**	9

4.4.4. Classification

The increase in the number of airports assigned to Cluster 6 forms a situation in which not all airports are secondary for their catchment area. Still, the connecting factor within this cluster is the degree of rail connections. Therefore, this cluster can be classified as *Land-connected Minor Airports*, showing the opposition of high performance on rail services, while a low performance of air services is identified. For the other clusters, there is no reason to see a different interpretation of their classification. There-

Table 4.21: Classifications after implementing scenario 3

C#	Class	n	Airports
C5	**Main Intermodal Hubs**	4	*CDG, FRA, AMS, LHR*
C7	**Central International Hubs**	9	*VIE, BRU, ORY, CGN, DUS, MUC, MXP, ZRH, LGW*
C2	**Peripheral International Hubs**	10	*CPH, ATH, FCO, OSL, WAW, LIS, BCN, MAD, ARN, MAN*
C6	**Land-connected Minor Airports**	22	*BVA, LCY, SEN, CRL, LYS, BSL, SXF, HAM, NUE, LEJ, STR, TXL, HAJ, DTM, BGY, LIN, BLQ, EIN, GVA, BHX, LTN, STN*
C1	**Central Minor Airports**	35	*Remaining airports*
C3	**Peripheral Minor Airports**	22	*BIO, VLC, SVQ, CLJ, BMA, SOF, SKG, BRI, VNO, TGD, BGO, TRD, SVG, OPO, OTP, BEG, ALC, AGP, SCQ, GOT, MMX, ABZ*
C4	**Disconnected Airports**	11	*LCG, GRX, XRY, SDR, BDS, KUN, TRF, NYO, BLL, AAL, WMI*

fore, the other clustering labels remain equal for this scenario. The complete structure for all classes is shown in Table 4.21. Its visualization is shown on the map in Figure 4.24.

4.5. Interpretation of classifications

The selection of European airports is classified for the current benchmark situation, with reassessed classifications after implementing multiple scenarios. These classifications aim to explore the possible effects on an intermodal network and its European nodes. It is important to identify the main findings from the base case to understand these effects and differences between the benchmark situation and the different scenarios. The classification of this benchmark scenario is based on a combination of key takeaways from the data.

4.5.1. Identifying classification keywords

Firstly, three levels of air services can be divided into *main*, *international*, and *minor* airports. These levels explain the air side variables of a cluster. The main level is assigned to the airports with frequencies that are outliers in the data, with very high values. The label of international airports is assigned to airports with offered frequencies in the mid-high range of the data set. These clusters also show a relatively high rate of intercontinental flights, emphasizing their international character in the aviation market. The minor airports show an overall low number of air services and destinations served from the airport.

A correlation between the location and assignment to specific clusters can be found by studying the geographical locations of the clusters. Comparing these findings shows a strong correlation with the level of rail service. The keywords *central* and *peripheral* are used to identify the clusters' geographical location. The airports in the central clusters are self-evidently mostly situated in central Europe. These locations have multiple directions to serve by train, which explains the better scoring values on the rail variables. Peripheral airports are situated more to the edges of the European study area. The number of airports within an HSR range, *AR rail*, contributes to this division. This metric was calculated for all airports considered on the mainland of Europe and, therefore, neglects possible connections over the eastern border of Europe in the direction of Russia and Turkey. However, the study focuses on implementing integral policies, most likely to occur within the European Union/Schengen area. The differentiation between the central and peripheral clusters distinguishes airports with a high potential for good connections with a large-scale international network versus airports that need to be specifically connected to such networks. This can be used to identify airports that require extra attention if it is desired to include them in intermodal plans.

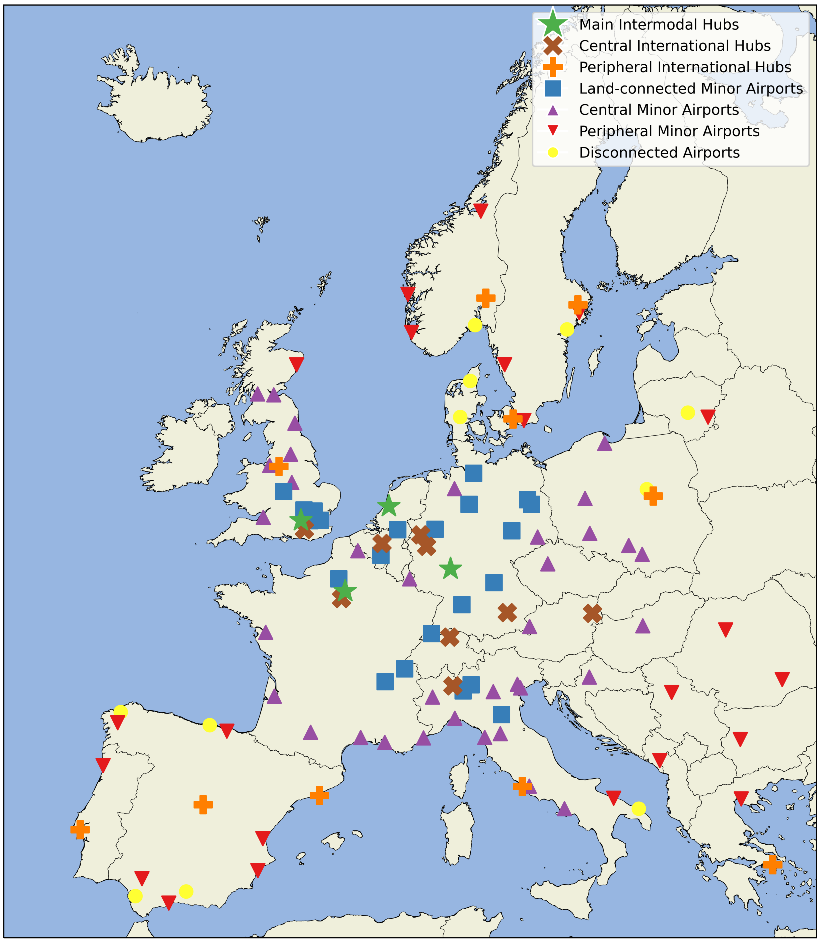


Figure 4.24: Scenario 3 in seven classes

Furthermore, the keyword *hubs* indicates a good connection between the air and rail side. These clusters perform highly on both the air and the rail side variables. Moreover, they are distinguished by the on-site availability of railway stations. For example, three out of four airports in the *Main Hubs*-cluster feature a railway station connected to HSR networks. Also, the travel time to the main railway network (SAC) is fairly low, indicating a good connection between the airport and FUA.

Combining these keywords enables the clusters to be qualitatively labeled according to their distinguishing characteristics. Only one cluster remains unlabeled when the classification is limited to these keywords. Cluster 4 is an extremely low-performing cluster, i.e. low number of services in rail and aviation, of airports with a very high travel time to the linked FUA. These could be classified as minor airports. However, due to the low level of service on all factors, the final classification for these low-impact airports is labeled *Disconnected Airports*. The qualitative labels assigned to the clusters give a good view of the role of an airport in the intermodal network. The classes that use the keyword *hubs* can be considered airports where it can realistically perform an intermodal trip.

Considering the study's scope, the interpretation of the classification is mainly based on the quantitative characteristics of the offered services and the geographical location. It can vary significantly which airports offer additional travel mode integration services, where long-distance train services are directly connected, or which airports counter infrastructural limitations. These questions can not be answered for all airports in this study. However, the classification helps identify airports for which it is useful to answer these questions.

4.6. Impact of policies on airport classifications

This section discusses the general impact of the policy implementations on the intermodal network. The interpretation of the distinguishing factors in classification helps the assessment of these scenarios. The labeling of the clusters is constant for most clusters through the changes between scenarios. As discussed in the individual classification of the clusters, the third scenario, that includes railway improvements, contains a more extensive cluster for the originally so-called *Secondary Airports*. However, due to this cluster growth, the term secondary airports, which relates to larger airports in the same FUA, does not apply to this scenario. Other airports with a similar level of service, which are the major airports of the FUA, join the cluster in this scenario as well. This removes the geographical determinant of the other nearby airport(s) and shifts the distinguishing factor mainly to the rail services in the area. Still, with a low level of air connections, this cluster is identified as a minor airport with a high level of connection with the rail network from its area. Therefore, it is labeled as *Land-connected Minor Airports*, remaining focused on the rail service but removing the determinant of being secondary airports in the area.

4.6.1. Large shifts in airport classification

With the interpretation of all other classes remaining constant through the implementation of the scenarios, the impact of the policies is mainly concentrated on individual cases. For each scenario, the individual cluster shifts are identified. These tables show that particularly Clusters 1 and 7 (*Central Minor Airports* and *Central International Hubs* respectively) tend to push off airports to other clusters. In scenario 1, where rail service is extended, most airports are moved to Cluster 6 (*Central Secondary Airports / Land-connected Minor Airports*). This effect is also seen in scenario 3, which includes the extension rail services. These airports have a relatively low value on the average shortest path and are well-connected to the railway network. Their addition to Cluster 6 improves the air-side score of this cluster, including an increase in the ratio of intercontinental flights.

The average frequency per unique intercontinental destination is calculated to assess the level of service per intercontinental destination. All scenarios are compared in Table 4.22 with the average total frequency for all international destinations. For each scenario, these values are compared to the highest-scoring cluster, the main hubs of Cluster 5. The percentages show that Scenarios 1 and 3 cause a large increase in the service of intercontinental destinations in Cluster 6. This shows that these airports have a high service level in the rail network and a high frequency on their intercontinental destinations. Although these airports serve fewer intercontinental destinations than the classified hubs, they can facilitate intermodal travel to their specific destinations. This extra analysis emphasizes the change in function of Cluster 6. In Scenarios 1 and 3, the airports in this class can be considered semi-hubs for a few destinations and, thus, facilitate intermodality.

In Scenario 2, the same occurs in Cluster 7. *Central Minor Airports* from Cluster 1 are shifting to Cluster 7, forming a class with a high ratio of intercontinental flights and high intensity on the intercontinental connections. In this scenario, the higher focus on longer-distance flights moves these airports from a *Minor Airport* to an *International Hub*, including them in the group of airports that can be considered suitable for an intermodal network. These Swiss airports Basel (BSL) and Geneva (GVA), the German Leipzig (LEJ) and Nürnberg (NUE), and Lyon (LYS) are all well situated in Europe as central airports to fulfill this role in the European network. Birmingham (BHX), which also moves to this cluster, is located less centrally but can still play a role on a lower level for Great Britain.

4.6.2. Individual changes in classification

Apart from the larger shifts of clusters, some airports shift individually. The airports of Warsaw (WAW), Stockholm-Bromma (BMA), and Malmö (MMX) all involve individual cluster shifts. Warsaw Airport changes from a *Central Minor Airport* to a *Peripheral International Hub* in all scenarios. The implementation of the policies showed an emphasis on specific airport characteristics in multiple cases. Also, the case of Warsaw emphasizes the geographical location. The airport seems to be a borderline case of being an international or minor airport, considering the shift from a minor airport in the central ge-

Table 4.22: Intercontinental flights per destination

	Benchmark		Scenario 1		Scenario 2		Scenario 3	
	IF / Dest	% of max	IF / Dest	% of max	IF / Dest	% of max	IF / Dest	% of max
C1	546	39%	690	45%	801	44%	1048	52%
C2	1067	**77%**	1447	**95%**	1387	77%	1796	**89%**
C3	844	61%	1305	85%	976	54%	1252	62%
C4	156	**11%**	282	**18%**	368	**20%**	526	**26%**
C5	1382	**100%**	1528	**100%**	1812	**100%**	2019	**100%**
C6	292	21%	1166	76%	462	25%	1556	77%
C7	1004	73%	1198	78%	1518	**84%**	1797	**89%**

*For each scenario the **best ratio**, **2nd best ratio**, and **worst ratio** are marked.*

ographic context and becoming a hub when considered peripheral. With the planning of CPK airport near Warsaw Centralny Port Komunikacyjny 2023, it is shown that Warsaw intends to fulfill a larger role as an intermodal hub in the future, and in that case, would be most likely assigned to an *International Hub* class when analyzed in a similar study.

The airports Stockholm-Bromma and Malmö both move from the *Central* to the *Peripheral Minor Airports*. The Swedish airports have peripheral characteristics in their number of airports within the HSR range and the rail network and frequencies. Both cases are affected by the stronger emphasis on the low scores by the scenario implementation. Considering the variable values for these airports, both are borderline cases between the central and peripheral clusters. However, considering intermodal possibilities, it is indifferent to these airports to which class they are assigned, as both indicate minor airports with a low potential for intermodal possibilities.

Meanwhile, Athens Airport is moved from a *Peripheral International Hub* to a *Peripheral Minor Airport* in Scenario 1. The extension of the rail services in this scenario strongly highlights Athens' meager score of its rail variables. The FUA of Athens only serves 28 weekly train services in the long-distance context. The air side variables of the airport score fairly high, with a total number of 180,000 annual flights. As an airport, Athens can be considered to play an internationally large role. In intermodality, however, the airport does not contribute nor connect to the European network and plays no significant role.

The individual examples show the difficulties of the simplifications done by clustering the airports en masse. Airports that appear to be borderline cases between different clusters can be categorized differently by changes in the data. For some airports, this is explainable, but for individual cases such as Athens, the differences in variable characteristics are so large that the simplification fails to describe the airport in a correct cluster.

4.7. Translating hub classification to effect on intermodality

Different factors distinguish the classifications for each scenario: the level of air services, their geographical location, and the level of rail services, including the connection between the airport and the main railway network. The qualitative labeling of these clusters helps identify the intermodal possibilities for the airports. To translate the classification of airports into their intermodal possibilities, the classes can be grouped on their role in an intermodal network. Table 4.23 shows the division of the clusters per group of intermodal possibilities. The classification keyword *hubs* indicates the airports that are ready for substitution, with high services on the air and rail side and a strong connection with the railway network. Cluster 6, which changes its label depending on the scenario's effects, shows potential for intermodality when classified as a *land-connected airport*. These airports have a medium-level service on the air side and lack a strong connection with the main railway network. However, the FUA they're situated in already has a strong connection to the railway network. The analysis of frequency per unique connection in Table 4.22 showed a fairly high performance of this class for the destinations the airports serve. If these airports gain a stronger connection with the rail network, for example, by constructing on-site railway stations, they can play a small role in an intermodal network. The other clusters have a service level that is too low to offer intermodal possibilities.

Table 4.23: Classes grouped by intermodal possibilities

High intermodality	Main Hubs (C5) Central International Hubs (C7) Peripheral International Hubs (C2)
Potential intermodality	Land-Connected Minor Airports (C6*)
No intermodality	Central Minor Airports (C1) Peripheral Minor Airports (C3) Disconnected Airports (C4) Central Secondary Airports (C6*)

Cluster 6 is classified differently for different scenarios. This influences its position in intermodal possibilities.

To assess the target group's of the intermodal classified airports, the airports are linked to the passenger numbers from the Eurostat flight data (European Commission 2022). The total number of passengers traveling through the airports classified in one of the high intermodality clusters gives an insight into the effect of an increased contribution of intermodality to the network. These numbers are shown for each scenario in Table 4.24. In scenario 3, creating the group of potential intermodality and shifting airports from the intermodality clusters to this group causes a problematic interpretation of the changes in enabled intermodal trips. The number of passengers traveling through an airport from the high intermodality group decreases in these scenarios. However, when the potential intermodality is included, an increase in enabled passengers is achieved. This effect is subject to the interpretation of the possible intermodality. To be able to interpret the changes in these passenger numbers specifically, an individual analysis of each airport's role should be performed. However, this analysis shows that when the airports in the potential intermodality group have met the conditions to classify as high intermodal airports, the scenarios cause an increase in intermodality throughout the entire European network. The extension of rail services can potentially increase the number of passengers who can travel intermodaly by 1.9%, while the reduction of short-haul flights could stimulate intermodality for 4.3% more passengers. The combination scenario increases this effect to 14.3%, with a large factor of uncertainty, by the classification of the potential intermodality.

Table 4.24: Total number of passengers per group of intermodality

	High intermodality	Potential intermodality	No intermodality	Change [%]
Benchmark	928	-	478	-
Scenario 1	946		460	+1,9%
Scenario 2	968	-	438	+4,3%
Scenario 3	857	204	345	-7,7% / +14,3%

Number of passengers in millions. Changes are given for high intermodality and total intermodality, including potential intermodality.

For all scenarios, a total increase of (potentially) enabled passengers is found. The large shifts discussed in subsection 4.6.1 are the main reason for this growth. These airports can therefore be identified as potential international hubs in case of policy implementation. The large increase in scenario 3 is particularly caused by the relatively few airports leaving the high-intermodality *Central International Hubs* and the large group of airports moving from *Central Minor Airports* to the potential intermodal group. Most airports moving to the *Land-connected Airports* in scenario 3 are moving to the *Central*

International Hubs in scenario 2, causing the increase of enabled passengers for this scenario.

Identifying the airports with the potential to classify as intermodal hubs indicates that focusing on these airports has the largest impact on the networks' intermodality. However, solely focusing on these airports could be counterproductive. The central situation of most airports that have the potential to become a hub also means that these airports could have several alternatives for travel. Partly discouraging short-haul flights would probably encourage people to opt for a different airport and still use the trusted modes of transport. Therefore, for these policies, an integral approach for the entire intermodal network is needed. Initiatives like the short-haul flight ban in France would contribute more when imposed by a larger collective, such as the European Union or the slot allocation authority.

On the other hand, the extension of rail services can focus first on the main connections of the identified airports. This scenario contains an encouraging approach, for which extra options in the particular airport will only improve its attractiveness en could potentially enable intermodality. In the same context, connecting *Minor Airports* between themselves would not contribute to the intermodal network significantly. When the railway network is improved or new rolling stock can be assigned to routes, it would be good to consider the access routes to the airports (potentially) classified as high intermodality airports.

5 Conclusion & Discussion

This chapter presents the conclusion of the network analysis performed in this study. The research aims to explore the role of airports in a European intermodal network and investigate the effects that possible policy scenarios have on this network. The study is purely focused on the network's perspective of intermodality. Therefore, integrated air/rail products and improvement of book-ability are recognized as important factors to a successful air/rail integration but are not considered during the network analysis. The conclusion to this study is obtained by answering the constructed research questions step-by-step to finally answer the main research question *"Where lies the potential of air/rail integration policies within a European intermodal network?"* The conclusion also considers the factors important to obtain a strong intermodal network but were outside the scope of the technical analysis.

5.1. Defining variables for intermodal assessment

This study started with the question: *"What variables can quantify the degree of air/rail intermodality at European airports?"* Answering this question helped build a basis to gather airport data and use this in a network analysis. From the literature, three main components stood out for successful integration: superior intermodal services in the network, physical facilities for a seamless transfer, and network support by integrated ticketing and extra integration services. Coupling these topics with the available data on European airports, the variables could be divided into air services, rail services, and network, and the physical characteristics of the integration.

The air services were defined mostly by the frequency of offered flights. Because the studied type of intermodality, focusing on transfer passengers, also valued intercontinental characteristics of the offered services, two variables in that context were added. Firstly, the total number of intercontinental flights and the ratio of this frequency compared to the total represent the airports' field of focus. Lastly, the number of destinations for each airport showed the variety of destinations each offers.

The rail side variables were obtained from extensive European international rail network research. This data set determines the number of intercity and long-distance trains as an important frequency-defining factor. Also, two network indicators are used to assess the rail side of the airport. The average shortest path in the European rail network is calculated for each airport. Transfer penalties are an important determinant in modal choice, which indicates the importance of traveling to, in this case, the airport as directly as possible. Finally, the node connectivity is assessed. This variable includes travel times, directness, and other travel-impeding factors.

The physical characteristics of the airport are based on the connection between the airport and the railway network. First, the availability of an on-site railway station is categorized as not available, conventional station and HSR stations. Also, the connection with the main rail network is assessed. As most services are OD-focused and originate from the main railway stations of the nearby cities, the travel time to this station is used to assess the connection. Furthermore, the geographical situation of the airport is considered. This location is linked to the potential distance for air/rail substitution. Therefore, the location is assessed by calculating the number of airports within a feasible HSR range. The variety of variables offers the possibility to assess the airports on different themes, enabling a classification with different combinations of these factors.

5.2. Network analysis

The analysis of the benchmark network showed possibilities to cluster the airports in seven clusters, allowing the network to be distinguished on a variety of factors. The classification of this benchmark network answers the sub-question: *”How can the airports be classified in the current European intermodal network?”* The clustering results clearly distinguish the different variables’ performance and the variable values’ combinations. The assessed variables are focused on air services, rail services, the rail network, and physical integration characteristics. The 7-cluster results show a strong influence from the geographical location of the airports on the cluster they are assigned to. In the classification of the airports, the keywords *central* and *peripheral* can describe the geographic situation of the clusters, while *main*, *international*, and *minor* can identify the level of air services that are offered at the airports. A combination of high air and rail service levels and a strong physical connection with their FUA could be classified as a *hub*. The classification also identified a cluster that can be considered disconnected from the main air and rail networks and served a small role in Europe despite the selection criteria for the airport selection before starting the analysis.

The benchmark network was analyzed to form a comparison for assessing future scenarios. Three constructed scenarios focus on different aspects of intermodality. These scenarios were carefully considered from governmental visions of transport networks and expert interviews with IenW. The construction of scenarios answers the question *”What European plans can affect the intermodal network and how?* The first scenario assesses improving the rail services offered by optimizing border corridor capacity and improving international and long-distance travel. The extension of the rail services involves relatively small improvements on the available infrastructure, and is mainly focused on improving the frequency of services, by acquiring more rolling stock and puting a stronger focus on the international and long-distance connections. The second scenario investigates imposing a ban on short-haul flights and stimulating air/rail substitution. This scenario increases the rate of intercontinental travel. The third scenario combines two previous scenarios to simulate an integral approach to air/rail intermodality policies.

The assessment of the different scenarios showed a fairly constant classification. The *Main Hubs* class, consisting of the four large European airports, was a constant factor through all scenarios, performing high on all defining variables. Also, in each scenario, almost all airports could be classified by combinations of the mentioned keywords considering the number of air services and the geographical location. One cluster has changed classification during the scenario classification. The *Central Secondary Airports*-class from the benchmark scenario grew to a larger cluster when combining the constructed scenarios, which changed the perspective on the cluster to *Land-connected Minor Airports*. Apart from the changes of entire classes, individual airports shift between the clusters in each scenario. This causes individual airports to shift from a *minor airport* to an *international hub*, which changes their role in intermodality, as was the case for Warsaw (WAW), for example. Also, a group of airports made the same shift, changing their clustering to a classification considered as an intermodal possibility.

As combinations of defining keywords could construct the classifications, the classes could also be grouped based on their value in intermodality. All classes considered hubs can be considered to have a high value in intermodality. The cluster that changed to *land-connected airports* is labeled as a potential value in intermodality for the scenarios it takes that classification, as the level of air services is relatively low. Still, it serves a small selection of destinations on a high level. The other minor airports are considered to have no real value in the intermodal network. The airports in each group are linked to their annual number of passengers to assess the impact of the airports on the total transport network, answering the question: *”What effect does change in classification have on the potential of intermodal travel in Europe?”*. If also the potential intermodal airports have an added value, the scenarios show increased enabled passengers for intermodality of 1.9%, 4.3%, and 14.3%, respectively.

5.3. Linking network analysis to transport reality

The numerical results and the consequential answers to the sub-research questions are deducted from a mathematical method. To answer the question, *"Where lies the potential of air/rail integration policies within a European intermodal network?"*, a link should be created between the mathematical model and reality. Translating these results to advice that can be used in reality, the limitations should be considered in determining to what extent the results can directly be used.

Firstly, the method starts with a statistical analysis (PCA) of the obtained data for the chosen variables. This treatment simplifies the information within the data and loses the ability to assign weights to specific variables considered more important than other variables. In this study, it could have been useful to assign a stronger weight to variables dependent on the currently available infrastructure to prioritize clustering based on the possibilities within the current infrastructure. On the other hand, the PCA was used to avoid arbitrarily assigning weights to different variables. If it is desired to assign weights to specific variables, a well-founded argumentation should be available to explain the chosen weights. The tool user could construct these weights to fit the desired outcomes better. Also, this can be assessed by prioritizing the used variables and only include the variables that have been found to be important factors from research specified to the desired use of the tool.

Following the statistical data treatment, the airports are clustered using AHC. This method is strongly number-based, grouping different airports without considering individual possibilities. Some specific airports deal with different reasons for a higher or lower value on the variable. For example, Schiphol cannot expand its services due to environmental reasons. Many other airports in Europe are still able to expand their services. Pinpointing these airport-specific characteristics is not possible using a number-based clustering technique. Also, the clustering results throughout the different scenarios showed that some airports are very sensitive to changes in the data. These borderline cases in the data are hard to put in the correct cluster because there's no possibility for individual assessments. Creating more clusters to catch these specific cases would not help this issue for fairly evenly distributed data. Firstly, it will create more clusters, thus more borders, thus more possibilities for borderline cases. Also, adding more clusters would reduce the clustering value by reducing the distinction between the clusters.

Overall, the method enables one to create insights into the current situation of the offered services in Europe. It can identify the strong outliers in the available data. However, the descriptive statistics of the data showed large groups of medium-performing airports for multiple variables. In these groups, it is hard to create a further distinction. To assess strategic decisions for these airports, more specific research should be performed, with a more individual approach, considering the possibilities of the individual airports. This individual approach can also help assess whether constructed scenarios are possible at all airports or clusters and how the data could be modified to create a new clustering.

Also, the data that is used can be specified further. For example, intercontinental connections are calculated very binary but do not differentiate between specific destinations. For some countries, specific connections can be economically or historically more important than others. An individual assessment for airports can help identify these important connections and distinguish airports on a lower level.

Despite the limitations in creating a more extensive differentiation between airports on a lower level, the method can capture the status of airports on a multimodal level. It can differentiate between large airports connected poorly or very well to the land-based rail network and smaller airports. Compared to the Chinese case study (Chen et al. 2022), this study equaled the importance of rail and air travel to assess the performance of both modes as equally contributing modes. The Chinese study mainly focused on the complementary effect of the availability of HSR stations in the catchment area. This European study also considered the rail network's available services and network performances. If this method were applied to the Chinese study, a more specific distinction could be created on the performance of the rail stations instead of assuming all HSR stations contribute equally to the rail network, independent of their size.

Assessing the limitations and strengths of the used methodology, a link can be made between the model and reality to answer the main research question: *"Where lies the potential of air/rail integration policies within a European intermodal network?"* To answer the question, three example scenarios are constructed. One of the scenarios reduced the short- and medium-haul flights, which resulted in a shift of several airports from the classification as a minor airport to international hubs:

- Birmingham (BHX)
- Lyon (LYS)
- Basel (BSL)
- Geneva (GVA)
- Nürnberg (NUE
- Leipzig (LEJ)
- Warsaw (WAW)

For this scenario, the answer would lie in these specific airports, as the implemented policy of short-haul flights changes the positions of these airports in the intermodal network. However, as discussed earlier, the individual cases of these airports should be considered to see whether the individual characteristics of these airports also allow the airports to facilitate that contribution to the network.

While the first scenario, where rail services are improved, has only changed the position of Warsaw (WAW) to a larger contribution, the third scenario changed the interpretation of the classification. Similar airports, as in the second scenario, leave the classification of *minor airports* behind but enter a classification with a potential contribution to intermodality. These airports are fairly small and should be considered individually to see their abilities to contribute to an intermodal network.

The general answer to the main research question is that the potential lies in the borderline cases of the clustering process. These airports are affected by the changes in the data set and, therefore, are worth further investigation, with a more individual approach, to assess the potential contribution to intermodality. Also, the airports considered as *Main* and *International Hubs* contribute to intermodality with their level of services. Therefore, these airports have great potential in the intermodal network when they consider offering additional services neglected in this study, such as integral ticketing, timetable synchronization, and luggage handling.

5.4. Discussing different intermodality perspectives

The implementation of the scenarios showed the difficulties of the balancing act occurring in making the transport sector more sustainable. Making transport modes greener is part of technical innovations, which always take some time to be adopted. The rail and aviation sectors are also relatively slow modes in terms of innovation. Air transport is heavily studying electrical and hydrogen-fueled aviation. However, the market comes across the massive distances that airplanes must travel, making the fossil fuel option still the most feasible for long distances. The railway sector depends on complicated infrastructures. Historically, the technical characteristics of railways differ per country, making cross-border train travel more complicated and limited to dedicated infrastructure.

These difficulties are considered in the construction of the scenarios. The rail service extension is considered low-cost, limiting the need for new infrastructure. This is the cheapest option and a more environmentally friendly start for improvement, given the doubts about the Life Cycle Analysis of train travel, including the used infrastructure. Maximizing the use of the cross-border and long-distance capacity couples the factor for intermodality of high frequencies with the feasibility of policy execution.

The airside scenario combines the call for banning short-haul flights, as performed in France lately, with the expectation of airports' reactions. As airports are assigned a specific number of slots they may use, airports are expected to react to short-haul flight removal by performing longer-distance flights instead. Short-haul flights are more polluting per kilometer due to the take-off procedures. In total, however, long-haul flights account for a much larger share of pollution in aviation. From the airports' perspective, it is understandable to fill the released slots with different OD pairs that are allowed. Also, the demand for mobility keeps increasing and can, in this manner, be supplied within the same number of flights. Therefore, cooperation between air and rail companies can be a compromise in increasing mobility and limiting the environmental impact.

To improve intermodality and increase the air/rail substitution rate, Europe has an ambitious vision of an integrated air/land network with several large corridors. However, the plans of the TEN-T originated in 1990. The full execution of the plan has not been realized yet, and recently, the European Council reviewed the plan and extended it. The aim is to finish the extended core network by 2040 and the comprehensive network by 2050. This means the current situation should lie exactly between the plans' origin and the full network's planned completion. These long-term infrastructural plans are commonly exposed to long delays and changing requirements. It is questionable if large-scale European projects can successfully construct such a network or if cutting it into smaller pieces is better. Both approaches have their benefits and disadvantages.

The large-scale TEN-T shows the difficulties of international governance. Infrastructural alignment is one example, but railway operations and airlines are subject to complicated governance. While Europe prefers a free market competition for railway operators, different countries still have different approaches to governance and the division of responsibilities for railway maintenance and operations. On the airside, governments have limited influence on the routes that are served by the airlines. Both complications are a limitation in successfully coordinating an international intermodal network. The number of stakeholders involved in such operations creates a loop of dependencies and misaligned interests. These differences should be resolved as much as possible to achieve an optimal intermodal network.

An important factor that is often mentioned in intermodality is the integration of tickets. This remains an important factor, independent of the policies applied from the network's perspective. Simplifying booking a ticket and taking the intermodal trip is essential to successfully integrating two travel modes. Ideally, obtaining a ticket for an intermodal trip would be as easy as finding the cheapest or quickest air journey from A to B. This requires cooperation between all railway operators in the network and the possibility to access timetables far in advance. Apart from the integrated ticketing, additional services can contribute even further. One example is handling luggage like a passenger transferring between two airplanes. Airports are fully equipped to handle this flow of baggage. However, handling baggage between trains and airplanes needs an extra link in this chain. Often, when an on-site station is available, this does not allow a connection between the train and the baggage handling area of the airport. Also, additional space is needed to load and unload baggage quickly and from the train. Another example, following Brexit, is the additional facilities for passport controls when boarding a train to the United Kingdom. This reduces the possibility for train services to access all desired airports. These difficulties limit the possibilities for such additional services.

If air/rail intermodality could become a successful alternative for long-distance traveling, it would positively impact the balance between increasing mobility and reducing GHG. However, the effect of the intermodality is relatively low, considering the largest share of air travelers are OD passengers. However, the introduction of intermodality could also positively affect direct air-rail substitution by improving services and familiarity with long-distance traveling by train. Overall, the improvement of intermodal possibilities has big potential. Compared to the direct substitution of OD travel, the effect is fairly small. However, it can also contribute to this direct substitution and create a custom to use the train for more long-distance journeys. Moreover, implementing stimulating policies can enable European people to travel intermodal, which is not possible yet. This could open a new mode of travel and experience for millions of people in Europe.

5.5. Recommendations

This study assesses the potential of intermodality from a network perspective. The data used forms a good basis for analyzing an intermodal network. However, in the data, there was some misalignment between the available data for airports and the railway data, which caused the exclusion of several airports and the artificial link between some airports and their FUA. Access to commercial railway databases and the air data of non-EU countries could contribute to a more complete data set. Furthermore, the railway data was obtained directly from research, which caused a lack of underlying data for some metrics. Therefore, the data did not facilitate a precise adaptation of the data, which limited the possibilities for scenario implementation. Also, considering the studies' scope, the link between the study area and the rest of the world can be modeled further. The rail connection between the eastern parts of Europe and areas in western Asia can change the position of the airports in that part of Europe.

This limitation also caused a shallow form of differentiation between airports, also caused by the size of the study's scope, containing a lot of European airports. The individual situation of each airport requires a smaller study to differentiate the implemented scenarios for each airport. The performed research assumed equal treatment for all airports and limited differentiation to applying scenarios to specific air and rail services categories. Personalizing these implementations would have been too time-consuming, given the scope of the research.

One of the implications of simplifying scenario implementation is seen in Scenario 2, where short-haul flights are substituted, and the released slots are distributed over the remaining airports. This redistribution could cause an artificially high number of long-distance flights for airports where most of the flights are served at a small distance. Further research into the expected replacement flights for short-haul flights would create a more reliable scenario.

The research into intermodality of air and rail travel can be further improved if there are more findings available on the influence of additional services, or 'soft attributes'. The air and rail modes are assessed mostly on the frequency of their services and some network properties. The integration between the two is mostly physically orientated because there's little scientific knowledge on the quantitative impact of, for example, integrated ticketing and luggage handling on the experienced integration that can be used in this quantitative analysis. Research on the transformation of qualitative services to quantitative scoring also helps assess the implementation of these important services in the analysis.

Finally, this research assumes the physical integration between the modes fairly linearly by the travel time to the city center. However, it is not considered what travel time to the main railway network is assumed to be sufficient, considering it to be an integration of the modes, even if additional services are included. Research on the willingness to travel to a train station can help categorize travel times and better assess the accessibility of the railway network.
Furthermore, the availability of (HSR) stations at the airport creates a three-step assessment. However, it would be interesting to see the relation between the availability of low-service on-site stations and high-service stations available within a particular travel time. This would result in a performance score for physical integration of the two modes, where a variety of solutions would be able to create a strong integration.

Bibliography

Beeravelli, Arun Datta et al. (2022). *Re-thinking Air-Rail Connec-tivity Joint Interdisciplinary Project*. Tech. rep. Delft Univesity of Technology, NACO.

Bergantino, Angela Stefania and Leonardo Madio (Mar. 2020). "Intermodal competition and substitution. HSR versus air transport: Understanding the socio-economic determinants of modal choice". In: *Research in Transportation Economics* 79, p. 100823. ISSN: 0739-8859. DOI: 10.1016/J.RETREC.2020.100823.

Brida, Juan Gabriel et al. (2017). "Air and HST Multimodal Products. A Segmentation Analysis for Policy Makers". In: DOI: 10.1007/s11067-017-9352-3.

Bruno, Francesco (2022). *The Connectivity of the Long-distance Rail and Air Transport Networks in Europe*. URL: http://resolver.tudelft.nl/uuid:d2322bfd-248e-411c-8e75-c8ef295de9a5.

Cangelosi, Richard and Alain Goriely (Jan. 2007). "Component retention in principal component analysis with application to cDNA microarray data". In: *Biology Direct* 2. ISSN: 17456150. DOI: 10.1186/1745-6150-2-2.

Centralny Port Komunikacyjny (2023). *Centralny Port Komunikacyjny*. URL: https://www.cpk.pl/en/.

Chen, Enhui, Amanda Stathopoulos, and Yu (Marco) Nie (Nov. 2022). "Transfer station choice in a multimodal transit system: An empirical study". In: *Transportation Research Part A: Policy and Practice* 165, pp. 337–355. ISSN: 09658564. DOI: 10.1016/j.tra.2022.09.014.

Chen, Yuting et al. (2022). "Airport classification in Chinese multi-airport regions: An interaction network perspective between aviation and high-speed rail". In: *European Journal of Transport and Infrastructure Research* 22.2, pp. 1–21. ISSN: 15677141. DOI: 10.18757/ejtir.2022.22.2.6043. URL: https://journals.open.tudelft.nl/ejtir/article/view/6043/5264.

Deutschlandtakt (2023). *Deutschlandtakt*. URL: https://www.deutschlandtakt.de.

Dobruszkes, Frédéric, Giulio Mattioli, and Laurette Mathieu (Oct. 2022). "Banning super short-haul flights: Environmental evidence or political turbulence?" In: *Journal of Transport Geography* 104. ISSN: 09666923. DOI: 10.1016/j.jtrangeo.2022.103457.

Donners, Barth, Elmer van Buuren, and Rémy Rijniers (2018). "Steden verbonden, trein of vliegtuig, effecten van substitutie". In: *Colloquium Vervoersplanologisch Speurwerk*.

Donners, Barth and Martijn Heufke Kantelaar (Nov. 2019). *Emissies van korte afstandsvluchten op Nederlandse luchthavens*. Tech. rep. Royal HaskoningDHV.

Duarte Costa, Joana and João António de Abreu Silva (2012). *Factors of air-rail passenger intermodality*. Tech. rep.

Dutta Baruah, Indraneel (Oct. 2020). *Cheat sheet for implementing 7 methods for selecting the optimal number of clusters in Python*. URL: https://towardsdatascience.com/cheat-sheet-to-implementing-7-methods-for-selecting-optimal-number-of-clusters-in-python-898241e1d6ad#:~:text=As%20highlighted%20by%20other%20cluster,the%20agglomerative%20hierarchical%20as%20well..

Equipo Barcelona (June 2021). *Más viajeros, más trenes, menos emisiones*. Tech. rep. Barcelona.

Eurocontrol (2004). *CARE II: The airport of the future: Central link of intermodal transport? - WP 1 - Review of the current intermodality situation*. Tech. rep.

Euronews (Dec. 2022). *It's official: France bans short-haul domestic flights in favour of train travel*. URL: https://www.euronews.com/green/2022/12/02/is-france-banning-private-jets-everything-we-know-from-a-week-of-green-transport-proposals.

European Commission (2019). *The European Green Deal*. Brussels. URL: https://eur-lex.europa.eu/resource.html?uri=cellar:b828d165-1c22-11ea-8c1f-01aa75ed71a1.0002.02/DOC_1&format=PDF.

— (2021a). *Action Plan to boost passenger rail*. URL: https://transport.ec.europa.eu/news/action-plan-boost-passenger-rail-2021-12-14_en.

— (2021b). *Core network corridors on the TEN-T*. URL: https://ec.europa.eu/transport/infrastructure/tentec/tentec-portal/site/en/maps.html.

European Commission (2022). *Eurostat Data Browser*. URL: https://ec.europa.eu/eurostat/databrowser/view/DEMO_R_PJANGRP3__custom_2790186/bookmark/table?lang=en&bookmarkId=168107c0-c56f-4adb-85b4-381126e7944b.
European Environment Agency (2022). *Transport and environment report 2022*. ISBN: 9789294805195.
European Parliament (Dec. 2021). "Union guidelines for the development of the trans-European transport network and repealing Decision No 661/2010/EU". In: *Official Journal of the European Union*. ISSN: 1315/2013.
Frankfurt Airport (2019). *Frankfurt Airport Air Traffic Statistics 2019*. Tech. rep. Fraport.
Givoni, Moshe and David Banister (Sept. 2006). "Airline and railway integration". In: *Transport Policy* 13.5, pp. 386–397. ISSN: 0967070X. DOI: 10.1016/J.TRANPOL.2006.02.001.
Grimme, Wolfgang and Martin Jung (July 2018). "Towards more sustainability? – The development of aviation emissions from Germany between 1995 and 2016". In: *Proceedings of the 22nd Air Transport Research Society World Conference*. URL: https://elib.dlr.de/121383/.
Groupe ADP (Jan. 2020). *Traffic at Paris Aéroport up by 2.5% in 2019, at 108 million passengers*. Tech. rep. Groupe ADP.
IATA (Jan. 2020). *Strengthening Cooperation on Standards for Intermodal Travel*. URL: https://www.iata.org/en/pressroom/pressroom-archive/2020-press-releases/2020-01-30-01/.
ICAO (2022). *Air traffic recovery is fast-approaching pre-pandemic levels*. URL: https://www.icao.int/Newsroom/Pages/Air-traffic-recovery-is-fastapproaching-prepandemic-levels.aspx#:~:text=ICAO%27s%20latest%20analyses%20reveal%20that,of%20its%202019%20high%20point..
Jernbanedirektoratet (Jan. 2020). *Jernbanedirektoratets perspektivanalyse*. Tech. rep. Oslo: Jernbanedirektoratet.
Jolliffe, I.T. (2002). *Principal Component Analysis*. 2nd ed., pp. 112–133.
Kroes, Eric and Fons Savelberg (Apr. 2019). "Substitution from Air to High-Speed Rail: The Case of Amsterdam Airport". In: *Transportation Research Record: Journal of the Transportation Research Board* 2673.5, pp. 166–174. ISSN: 21694052. DOI: 10.1177/0361198119839952. URL: https://journals.sagepub.com/doi/10.1177/0361198119839952.
Maćkiewicz, Andrzej and Waldemar Ratajczak (Mar. 1993). "Principal components analysis (PCA)". In: *Computers & Geosciences* 19.3, pp. 303–342. ISSN: 0098-3004. DOI: 10.1016/0098-3004(93)90090-R.
Maertens, Sven and Wolfgang Grimme (June 2015). *How to assess the percentage of transfer passengers at airports?* Tech. rep. Cologne: Institute for Air Transport and Airport Research.
Maklin, Cory (Dec. 2018). *Hierarchical Agglomerative Clustering Algorithm Example In Python*. URL: https://towardsdatascience.com/machine-learning-algorithms-part-12-hierarchical-agglomerative-clustering-example-in-python-1e18e0075019.
Ministerie van Infrastructuur en Waterstaat (Nov. 2020). *Luchtvaartnota 2020-2050*. Tech. rep. Ministerie van Infrastructuur en Waterstaat.
Ministerie van Infrastructuur en Waterstaat et al. (Nov. 2020). *Actieagenda Trein en Luchtvaart*. Tech. rep.
NOS (Sept. 2023). *Ruzie over krimp Schiphol loopt hoog op, VS dreigt met sancties*. URL: https://nos.nl/artikel/2488777-ruzie-over-krimp-schiphol-loopt-hoog-op-vs-dreigt-met-sancties.
Priemus, Hugo and Wil Zonneveld (2003). "What are corridors and what are the issues? Introduction to special issue: The governance of corridors". In: *Journal of Transport Geography* 11.3, pp. 167–177. ISSN: 09666923. DOI: 10.1016/S0966-6923(03)00028-0.
ProRail (2021). *Integrale Mobiliteitsanalyse 2021*. Tech. rep. URL: www.prorail.nl.
Raad voor de Leefomgeving en Infrastructuur (July 2020). *Verzet de wissel: naar beter internationaal reizigersvervoer per trein.* Tech. rep. URL: www.rli.nl.
Railtech (Dec. 2020). *New European mobility strategy aims to triple high-speed rail traffic*. URL: https://www.railtech.com/policy/2020/12/10/new-european-mobility-strategy-aims-to-triple-high-speed-rail-traffic/.
— (Apr. 2022). *Romania to study first high-speed railway and revamp rail infrastructure*. URL: https://www.railtech.com/infrastructure/2022/04/28/romania-plans-to-study-first-high-speed-railway-and-revamps-rail-infrastructure/.

Reiter, Vreni, Augusto Voltes-Dorta, and Pere Suau-Sanchez (Dec. 2022). "The substitution of short-haul flights with rail services in German air travel markets: A quantitative analysis". In: *Case Studies on Transport Policy* 10.4, pp. 2025–2043. ISSN: 22136258. DOI: 10.1016/J.CSTP.2022.09.001.

Resource Systems Group, Inc. et al. (2015). *Integrating aviation and passenger rail planning*. ISBN: 9780309308342.

Ryerson, Megan S and Hyun Kim (2013). "Integrating airline operational practices into passenger airline hub definition". In: *Journal of Transport Geography* 31, pp. 84–93. ISSN: 0966-6923. DOI: https://doi.org/10.1016/j.jtrangeo.2013.05.013. URL: https://www.sciencedirect.com/science/article/pii/S0966692313001099.

Savelberg, Fons and Maarten De Lange (2018). *Substitutie mogelijkheden van luchtvaart naar spoor*. Tech. rep. Kennisinstituut voor Mobiliteitsbeleid.

Schiphol (Dec. 2019). *Maandelijkse Verkeer & Vervoer cijfers december 2019*. Tech. rep. Amsterdam: Royal Schiphol Group.

Song, Fangqing, Stephane Hess, and Thijs Dekker (June 2018). "Accounting for the impact of variety-seeking: Theory and application to HSR-air intermodality in China". In: *Journal of Air Transport Management* 69, pp. 99–111. ISSN: 0969-6997. DOI: 10.1016/J.JAIRTRAMAN.2018.02.008.

Sperry, B.R. et al. (2012). *Intercity passenger rail access to airports*. 2300, pp. 22–30. DOI: 10.3141/2300-03.

Statista (Dec. 2020). *Share of passengers transferring and terminating at Heathrow Airport in the United Kingdom (UK) between 2002 and 2019*. URL: https://www.statista.com/statistics/303920/flight-transfers-passengers-terminating-at-heathrow-airport-uk/.

Statista Research Department (June 2023). *Number of scheduled passengers boarded by the global airline industry from 2004 to 2022*. URL: https://www.statista.com/statistics/564717/airline-industry-passenger-traffic-globally/.

The European Rail Research Advisory Council (2011). *Rail route 2050: The Sustainable Backbone of the Single European Transport Area*. Tech. rep. URL: www.errac.org.

Wang, Yiqiao et al. (Oct. 2020). "Travel time analysis in the Chinese coupled aviation and high-speed rail network". In: *Chaos, Solitons & Fractals* 139, p. 109973. ISSN: 0960-0779. DOI: 10.1016/J.CHAOS.2020.109973.

Witlox, Frank et al. (Dec. 2022). "Changing tracks: identifying and tackling bottlenecks in European rail passenger transport". In: *European Transport Research Review* 14.1. ISSN: 18668887. DOI: 10.1186/s12544-022-00530-9.

Wortel, Monique van and Mariska Kaptein (Jan. 2023). *Air-Rail initiatives at national and at European level*. Tech. rep. Ministry of Infrastructure and Water Management.

Wu, Shuping and Dan Han (Dec. 2022). "Accessibility of high-speed rail (HSR) stations and HSR–air competition: Evidence from China". In: *Transportation Research Part A: Policy and Practice* 166, pp. 262–284. ISSN: 09658564. DOI: 10.1016/J.TRA.2022.10.015.

Yuan, Zhiyi, Changgui Dong, and Xunmin Ou (Jan. 2023). "The substitution effect of high-speed rail on civil aviation in China". In: *Energy* 263, p. 125913. ISSN: 0360-5442. DOI: 10.1016/J.ENERGY.2022.125913.

Zakaria Jaadi (Mar. 2023). *A Step-by-Step Explanation of Principal Component Analysis (PCA)*. URL: AStep-by-StepExplanationofPrincipalComponentAnalysis(PCA).

Zhang, Fangni, Daniel J. Graham, and Mark Siu Chun Wong (Dec. 2018). "Quantifying the substitutability and complementarity between high-speed rail and air transport". In: *Transportation Research Part A: Policy and Practice* 118, pp. 191–215. ISSN: 09658564. DOI: 10.1016/J.TRA.2018.08.004.

Zhang, Q., H. Yang, and Q. Wang (2017). "Impact of high-speed rail on China's Big Three airlines". In: *Transportation Research Part A: Policy and Practice* 98, pp. 77–85. DOI: 10.1016/j.tra.2017.02.005.

A

Airport selection

Table A.1: Overview of all airports in the airport selection

Airport name	**IATA Code**	**ICAO Code**	**Municipality**	**Country**
Salzburg Airport	SZG	LOWS	Salzburg	Austria
Bilbao Airport	BIO	LEBB	Bilbao	Spain
A Coru?a Airport	LCG	LECO	Culleredo	Spain
Federico Garcia Lorca Airport	GRX	LEGR	Granada	Spain
Jerez Airport	XRY	LEJR	Jerez de la Forntera	Spain
Valencia Airport	VLC	LEVC	Valencia	Spain
Santander Airport	SDR	LEXJ	Santander	Spain
Sevilla Airport	SVQ	LEZL	Sevilla	Spain
Montpellier-M?diterran?e Airport	MPL	LFMT	Montpellier/ M?diterran?e	France
Paris Beauvais Till? Airport	BVA	LFOB	Beauvais/Till?	France
Lille-Lesquin Airport	LIL	LFQQ	Lille/Lesquin	France
Nantes Atlantique Airport	NTE	LFRS	Nantes	France
Brindisi ? Salento Airport	BDS	LIBR	Brindisi	Italy
Peretola Airport	FLR	LIRQ	Firenze	Italy
Kaunas International Airport	KUN	EYKA	Kaunas	Lithuania
Sandefjord Airport, Torp	TRF	ENTO	Torp	Norway
Cluj-Napoca International Airport	CLJ	LRCL	Cluj-Napoca	Romania
Stockholm Skavsta Airport	NYO	ESKN	Stockholm / Nyk?ping	Sweden
Stockholm-Bromma Airport	BMA	ESSB	Stockholm	Sweden
London City Airport	LCY	EGLC	London	United Kingdom
Southend Airport	SEN	EGMC	Southend	United Kingdom
Vienna International Airport	VIE	LOWW	Vienna	Austria
Brussels Airport	BRU	EBBR	Brussels	Belgium
Brussels South Charleroi Airport	CRL	EBCI	Brussels	Belgium
Sofia Airport	SOF	LBSF	Sofia	Bulgaria
Zagreb Airport	ZAG	LDZA	Zagreb	Croatia

Table A.1 continued from previous page

Airport name	**IATA Code**	**ICAO Code**	**Municipality**	**Country**
V?clav Havel Airport Prague	PRG	LKPR	Prague	Czech Republic
Billund Airport	BLL	EKBI	Billund	Denmark
Copenhagen Kastrup Airport	CPH	EKCH	Copenhagen	Denmark
Aalborg Airport	AAL	EKYT	Aalborg	Denmark
Bordeaux-M?rignac Airport	BOD	LFBD	Bordeaux/M?rignac	France
Toulouse-Blagnac Airport	TLS	LFBO	Toulouse/Blagnac	France
Lyon Saint-Exup?ry Airport	LYS	LFLL	Lyon	France
Marseille Provence Airport	MRS	LFML	Marseille	France
Nice-C?te d'Azur Airport	NCE	LFMN	Nice	France
Charles de Gaulle International Airport	CDG	LFPG	Paris	France
Paris-Orly Airport	ORY	LFPO	Paris	France
EuroAirport Basel-Mulhouse-Freiburg Airport	BSL	LFSB	B?le/Mulhouse	France
Berlin-Sch?nefeld Airport	SXF	EDDB	Berlin	Germany
Dresden Airport	DRS	EDDC	Dresden	Germany
Frankfurt am Main Airport	FRA	EDDF	Frankfurt am Main	Germany
Hamburg Airport	HAM	EDDH	Hamburg	Germany
Cologne Bonn Airport	CGN	EDDK	Cologne	Germany
D?sseldorf Airport	DUS	EDDL	D?sseldorf	Germany
Munich Airport	MUC	EDDM	Munich	Germany
Nuremberg Airport	NUE	EDDN	Nuremberg	Germany
Leipzig/Halle Airport	LEJ	EDDP	Leipzig	Germany
Stuttgart Airport	STR	EDDS	Stuttgart	Germany
Berlin-Tegel Airport	TXL	EDDT	Berlin	Germany
Hannover Airport	HAJ	EDDV	Hannover	Germany
Bremen Airport	BRE	EDDW	Bremen	Germany
Dortmund Airport	DTM	EDLW	Dortmund	Germany
Eleftherios Venizelos International Airport	ATH	LGAV	Athens	Greece
Thessaloniki Macedonia International Airport	SKG	LGTS	Thessaloniki	Greece
Budapest Liszt Ferenc International Airport	BUD	LHBP	Budapest	Hungary
Bari Karol Wojty?a Airport	BRI	LIBD	Bari	Italy
Malpensa International Airport	MXP	LIMC	Milan	Italy
Il Caravaggio International Airport	BGY	LIME	Bergamo	Italy
Turin Airport	TRN	LIMF	Torino	Italy
Genoa Cristoforo Colombo Airport	GOA	LIMJ	Genova	Italy
Milano Linate Airport	LIN	LIML	Milan	Italy
Bologna Guglielmo Marconi Airport	BLQ	LIPE	Bologna	Italy
Treviso-Sant'Angelo Airport	TSF	LIPH	Treviso	Italy

Table A.1 continued from previous page

Airport name	IATA Code	ICAO Code	Municipality	Country
Verona Villafranca Airport	VRN	LIPX	Verona	Italy
Venice Marco Polo Airport	VCE	LIPZ	Venice	Italy
Ciampino?G. B. Pastine International Airport	CIA	LIRA	Rome	Italy
Leonardo da Vinci?Fiumicino Airport	FCO	LIRF	Rome	Italy
Naples International Airport	NAP	LIRN	N?poli	Italy
Pisa International Airport	PSA	LIRP	Pisa	Italy
Vilnius International Airport	VNO	EYVI	Vilnius	Lithuania
Luxembourg-Findel International Airport	LUX	ELLX	Luxembourg	Luxembourg
Podgorica Airport	TGD	LYPG	Podgorica	Montenegro
Amsterdam Airport Schiphol	AMS	EHAM	Amsterdam	Netherlands
Eindhoven Airport	EIN	EHEH	Eindhoven	Netherlands
Bergen Airport Flesland	BGO	ENBR	Bergen	Norway
Oslo Gardermoen Airport	OSL	ENGM	Oslo	Norway
Trondheim Airport V?rnes	TRD	ENVA	Trondheim	Norway
Stavanger Airport Sola	SVG	ENZV	Stavanger	Norway
Gda?sk Lech Wa?sa Airport	GDN	EPGD	Gda?sk	Poland
Krak?w John Paul II International Airport	KRK	EPKK	Krak?w	Poland
Katowice International Airport	KTW	EPKT	Katowice	Poland
Modlin Airport	WMI	EPMO	Warsaw	Poland
Pozna?-?awica Airport	POZ	EPPO	Pozna?	Poland
Warsaw Chopin Airport	WAW	EPWA	Warsaw	Poland
Copernicus Wroc?aw Airport	WRO	EPWR	Wroc?aw	Poland
Francisco de S? Carneiro Airport	OPO	LPPR	Porto	Portugal
Humberto Delgado Airport (Lisbon Portela Airport)	LIS	LPPT	Lisbon	Portugal
Henri Coand? International Airport	OTP	LROP	Bucharest	Romania
Belgrade Nikola Tesla Airport	BEG	LYBE	Belgrade	Serbia
Alicante International Airport	ALC	LEAL	Alicante	Spain
Barcelona International Airport	BCN	LEBL	Barcelona	Spain
Adolfo Su?rez Madrid?Barajas Airport	MAD	LEMD	Madrid	Spain
M?laga Airport	AGP	LEMG	M?laga	Spain
Santiago de Compostela Airport	SCQ	LEST	Santiago de Compostela	Spain
Gothenburg-Landvetter Airport	GOT	ESGG	Gothenburg	Sweden
Malm? Sturup Airport	MMX	ESMS	Malm?	Sweden

Table A.1 continued from previous page

Airport name	**IATA Code**	**ICAO Code**	**Municipality**	**Country**
Stockholm-Arlanda Airport	ARN	ESSA	Stockholm	Sweden
Geneva Cointrin International Airport	GVA	LSGG	Geneva	Switzerland
Z?rich Airport	ZRH	LSZH	Zurich	Switzerland
Birmingham International Airport	BHX	EGBB	Birmingham	United Kingdom
Manchester Airport	MAN	EGCC	Manchester	United Kingdom
Bristol Airport	BRS	EGGD	Bristol	United Kingdom
Liverpool John Lennon Airport	LPL	EGGP	Liverpool	United Kingdom
London Luton Airport	LTN	EGGW	London	United Kingdom
London Gatwick Airport	LGW	EGKK	London	United Kingdom
London Heathrow Airport	LHR	EGLL	London	United Kingdom
Leeds Bradford Airport	LBA	EGNM	Leeds	United Kingdom
Newcastle Airport	NCL	EGNT	Newcastle	United Kingdom
East Midlands Airport	EMA	EGNX	Nottingham	United Kingdom
Aberdeen Dyce Airport	ABZ	EGPD	Aberdeen	United Kingdom
Glasgow International Airport	GLA	EGPF	Glasgow	United Kingdom
Edinburgh Airport	EDI	EGPH	Edinburgh	United Kingdom
London Stansted Airport	STN	EGSS	London	United Kingdom

C

Scientific Paper

European air/rail intermodality network potential: a classification of hub integration

D. Nijhof, O. Cats, A. Bombelli, W.W. Veeneman, B.J.H.F. Donners
Department of Transport, Infrastructure and Logistics, Delft University of Technology, Delft, The Netherlands

***Abstract*—Following the European Green Deal, transport-related emissions must be reduced by 90% in 2050. While aviation contributes to a significant share of these emissions, the substitution of flights for rail travel is often considered to achieve a reduction in emissions of greenhouse gasses. Mostly, travelers consider this option when the full trip can be substituted directly. However, possibilities also lie in intermodal travel, where specific journey parts are substituted. To promote this way of travel, the two modes should have a strong level of service, and there should be a significant level of integration between the two modes.**

In this paper, a method is presented to assess the position of airports' in a European intermodal network. This tool helps identify airports with the potential to facilitate intermodal travel in Europe. The service levels in both the air and the rail network, the physical integration between the modes, and the geographical location of airports are considered in different variables to assess the network. A Principle Component Analysis identifies the shared information between the different variables. The Principle Components are then clustered using an Agglomerative Hierarchical Clustering. With this clustering, airports are grouped based on their characteristics of intermodal factors, and different roles in an intermodal network can be identified. Comparing the descriptive statistics of each cluster helps assign qualitative labels to the cluster to classify the airports' roles. After creating the benchmark with data from the current network situation, scenarios are constructed to assess the effects of implementing different policies regarding intermodality. Extending rail services and/or reducing short-haul flights are policies formed into scenarios in which the underlying data is modified. The change of clustering and classifications caused by the data modification helps identify potentially interesting airports in the network.

The results show that 20 to 30 of the 113 considered airports show a substantial role in intermodality. Also, the findings identify airports that change their position due to the scenario implementations and offer the potential for a significant role in an intermodal network. Finally, comparing the classifications in the different scenarios shows that the constructed scenarios can enable up to 14.3% extra passengers to travel intermodal.

***Keywords*—Airport classification, Intermodality, Air/Rail integration, Hubs, Principle Component Analysis, Agglomerative Hierarchical Clustering, Europe**

I. Introduction

By the aim of reducing transport-related greenhouse gasses emissions, the substitution of flights for rail is an important topic in governmental policies. Apart from direct substitution, also partial substitution can be achieved in multi-legged flights. By integrating air and rail, an intermodal network can be constructed, to achieve the intermodality between the two modes and to promote the partial substitution of flights. This paper aims to identify the role of European airports in an intermodal network, and to show the airports that have potential to play an important role in intermodality. This is done by a clustering analysis, based on multiple variables considering air and rail services and the physical integration between the two modes.

The paper is structured as follows. Section II shows the available literature in this domain, and defines intermodality. Also, an example of clustering involving air and rail travel is introduced. The methods of the classification process are elaborated in section III, including data collection, statistical treatment through a Principle Component Analysis (PCA) and the Agglomerative Hierarchical Clustering (AHC). A Benchmark scenario is constructed in section IV, using data from 2019. In section V, scenarios of different policy implementations are constructed. Based on these scenarios the data is modified and the airports are reclustered and reclassified. The conclusions from this study are presented in section VI.

II. Literature Review

Total passenger transport demand in Europe has increased by 20% between 2000 and 2019. The demand for air traffic specifically has been globally increasing [1] and even increased by 86% in Europe between 2000 and 2019 [2] until the COVID-19 pandemic and recovered quickly towards pre-pandemic levels in 2022 [3]. While on one side (air) mobility demand is increasing, on the other hand, many different climate agreements are made to reduce the impact of travel on greenhouse gas (GHG) emissions. The Paris Agreement by the United Nations in 2015 has led to the long-term focused European Green Deal [4], agreeing to reduce GHG emissions by 55% in 2030 compared to 1990. Transportation has a significant role in GHG emissions, and its share even relatively increased by 33% from 2000 to 2019 to 29% of all GHG emissions in Europe [2]. To reduce the transport sector's impact, The European Green Deal also states that transport-related emissions should be reduced by 90% in 2050.

A. Substitution of air travel

Aviation accounts for 13,9% of total GHG emissions in European transport and is (after road transport with 43.9%) the second largest GHG emitter. As aviation is one of the largest emitters of GHG, it is relevant to investigate what rail transport in Europe can do compared to air travel. Rail transport is widely recognized as a sustainable travel mode and an important topic for national governments and the European Union regarding mobility improvement [5]. Much research has been conducted on the substitution of air travel for rail transport, which shows potential in a range from 5% up to 75% of flight reduction [6][7][8][9][10]. These results strongly depend on the choice of scenarios, parameters, and timeline but show that rail transport can be considered a good option to substitute flights for rail travel.

Most research considers direct substitution, focused on origin-destination (OD) travel. Generally, these direct substitutions cover short- and medium-haul flights with a competitivity range of up to 5 hours of train travel time [11]. A large share of these competitive flights are performed within a range of 750 kilometers, in Schiphol Amsterdam's case approximately 25% [12], and are a feasible option for direct substitution. However, long-haul flights have no feasible option for direct substitution (yet). At Schiphol, 19% of the flights in 2019 were performed on long-haul flights, accounting for 29% of the total number of passengers in that year [13]. Furthermore, 36% of Schiphol's passengers are transferring to a new flight. In that case, whether the incoming and outgoing flights are short-, medium- or long-haul flights is unknown. However, a significant share of Schiphol's passengers is transferring. Also, as most long-haul flights from Europe are intercontinental flights, many passengers cannot substitute their full journey with a different mode.

Apart from Schiphol, Frankfurt (54% [14]), Paris Charles de Gaulle (23%; [15]) and London Heathrow (34%; [16]) are known for their share of transfer passengers, based on pre-COVID-19 data from 2019. Passengers who connect in a hub airport generally transfer between two flight legs. Some have an origin and destination that allow them to substitute their journey for a complete rail trip. However, some have an origin or destination not on the European mainland or the UK, connected to the mainland by rail. When one of the legs of their journey is performed on the mainland or in the UK, it would be possible to perform a partial substitution. This however, should rather be achieved by improving intermodal possibilities instead of simply banning short-haul flights. When one of the legs of this trip isn't available because of regulations banning short-haul flights, that leaves a traveler two options. Firstly, smart ticketing platforms can find other travel options via plane via a hub further from the origin or destination. As a result, none of the legs is considered a short-haul flight and thus banned. An example would be a flight from München to JFK, New York, via Frankfurt. Suppose the short-haul leg from München to Frankfurt is restricted. In that case, travelers will still have the option to fly via London Heathrow, Amsterdam Schiphol, or Paris Charles de Gaulle, depending on the threshold distance for the short-haul flight ban. The second option is to improve the attractiveness of other, more sustainable travel options. If intermodal options are promoted, this could fill the gap that originated from the short-haul flight ban. To achieve this, intermodal products should be improved and easily accessible for travelers to be a feasible option.

B. *Intermodality*

The European Commission has defined intermodality as the characteristic of a transport system that integrates at least two different transport modes [17]. The access and egress legs of a journey are excluded from this concept because intermodality is focused on the main legs of a journey. In a partly intercontinental multi-leg journey, the European part could be substituted and thus become an intermodal trip. To do so, airports should focus on their role as intermodal hubs, accommodating air and rail travel and transfer possibilities that are as seamless as possible. Some research explores important factors to improve this intermodal hub function and vary from coordination of services (frequencies) of air and rail to the improvement of ticketing and baggage handling conveniences [18][19][10][9]. They also recognize that the same factors apply to choices for intermodal travel as to the substitution of air for rail travel; travel times, frequencies, and connecting times form an important basis for mode choice. These factors can be divided into two themes. One focused on services, from ticketing and baggage conveniences to the offered air and rail services and their frequencies. The second theme focuses on physical characteristics, which mainly determine travel times, e.g., high-speed rail (HSR) infrastructure availability and the geographic situation in the air and rail networks. These variables can be assessed for each (significant) airport in Europe to create insight into the characteristics of these airports.

Intermodality can be applied to various travel modes. This research assesses the intermodality between air and rail transport in Europe. This section elaborates on the specific intermodality for these travel modes and the factors that influence the intermodality of a network from a transport node's perspective. There are various interpretations of intermodality in literature [20][21]. This study focuses on integrating the airport into an extensive ground network, as a basis for defining an intermodal network, and reviews intermodality to substitute short-haul flights by improving air/rail intermodality possibilities. Such intermodality would mostly apply to passengers who normally travel by a multi-leg journey.

There are three main components for the successful integration of air and rail transport [22]:

1) There are superior intermodal services from the network perspective.
2) Physical facilities are available for a seamless transfer.
3) The network is supported by integrated ticketing and extra service provisions.

In several governmental recommendations, the improvement of international rail transport, as an important side of the intermodal travel, is discussed, resulting in different important aspects [23]. These aspects include infrastructure, traffic regulation systems, recommendations for train operators, and improvements in ticketing and passenger rights. This should be achieved and integrally executed by better international coordination of the international railway network. The network assessment considers the factors determining physical and transport services, leaving the additional services outside the research scope. Therefore, the defining factors are based on components one and two from the list of defined components.

To offer a superior intermodal service from the network perspective, both sides of the intermodality should provide high-quality services. One of the most critical factors for multi-modal accessibility is the frequency of services [24]. From a node's perspective, the total frequencies of air and rail travel influence this accessibility. On the airside, there's a difference between substitutable and non-substitutable flights. Most multi-legged journeys contain a long-distance, intercontinental leg unsuitable for substitution. An airport offering a relatively high number of these flights will likely function as an air hub. To cover these characteristics, the frequency of intercontinental flights and the ratio of these intercontinental flights compared to the total number of flights are considered to be important factors as well. Another aspect of mode choice is the offer of a variety of services. When there are more options, it is more likely that a traveler will use the node. This characteristic can be covered by the number of destinations an airport serves.

The service level of the rail side determines the other side of intermodality. In this case, the frequency of services is also one of the most critical factors. In addition, as the rail legs compete with the more common air services, the rail service should be convenient. Travel convenience is based on different aspects but is mainly determined by comfort, travel time, waiting time, and number of transfers. The number of transfers can be reviewed from the network perspective, looking at the shortest paths between nodes. Other travel (in)convenience factors can be included in a characteristic like travel impedance or generalized travel costs.

The interaction between air and rail travel can be reviewed based on the potential competition and substitution per link in OD matrices. Research evaluated these potentials for European Functional Urban Areas [25]. The number of links suitable for competition or substitution originating from a specific area shows the potential of flight substitution from and to this airport. The second component for successful integration considers the physical integration between modes. Examples, such as Frankfurt and Schiphol Amsterdam, show the convenience offered by on-site railway stations. The type of services provided at a station at the airport also indicates the integration into a ground network. An airport station can be connected to an urban metro-like network to access the nearby urban area or directly integrated into the national network. There's also a difference on the national level whether this network offers regular national services or long-distance and high-speed services. The connection to the ground network can also be assessed by the travel time to the main station in the related urban area. This parameter contains a mixture of the physical distance to the urban area and the service level of the connection. With low travel times to a central station, the integration between the airport and the main rail network can be considered higher.

Finally, the geographical location can be assessed to assess the potential of an airport without considering existing rail infrastructure. Because water bodies, mountain ranges, or other natural barriers can influence travel time over a simple geodesic distance for ground travel, the geographic location can be modeled by car travel times between two locations. The road network is the most advanced ground-level network and can be a good measure of travel distances.

C. Airport role in an intermodal network

As every airport has a certain role in the European air network, they would also have a role in a European intermodal network. In air networks, the concept of hubs is already widely known. Most airports function within a hub and spoke network for mostly economic reasons. Some research shows that airports within a network can be clustered and classified based on the airports' data [26]. Also, a link to rail travel can be made considering its complementary effect, by the availability of HSR stations in the catchment area of airports [27]. The results of this study show an example for implementation on the European study (Figure 1). Towards an even more elaborated intermodality classification, a wider variety of variables could be added to the clustering, as discussed earlier. Applying this to a European case, the classification of the airports' intermodality could identify the role of specific airports and their potential in an intermodal European network.

When the improvement of the intermodal network in Europe can lead to a reduction in short-haul flights, this can be reviewed from different perspectives. At first, reducing the relatively more polluting short-haul flights (per kilometer) by rail will directly reduce GHG emissions because rail transport is a more sustainable alternative than air travel. However, when short-haul flights are canceled, this will free the so-called slots at the airport and enables an airport to assign these slots to flights over a longer distance [28]. Considering the mobility perspective, this will improve overall mobility because (more) destinations can be reached on a higher frequency. Whether this mobility improvement will satisfy the existing demand for these destinations or enable a latent demand by increasing travel options is questionable [29]. Meanwhile, environmentally considered, these long(er)-haul flights have a higher load on GHG emissions [30], and thus will cause a higher total GHG emission in aviation. When short-haul flights are substituted, a good balance between improving total mobility and reducing GHG emissions by the

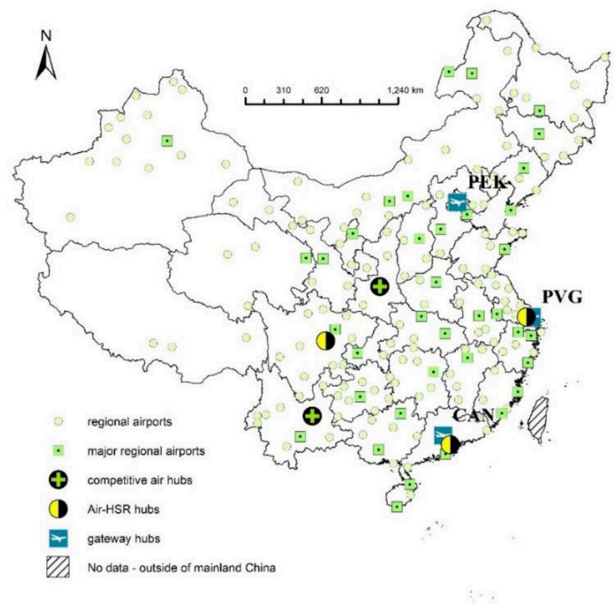


Fig. 1. Classification example for Chinese air/rail interaction [27].

transport sector should be found. With the Paris Agreement 2015, governmental institutions have a hard deadline to achieve the set emission goals. Still, they the balance they have to find between environmental and mobility perspectives is interesting to consider when interpreting the expected impact of intermodality encouragement.

III. Methodology

This analysis in this study aims to identify the location of the intermodal potential and focuses on the position of the individual airports in the larger network. A method is constructed to transfer the factors of intermodality into insights into the European intermodal network. To interpret the role of airports in the network, a shift from quantitative network data towards a classification of airports' positions in the network should be made. The total model is built by several steps of data treatment and sequent modeling techniques, shown in Figure 2 and listed as follows:

1) Data treatment & calculations
2) Principal Component Analysis
3) Agglomerative Hierarchical Clustering
4) Classification
5) Scenario implementation
6) Intermodal Possibilities Evaluation

The air traffic data is retrieved from the Eurostat Browser [31]. Because the airport's role in the intermodal network is assessed, only airports connected to the rail network are included in the data set. Before collecting data additional for the clustering, selection criteria are constructed to limit the size of the data set. The selection is evaluated on annual passenger data taken from the pre-COVID year of 2019 to narrow down the final selection. Two thresholds are introduced to only include airports of significance in this

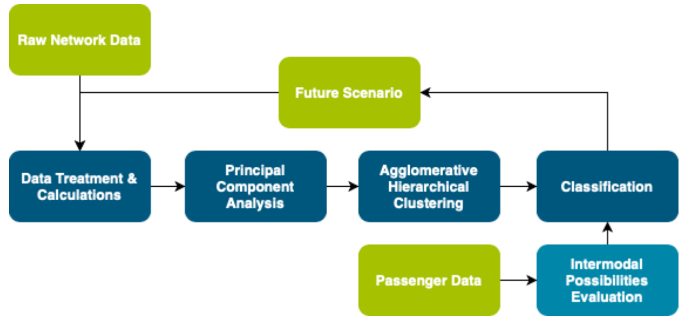


Fig. 2. Modeling structure

study. The first is based on the total number of passengers. Therefore the annual passenger numbers for each connection are summed for each origin airport, for which a lower bound of one million annual passengers is set. Also, every individual connection is reviewed on the number of passengers. A reasonable boundary for substitution would be when all passengers almost fill one large high-speed train. A Eurostar, or a double Thalys, train has a capacity of 750 people. With a load factor of 90%, a daily service would transport approximately 250,000 people. Therefore, an airport will only be excluded if at least one connection transports over 250,000 people.

For the preliminary selection of airports, rail data is collected from a database containing rail information for 125 Functional Urban Areas (FUAs) [25]. These areas partly overlap with the airports from this study's pre-selection. To connect the airports to the rail data a manual allocation of airports to the urban areas is performed. A set of 22 airports could not directly be linked to any of the FUAs, of which some are excluded from the data set, and some are linked to another FUA. The final selection of airports contains 113 airports in Europe that are accessible by land.

A. *Used variables*

From the definition of defining factors in network intermodality, the frequency and type of services, the geographical location, and the physical integration of travel modes in a node are the main aspects of mode integration. The combination of these factors supports the chosen variables used in this study. The variables are divided in air services, rail services and physical indicators.

Air services The first variable on air services is calculated by the total number of flights departing at an airport (eq. 1). This represents the total frequency at the airport,

$$f_i^{air} = \sum_j n_{ij}^{flights} \quad \forall \quad i \in S^O, j \in S^D \qquad (1)$$

where:

S^O = Set of origin airports in Europe,
S^D = List of all unique destinations from origin airports in Europe,
$n_{ij}^{flights}$ = Annual number of flights from airport i to airport j.

Secondly, the annual number of intercontinental flights (eq. 2) and the ratio compared to the total number of flights (eq. 3) is calculated to show the connection with intercontinental destinations and the focus area of the airport,

$$f_i^{IC\,air} = \sum_j n_{ij}^{flights} \quad \forall \quad i \in S^O, j \in S^{IC} \tag{2}$$

$$r_i^{IC\,air} = \frac{f_i^{IC\,air}}{f_i^{air}} \quad \forall \quad i \in S^O \tag{3}$$

where:

S^{IC} = List of all unique intercontinental destinations from origin airports in Europe.

Finally, the number of unique destinations with a significant load is considered (eq. 4). This number represents the variety of services that are offered at the airport. A threshold of three weekly flights is assumed to filter the regular connections,

$$n_i^{SC} = \sum_{j \in S^D} X_{ij} \quad \forall \quad i \in S^O \tag{4}$$

where:

n_i^{SC} = Number of significant connections,
$X_{ij} = \begin{cases} 1, & \text{if } f_{ij} \geq 3 \text{ per week} \\ 0, & \text{otherwise.} \end{cases}$

Rail services The rail services variables consider the rail network from the study by Bruno [25]. The first variable represents the number of rail services departing at an FUA (eq. 5),

$$f_i^{rail} = \sum_{j \in S^{FUA}} n_{kj}^{rail} \quad \forall \quad i \in S^O \tag{5}$$

where:

S^{FUA} = Set of all 125 FUAs in the observed rail network,
n_{ij}^{rail} = Weekly number of rail services from FUA k to FUA j,
k = FUA k assigned to airport i.

Because the rail leg should replace a more common air leg, the ease of using the rail should be considered. A main determinant is the number of transfers need for a trip. Therefore, the average shortest path for all FUAs is considered (eq. 6). Also, the node connectivity as calculated by Bruno [25], is considered to assess the travel convenience including multiple determining factors. The node connectivity is calculated as the sum of the link connectivity on all links from one origin (eq. 7). The link connectivity represents an indicator that combines the effective frequency on the link with the travel impedance experienced (eq. 8),

$$SP_i = \frac{\sum_{j \in S^{FUA}} sp_{kj}}{N^{FUA}} \quad \forall \quad i \in S^O \tag{6}$$

where:

S^{FUA} = Set of all 125 FUAs in the observed rail network,
N^{FUA} = Number of FUAs in set S_{FUA},
sp_{kj} = Shortest path between FUA k and FUA j,
k = FUA k assigned to airport i.

$$NC_i = \frac{\sum_{j \in S^{FUA}} LC_{ij}}{N^{FUA}} \quad \forall \quad i \in S^O \tag{7}$$

$$LC_{ij} = \frac{\theta^f f_{ij}}{\theta^{IV} tt_{ij}^{IV} + \theta^{OV}(tt_{ij}^A + tt_{ij}^W + tt_{ij}^E) + \theta^T tt_{ij}^T} \quad \forall \quad i, j \in S^{FUA} \tag{8}$$

where:

LC_{ij} = Link connectivity between FUAs i and j,
θ = Weighing factor for time and frequency numbers,
tt_{ij}^x = Travel time for activity x, between FUAs i and j,
IV = In-vehicle,
OV = Out-of-vehicle,
A = Access,
W = Waiting,
E = Egress,
T = Transfer.

The physical integration between is assessed by three different variables. The first distinguishes the availability of a station at the airports (eq. 9). This is a linearly dummy-coded variable, to score the availability from 0 (no station) to 2 (HSR). Also, given that no airport in Europe is connected to all HSR connections of its FUA, the public transport travel time between the airport and the central railway station of the airport's related city is still important. This is calculated using the Google Maps Directions API, calculating the route by public transport between the airport and main railway station of the city. This represents the distance to the main railway network. Lastly, the physical location of the airport is considered. This is represented by the number of airports that can be found within a competitive range. In a range between 4 and 6 hours, high-speed rail is considered competitive with air transport [32]. For each airport the Open Source Routing Machine (OSRM) API is used to count the number of airports in the dataset that is situated within this competitive range.

$$SA_i = \begin{cases} 0, & \text{if no railway station is available} \\ 1, & \text{if a conventional railway station is available} \\ 2, & \text{if a high-speed railway station is available} \end{cases} \quad \forall \quad i \in S^O \tag{9}$$

B. PCA, AHC, and Interpretation

A PCA is performed to the data set to avoid the need of an arbitrary weighting of performance variables, such as the β-parameters in utility function. PCA is a method of multivariate statistics to reduce the dimensions of all observations while preserving as much information as possible [33] and finding the variability between, in this case, the airports. First, the data is standardized, to remove the difference in magnitudes of the variables, and treat all variables equally. Then, the variables are transformed in a number of principle components (PC). The optimal number can be determined by different methods, based on the eigenvalues and explained variances of the PCs.

Using the components resulting from the PCA, the AHC can be performed. A matrix is filled where each cell represents the increase in cluster variance when the clusters i ans j are merged, based on the squared distance between these clusters. This method within AHC is called *ward-linkage*, and minimizes the effects of noise between clusters, during the process. From the matrix, the lowest increase in variance is picked, to merge the accessory clusters. The process repeats until all data points are merged to one cluster. Afterwards, the optimal number of clusters is determined by the Silhouette Score [34] and the constructed dendrogram.

The results from the clustering are interpreted based on the descriptive statistics and geographical situation of each cluster. The interpretation helps to translate the quantitative results to qualitative results, and to construct a classification. Using the classifications, the roles of the airports can be identified. Also, the total number of passengers throughout Europe that travel via an airport with significant intermodal possibilities can be calculated to assess the potential of intermodal travel in Europe.

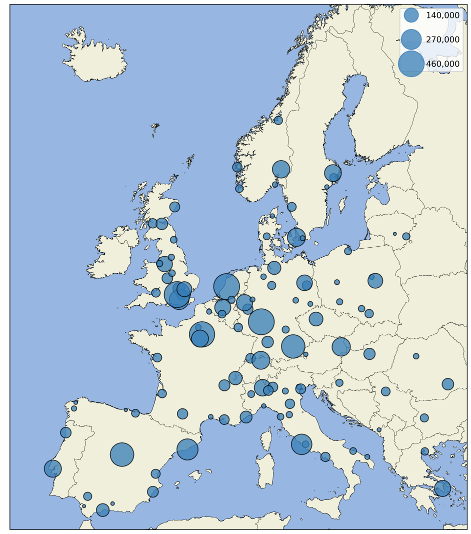


Fig. 3. Visualization of total air frequencies

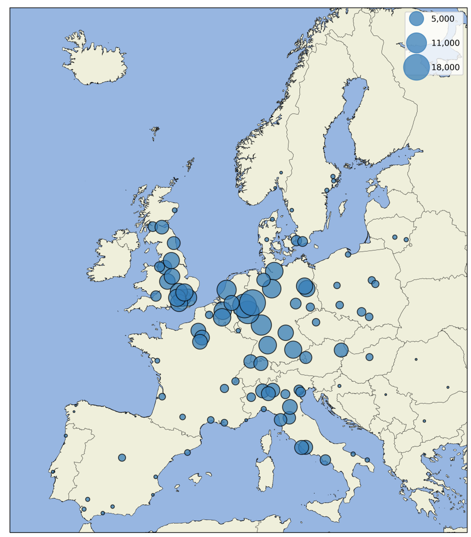


Fig. 4. Visualization of total rail frequencies

IV. Benchmark scenario

For the selection of airports, the variables are calculated based on the original data set and calculations in section III-A. Figures 3 and 4 show the geographical distribution of two important variables: total air and rail frequency. This identifies the larger airports, and the aiports with strong rail connections.

After performing the PCA, resulting in three unique PCs, and the AHC, the airports are divided into 7 clusters. The variable statistics for each cluster are shown in Table I. The corresponding clusters are plotted in Figure 5.

The clusters are classified to assess the meaning of the different clusters found and to see the effect of the changed clusters on the role of some airports. This classification adds a label to each cluster, showing the clusters' qualitative characteristics and assessing each cluster's added value more qualitatively compared to the quantitative analysis. Roughly four categories of clusters are shown, combined by high or low performance on air and rail services. Other variables, such as the station availability or the accessibility of the rail network, subsequently split these categories further.

The constant performance of the core four airports is notable in cluster 5. The already noted airports beforehand, London Heathrow, Paris Charles de Gaulle, Amsterdam Schiphol Airport, and Frankfurt Airport, are clustered with high-scoring values on all aspects. These core airports are labeled as *Main Intermodal Hubs*.

Clusters 2 and 7 also perform fairly well regarding the offered frequency of air services. They are distinguished by their offered services on the rail side. Cluster 2 has an average to low number of rail services. On the other hand, Cluster 2 does perform well on the availability of a railway station and its accessibility to the main railway network. The airports are relatively large, but, due to their location on the edge of Europe, they lack centrality, thus connection to the European rail network. Therefore, it can be classified as *Peripheral*

TABLE I
MEAN VARIABLE VALUES FOR 7 CLUSTERS

	TF air	IF air	IR air	NC air	AR rail	STA	SAC	TF rail	SP rail	CN rail	n
1	48.456	4.913	0,09	40	23	**0,24**	33	925	2,86	0,30	49
2	**237.474**	**39.747**	0,16	**100**	10	**1,33**	**31**	617	3,61	0,16	9
3	51.800	5.908	0,11	38	**7**	0,40	34	**102**	**4,02**	**0,04**	20
4	**13.775**	**469**	**0,03**	**19**	**7**	0,30	**96**	157	3,56	0,06	10
5	**452.354**	**135.437**	**0,30**	**181**	**32**	**1,75**	**26**	**2852**	**2,20**	**0,79**	4
6	43.983	2.624	0,06	46	**30**	0,33	71	**2827**	2,36	**0,81**	6
7	168.961	29.103	**0,17**	80	**30**	1,13	34	2676	**2,25**	0,69	15

For each variable the ***best value****,* ***2nd best value****, and* ***worst value*** *are marked.*

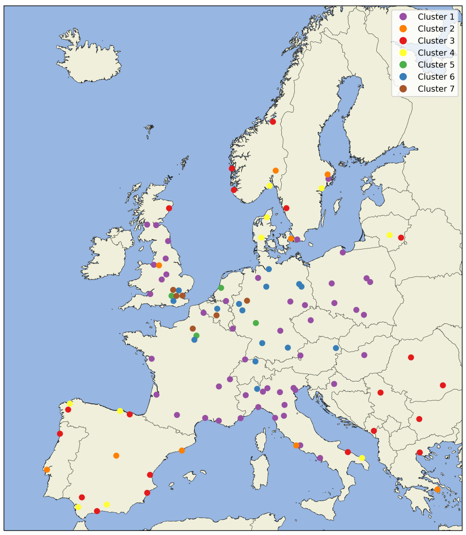


Fig. 5. Clustering of benchmark scenario

TABLE II
CLASSIFICATIONS OF THE BENCHMARK NETWORK

C#	Class	n	Airports
C5	**Main Intermodal Hubs**	4	*CDG, FRA, AMS, LHR*
C7	**Central International Hubs**	15	*VIE, BRU, ORY, SXF, HAM, CGN, DUS, MUC, STR, TXL, HAJ, MXP, ZRH, LGW, STN*
C2	**Peripheral International Hubs**	9	*CPH, ATH, FCO, OSL, LIS, BCN, MAD, ARN, MAN*
C6	**Central Secondary Airports**	6	*BVA, LCY, SEN, CRL, DTM, LTN*
C1	**Central Minor Airports**	49	*Remaining airports*
C3	**Peripheral Minor Airports**	20	*BIO, VLC, SVQ, CLJ, SOF, SKG, BRI, VNO, TGD, BGO, TRD, SVG, OPO, OTP, BEG, ALC, AGP, SCQ, GOT, ABZ*
C4	**Disconnected Airports**	10	*LCG, GRX, XRY, SDR, BDS, KUN, TRF, NYO, BLL, AAL*

International Hubs.

Cluster 7 contains airports with similar air services but a higher rail transport service. This cluster is the central counterpart of Cluster 2 and, therefore, can be classified as *Central International Hubs*.

The remaining clusters consist of airports with a lower level of air services. The cluster contains secondary airports such as London Luton, with low air services, but situated near large areas. These airports can be classified as *Central Secondary Airports*.

Cluster 3 and Cluster 4 are characterized by an extremely low level of rail services, because all airports are situated on the edges of Europe. The distinctive factor between these two clusters lies in the station accessibility, i.e., the connection to the railway network. Cluster 4's travel time to the studied railway network is extremely large. Combined with low services within every aspect of this study, they're classified as *Disconnected Airports*. Cluster 3 contains small airports on the edge of the European network. They can be classified as *Peripheral Minor Airports*.

Lastly, Cluster 1 is the largest calculated cluster. This cluster contains airports with low levels of services but a very central location in Europe. Also, with the low availability of on-site railway stations and low air and rail services, these airports can be classified as *Central Minor Airports*.

An overview of the classification and the airports contained in each class is shown in II. The airports are indicated by their IATA codes. The classification is also visualized in the map in 6 to see the geographical distribution of each class.

V. SCENARIO IMPLEMENTATION

To identify the potential of the European intermodal network, it is interesting to assess the effects that possible policies can have on the airports and their role in the network. To construct scenarios with a balance between having a significant impact, being affordable, and being a possible direction of decisions that will be made, examples of visions for European international travel are used. These scenarios are translated to a change in the data used for the benchmark situation. The adapted data can be used in the constructed analysis tool to construct new clustering for each scenario and assess the contribution to the network intermodality for each cluster and its airports by the mean values of the stated variables of intermodality.

Three different scenarios are constructed to assess potential effects of policy on the classification of airports in the intermodal network:

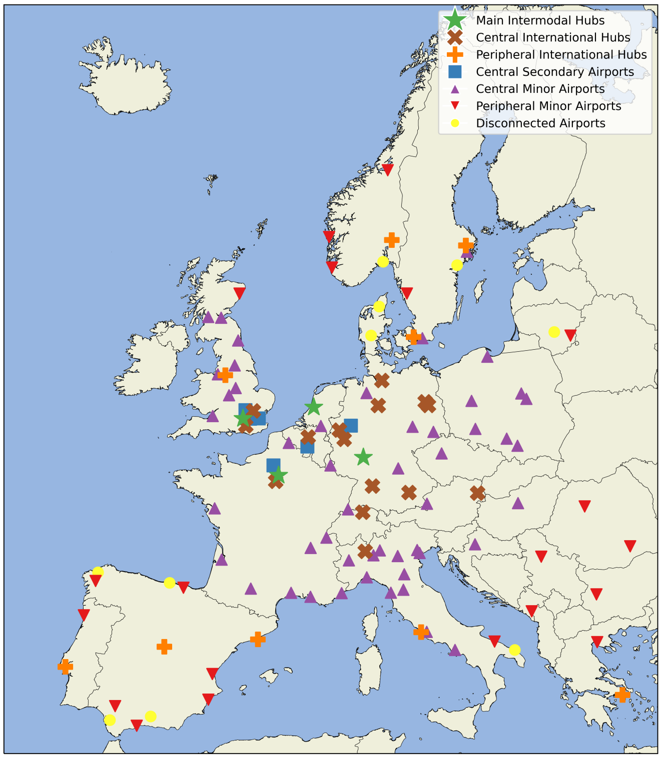


Fig. 6. Benchmark network in seven classes

1) Extension of rail services
2) Short-haul flight reduction
3) Combination of policies

A. *Scenario 1: Extension of rail services*

The first scenario involves the expansion of rail services, summarized in Figure 7. The existing track infrastructure is used, while safety and communication systems are improved to facilitate increased rail services, based on the Dutch Integral Mobility Analysis (IMA) [35], and extrapolated for Europe, after considering several visions from European countires [36][37][38][39][40][41][42]. This will show a relatively higher focus on international and longer-distance train travel than national short-distance trains.

Fig. 7. Scenario 1: Extension of rail services

From the IMA, two different multipliers are concluded. Intercity or medium-distance services are being quadrupled, while international long-distance services will double. These can be applied per range on the OD frequency matrix for rail services used for the benchmark scenario. The IC services are assumed to be from 1.5 to 3.5 hours (e.g. Amsterdam - Brussels), while the international services have a travel time of over 3.5 hours (Amsterdam - Paris, and further).

After implementing the data modifications, and re-executing the method's steps, some minor differences in clustering are found. The difference in data is presented in Table III, while the affected airports are identified in Table IV. This table indicates that a small shift has occurred from clusters 1 and 3 to cluster 2. These airports (Warsaw and Aberdeen) both are airports from a large group that only experienced an increase in rail frequencies of 100%. The small shift is also presented in the map in Figure 8.

TABLE III
DIFFERENCES HIGHLIGHTED FOR CHANGED CLUSTERS BETWEEN BENCHMARK AND SCENARIO 1

	TF air	IF air	IR air	NC air	AR air	STA	SAC	TF rail	SP rail	CN rail	n
B1	48.456	4.913	0,09	40	23	**0,24**	33	925	2,86	0,30	49
S1	46.331	4.575	0,09	39	23	**0,23**	**33**	2.922	2,86	0,30	48
B2	**237.474**	**39.747**	0,16	**100**	10	**1,33**	**31**	617	3,61	0,16	9
S2	**214.264**	**37.381**	**0,19**	**91**	10	**1,18**	**33**	1.745	3,51	0,16	11
B3	51.800	5.908	0,11	38	**7**	0,40	34	**102**	**4,02**	**0,04**	20
S3	50.884	4.517	0,09	39	**7**	0,42	34	**207**	**4,04**	**0,03**	19

TABLE IV
MOVED AIRPORTS IN SCENARIO 1

	Added	Removed
Cluster 1		*WAW*
Cluster 2	*WAW, ABZ*	
Cluster 3		*ABZ*
Cluster 4		
Cluster 5		
Cluster 6		
Cluster 7		

B. *Scenario 2: Short-haul flight reduction*

The scenario consists of actions that consider flight reduction and the reaction of airports and airlines, shown in Figure 9. At first, the example of the super-short-haul flight ban from France [43] is extrapolated throughout Europe. All flights below 500 kilometers will be banned in this scenario. The assumption in this scenario is made as if the European Union holds the power to ban international flights. Secondly, short- to medium-haul flights are heavily reduced by rail possibilities. Flights between airports within 500-1000 kilometer distance are reduced by 50%. It is assumed that improved railway services cause a shift in mode choice, but no ban on these flights is performed, which maintains half of the flights in this range. Lastly, the slots released by reducing short- and medium-haul flights are filled by offering higher frequencies on the remaining flights. It is assumed that the airports will provide the same ratio of medium- and long-haul continental flights compared to intercontinental flights.

Implementing this scenario in the data, and rerunning the classification process has affected all airside variables except the total frequency. Cluster 1 has become smaller after the implementation, while Cluster 7 has grown. Table V shows that the scenario mainly caused a redistribution of several

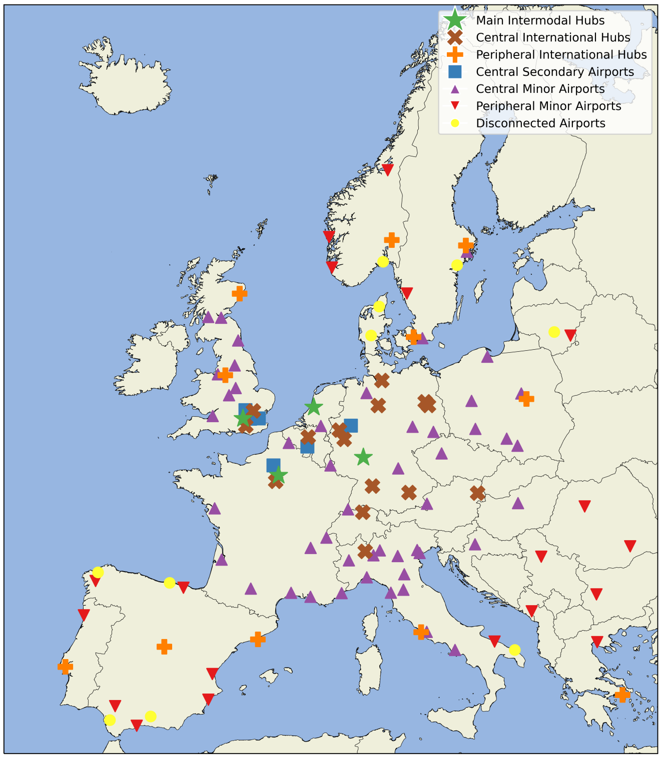


Fig. 8. Scenario 1 in seven classes

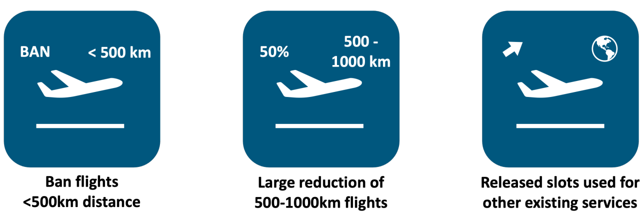


Fig. 9. Scenario 2: Shift in slot allocation

airports from Cluster 1. Clusters 4 and 5, also in this scenario, remain unchanged. Despite the several changes in airports' clustering, no difference in cluster interpretation is found. The differences in the values of the variables are highlighted in Table VI, for the clusters that have changed composition. The changed classification is visualized in Figure 10.

TABLE V
MOVED AIRPORTS IN SCENARIO 2

	Added	Removed
Cluster 1		*BSL, BHX, GVA, LEJ, BMA, LYS, WAW, MMX, NUE*
Cluster 2	*WAW*	
Cluster 3	*BMA, MMX*	
Cluster 4		
Cluster 5		
Cluster 6	*STN*	
Cluster 7	*BSL, BHX, GVA, LEJ, LYS, NUE*	*STN*

TABLE VI
DIFFERENCES HIGHLIGHTED FOR CHANGED CLUSTERS BETWEEN BENCHMARK AND SCENARIO 2

	TF air	IF air	IR air	NC air	AR air	STA	SAC	TF rail	SP rail	CN rail	n
B1	48.456	4.913	0,09	40	23	**0,24**	33	925	2,86	0,30	49
S1	43.946	6.407	0,15	36	22	**0,15**	**32**	894	2,89	0,30	40
B2	**237.474**	**39.747**	0,16	**100**	10	**1,33**	**31**	617	3,61	0,16	9
S2	**228.771**	**52.704**	0,23	**89**	10	**1,30**	**32**	627	3,52	0,17	10
B3	51.800	5.908	0,11	38	**7**	0,40	34	**102**	**4,02**	**0,04**	20
S3	49.821	6.829	0,14	33	**7**	0,36	35	**139**	**3,93**	**0,04**	22
B6	43.983	2.624	0,06	46	**30**	0,33	71	**2.827**	2,36	**0,81**	6
S6	59.501	5.083	0,07	55	30	0,43	70	**2.770**	2,39	**0,83**	7
B7	168.961	29.103	**0,17**	80	**30**	1,13	34	2.676	**2,25**	0,69	15
S7	139.295	37.951	**0,30**	63	**31**	1,05	**32**	2.276	**2,34**	0,59	20

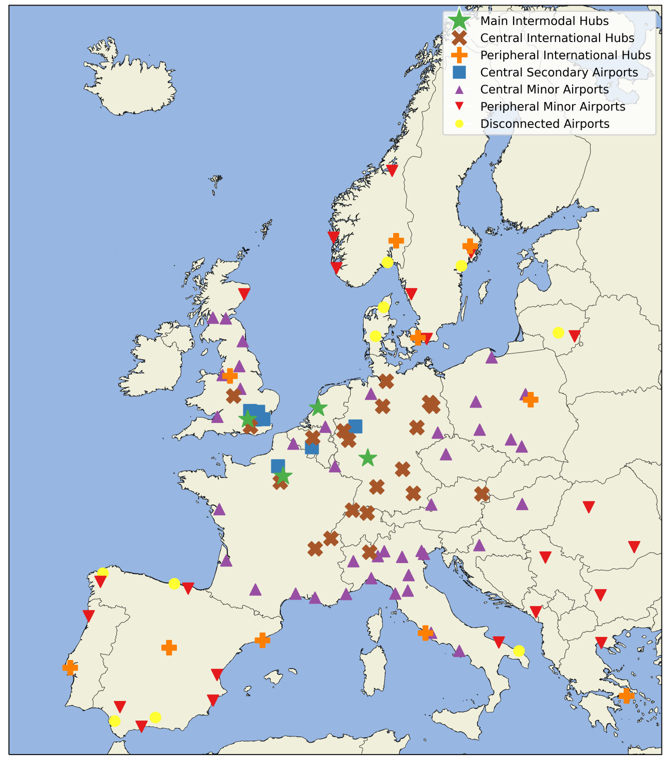


Fig. 10. Scenario 2 in seven classes

C. Scenario 3: Combination of air/rail policies

The previously addressed scenarios show the two sides of the air/rail integration. However, to achieve a higher contribution to the network intermodality, a combination of the two scenarios could cause another shift in the intermodal network. Combining adding extra services with the imposed reduction of short-haul flights creates a different scenario. This scenario involves an even more complex coordination of international stakeholders, including railway managers and operators, airports, airlines, and governments.

Scenario 3 combines both previous scenarios, considering changes in the air and rail factors. Also, the effects on the clustering seem to be a combination of both scenarios and the shifts that occurred in their implementation. Compared to the benchmark situation, Cluster 1 has become smaller, as happened after the implementation of scenario 2. However, these airports are redistributed over Cluster 6 instead of Cluster 7. As most airports join Cluster 6, this cluster can be interpreted differently than the *Central Secondary Airports* from the previous scenarios. The airports that have changed clusters are shown in Table VII.

TABLE VII
MOVED AIRPORTS IN SCENARIO 3

	Added	Removed
Cluster 1		*NUE, WAW, GVA, LYS, LEJ, BGY, LIN, BSL, MMX, BHX, EIN, BLQ, WMI, BMA*
Cluster 2	*WAW*	
Cluster 3	*BMA, MMX*	
Cluster 4	*WMI*	
Cluster 5		
Cluster 6	*NUE, GVA, STN, LYS, LEJ, BGY, TXL, LIN, BSL, BHX, EIN, BLQ, HAJ, SXF, STR, HAM*	
Cluster 7		*STN, TXL, HAJ, SXF, STR, HAM*

In this scenario, only Cluster 5, with the four main hubs, remains unchanged compared to the benchmark network. The changes within the variables for each cluster are shown in Table VIII. It is also noted that all airports leaving Cluster 7 join Cluster 6. London Stansted is one of the airports and matches the intercontinental flight variables (23.000, 15%) better with Cluster 6. The other shifts of clustering are not having a big impact on the variables' values. The increase in the number of airports assigned to Cluster 6 forms a situation in which not all airports are secondary for their catchment area. Still, the connecting factor within this cluster is the degree of rail connections. Therefore, this cluster can be classified as *Land-connected Minor Airports*, showing the opposition of high performance on rail services, while a low performance of air services is identified. For the other clusters, there is no reason to see a different interpretation of their classification. Therefore, the other clustering labels remain equal for this scenario. The new classification is shown in Table IX. The visualization of the new clustering is shown in Figure 11.

TABLE VIII
DIFFERENCES HIGHLIGHTED FOR CHANGED CLUSTERS BETWEEN BENCHMARK AND SCENARIO 3

	TF air	IF air	IR air	NC air	AR air	STA	SAC	TF rail	SP rail	CN rail	n
B1	48.456	4.913	0,09	40	23	**0,24**	33	925	2,86	0,30	49
S1	43.193	6.740	0,16	35	22	**0,17**	**32**	2.515	2,93	0,27	35
B2	**237.474**	**39.747**	0,16	**100**	10	**1,33**	**31**	617	3,61	0,16	9
S2	**228.771**	52.704	0,23	**89**	10	1,30	**32**	1.855	3,52	0,17	10
B3	51.800	5.908	0,11	38	**7**	0,40	34	**102**	**4,02**	**0,04**	20
S3	49.821	6.829	0,14	33	**7**	0,36	35	**351**	**3,93**	**0,04**	22
B4	**13.775**	**469**	**0,03**	**19**	**7**	0,30	**96**	157	3,56	0,06	10
S4	**14.053**	**1.041**	**0,07**	**20**	**7**	0,27	**94**	456	3,48	0,07	11
B6	43.983	2.624	0,06	46	**30**	0,33	71	**2827**	2,36	**0,81**	6
S6	69.539	13.104	0,19	48	30	0,55	42	6.986	2,47	0,62	22
B7	168.961	29.103	**0,17**	80	**30**	1,13	34	2676	**2,25**	0,69	15
S7	211.308	**58.475**	**0,28**	**89**	**31**	**1,33**	34	**8.462**	**2,19**	**0,67**	9

TABLE IX
CLASSIFICATIONS AFTER IMPLEMENTING SCENARIO 3

C#	Class	n	Airports
C5	**Main Intermodal Hubs**	4	*CDG, FRA, AMS, LHR*
C7	**Central International Hubs**	9	*VIE, BRU, ORY, CGN, DUS, MUC, MXP, ZRH, LGW*
C2	**Peripheral International Hubs**	10	*CPH, ATH, FCO, OSL, WAW, LIS, BCN, MAD, ARN, MAN*
C6	**Land-connected Minor Airports**	22	*BVA, LCY, SEN, CRL, LYS, BSL, SXF, HAM, NUE, LEJ, STR, TXL, HAJ, DTM, BGY, LIN, BLQ, EIN, GVA, BHX, LTN, STN*
C1	**Central Minor Airports**	35	*Remaining airports*
C3	**Peripheral Minor Airports**	22	*BIO, VLC, SVQ, CLJ, BMA, SOF, SKG, BRI, VNO, TGD, BGO, TRD, SVG, OPO, OTP, BEG, ALC, AGP, SCQ, GOT, MMX, ABZ*
C4	**Disconnected Airports**	11	*LCG, GRX, XRY, SDR, BDS, KUN, TRF, NYO, BLL, AAL, WMI*

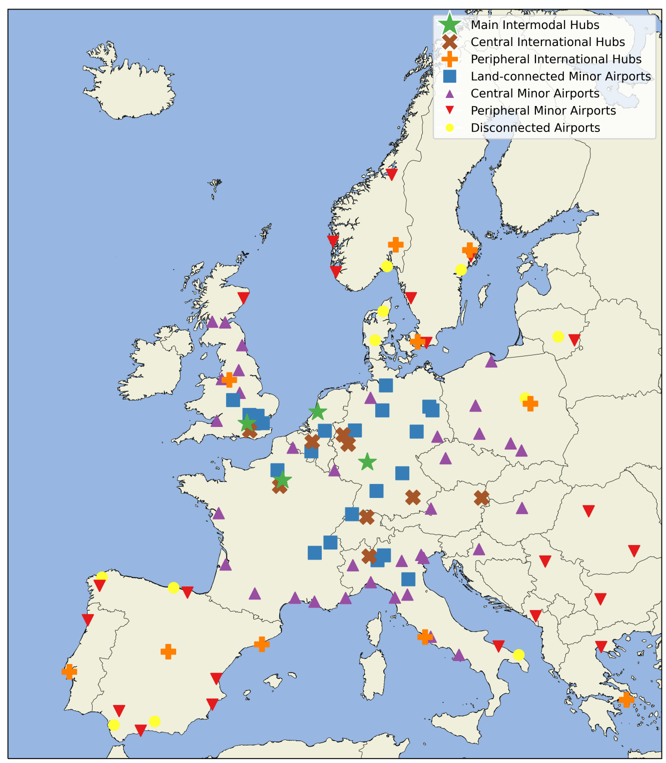


Fig. 11. Scenario 3 in seven classes

D. Interpretation of classification

The classifications in all scenarios are based on several factors. First of all, three levels of air services can be divided into *main*, *international*, and *minor* airports. These levels explain the air side variables of a cluster.

Also, a correlation between the location and assignment to specific clusters can be found by studying the geographical locations of the clusters. Comparing these findings shows a strong correlation with the level of rail service. The keywords *central* and *peripheral* are used to identify the clusters' geographical location.

Furthermore, the keyword *hubs* indicates a good connection between the air and rail side. These clusters perform highly on both the air and the rail side variables. Moreover, they are distinguished by the on-site availability of railway stations.

Only one cluster remains unlabeled when the classification is limited to these keywords. Cluster 4 is an extremely low-performing cluster, i.e. low number of services in rail and aviation, of airports with a very high travel time to the linked FUA. The classification for these low-impact airports is labeled *Disconnected Airports*.

Combining these keywords enables the clusters to be qualitatively labeled according to their distinguishing characteristics. The classes that use the keyword *hubs* can be considered airports where it can realistically perform an intermodal trip.

To translate the classification of airports into their intermodal possibilities, the classes can be grouped on their role in an intermodal network. Table X shows the division of the clusters per group of intermodal possibilities. The classification keyword *hubs* indicates the airports that are ready for substitution, with high services on the air and rail side and a strong connection with the railway network. Cluster 6, which changes its label depending on the scenario's effects, shows potential for intermodality when classified as a *land-connected airport*. These airports have a medium-level service on the air side and lack a strong connection with the main railway network. However, the FUA they're situated in already has a strong connection to the railway network. An analysis of frequency per unique connection showed a fairly high intercontinental performance of this class for the destinations the airports serve. If these airports gain a stronger connection with the rail network, for example, by constructing on-site railway stations, they can play a small role in an intermodal network. The other clusters have a service level that is too low to offer intermodal possibilities.

TABLE X
CLASSES GROUPED BY INTERMODAL POSSIBILITIES

High intermodality	Main Hubs (C5) Central International Hubs (C7) Peripheral International Hubs (C2)
Potential intermodality	Land-Connected Minor Airports (C6*)
No intermodality	Central Minor Airports (C1) Peripheral Minor Airports (C3) Disconnected Airports (C4) Central Secondary Airports (C6*)

Cluster 6 is classified differently for different scenarios. This influences its position in intermodal possibilities.

To assess the target group's of the intermodal classified airports, the airports are linked to the passenger numbers from the Eurostat flight data [31]. The total number of passengers traveling through the airports classified in one of the high intermodality clusters gives an insight into the effect of an increased contribution of intermodality to the network. These numbers are shown for each scenario in Table XI. This analysis shows that when the airports in the potential intermodality group have met the conditions to classify as high intermodal airports, the scenarios cause an increase in intermodality throughout the entire European network. The extension of rail services can potentially increase the number of passengers who can travel intermodaly by 1.9%, while the reduction of short-haul flights could stimulate intermodality for 4.3% more passengers. The combination scenario increases this effect to 14.3%, with a large factor of uncertainty, by the classification of the potential intermodality.

TABLE XI
TOTAL NUMBER OF PASSENGERS PER GROUP OF INTERMODALITY

	High intermodality	Potential intermodality	No intermodality	Change [%]
Benchmark	928	-	478	-
Scenario 1	946		460	+1,9%
Scenario 2	968	-	438	+4,3%
Scenario 3	857	204	345	-7,7% / +14,3%

Number of passengers in millions. Changes are given for high intermodality and total intermodality, including potential intermodality.

Identifying the airports with the potential to classify as intermodal hubs indicates that focusing on these airports has the largest impact on the networks' intermodality. However, solely focusing on these airports could be counterproductive. The central situation of most airports that have the potential to become a hub also means that these airports could have several alternatives for travel. Partly discouraging short-haul flights would probably encourage people to opt for a different airport and still use the trusted modes of transport. Therefore, for these policies, an integral approach for the entire intermodal network is needed. Initiatives like the short-haul flight ban in France would contribute more when imposed by a larger collective, such as the European Union or the slot allocation authority.

On the other hand, the extension of rail services can focus first on the main connections of the identified airports. This scenario contains an encouraging approach, for which extra options in the particular airport will only improve its attractiveness en could potentially enable intermodality. In the same context, connecting *Minor Airports* between themselves would not contribute to the intermodal network significantly. When the railway network is improved or new rolling stock can be assigned to routes, it would be good to consider the access routes to the airports (potentially) classified as high intermodality airports.

VI. CONCLUSION

The research aimed to explore the role of airports in a European intermodal network and investigate the effects that possible policy scenarios have on this network. The study is purely focused on the network's perspective of intermodality. The analysis of the benchmark network showed possibilities to cluster the airports in seven clusters, allowing the network to be distinguished on a variety of factors. The clustering results clearly distinguish the different variables' performance and the variable values' combinations. The assessed variables are focused on air services, rail services, the rail network, and physical integration characteristics.

In the classification of the airports, the keywords *central* and *peripheral* can describe the geographic situation of the clusters, while *main*, *international*, and *minor* can identify the level of air services that are offered at the airports. A combination of high air and rail service levels and a strong physical connection with their FUA could be classified as a *hub*.

The clustering and classification of the benchmark scenario was used to interpret the effect of scenario implementations. The assessment of the different scenarios showed a fairly constant classification. The *Main Hubs* class, consisting of the four large European airports, was a constant factor through all scenarios, performing high on all defining variables. Also, in each scenario, almost all airports could be classified by combinations of the mentioned keywords considering the number of air services and the geographical location. One cluster has changed classification during the scenario classification. The *Central Secondary Airports*-class from the benchmark scenario grew to a larger cluster when combining the constructed scenarios, which changed the perspective on the cluster to *Land-connected Minor Airports*. Apart from the changes of entire classes, individual airports shift between the clusters in each scenario. This causes individual airports to shift from a *minor airport* to an *international hub*, which changes their role in intermodality, as was the case for Warsaw (WAW), for example. Also, a group of airports made the same shift, changing their clustering to a classification considered as an intermodal possibility.

The classes could also be grouped based on their value in intermodality. All classes considered hubs can be considered to have a high value in intermodality. The cluster that changed to *land-connected airports* is labeled as a potential value in intermodality for the scenarios it takes that classification, as the level of air services is relatively low. If also the potential intermodal airports have an added value, the scenarios show increased enabled passengers for intermodality of 1.9%, 4.3%, and 14.3%, respectively.

The study aimed to identify the airports where potential lies for intermodal travel. In the scenarios, shifts of airports are shown from being a *minor airport* to *international hubs*. These borderline cases can be considered a contribution to intermodality. Also, the airports that were considered hubs through the entire analysis show, even more, potential, to promote intermodality in these airports. However the individual cases of these airports should be considered to see whether the individual characteristics of these airports also allow the airports to facilitate that contribution to the network. If these are considered to contribute to intermodality, offering additional services that are neglected in this study, such as integral ticketing, timetable synchronization, and luggage handling, can create strong intermodal nodes in the European network.

VII. Discussion & Recommendations

During this study, intermodality was assessed in a very quantitative way, neglecting softer attributes, like additional services. Also, all variables that were used, were considered equally important by the application of a PCA. It could have been useful to assign a stronger weight to variables dependent on the currently available infrastructure to prioritize clustering based on the possibilities within the current infrastructure.

Furthermore, the AHC is a strongly number-based method, grouping different airports without considering individual possibilities. The clustering results throughout the different scenarios showed that some airports are very sensitive to changes in the data. These borderline cases in the data are hard to put in the correct cluster because there's no possibility for individual assessments.

Overall, the method enables one to create insights into the current situation of the offered services in Europe. It can identify the strong outliers in the available data. However, the descriptive statistics of the data showed large groups of medium-performing airports for multiple variables. In these groups, it is hard to create a further distinction. To assess strategic decisions for these airports, more specific research should be performed, with a more individual approach, considering the possibilities of the individual airports. This individual approach can also help assess whether constructed scenarios are possible at all airports or clusters and how the data could be modified to create a new clustering.

The individual situation of each airport requires a smaller study to differentiate the implemented scenarios for each airport. The performed research assumed equal treatment for all airports and limited differentiation to applying scenarios to specific air and rail services categories. Personalizing these implementations would have been too time-consuming, given the scope of the research.

The research into intermodality of air and rail travel can be further improved if there are more findings available on the influence of additional services, or 'soft attributes'. The air and rail modes are assessed mostly on the frequency of their services and some network properties. The integration between the two is mostly physically orientated because there's little scientific knowledge on the quantitative impact of, for example, integrated ticketing and luggage handling on the experienced integration that can be used in this quantitative analysis. Research on the transformation of qualitative services to quantitative scoring also helps assess the implementation of these important services in the analysis.

Finally, this research assumes the physical integration between the modes fairly linearly by the travel time to the city center. However, it is not considered what travel time to the main railway network is assumed to be sufficient, considering it to be an integration of the modes, even if additional services are included. Research on the willingness to travel to a train station can help categorize travel times and better assess the accessibility of the railway network. Furthermore, the availability of (HSR) stations at the airport creates a three-step assessment. However, it would be interesting to see the relation between the availability of low-service on-site stations and high-service stations available within a particular travel time. This would result in a performance score for physical integration of the two modes, where a variety of solutions would be able to create a strong integration.

REFERENCES

[1] S. R. Department, "Number of scheduled passengers boarded by the global airline industry from 2004 to 2022," 6 2023. [Online]. Available: https://www.statista.com/statistics/564717/airline-industry-passenger-traffic-globally/

[2] E. E. Agency, *Transport and environment report 2022*, 2022.

[3] ICAO, "Air traffic recovery is fast-approaching pre-pandemic levels," 2022. [Online]. Available: https://www.icao.int/Newsroom/Pages/Air-traffic-recovery-is-fastapproaching-prepandemic-levels.aspx#:~:text=ICAO%27s%20latest%20analyses%20reveal%20that,of%20its%202019%20high%20point.

[4] E. Commission, "The european green deal," 2019. [Online]. Available: https://eur-lex.europa.eu/resource.html?uri=cellar:b828d165-1c22-11ea-8c1f-01aa75ed71a1.0002.02/DOC_1\&format=PDF

[5] ——, "Action plan to boost passenger rail," 2021. [Online]. Available: https://transport.ec.europa.eu/news/action-plan-boost-passenger-rail-2021-12-14_en

[6] V. Reiter, A. Voltes-Dorta, and P. Suau-Sanchez, "The substitution of short-haul flights with rail services in german air travel markets: A quantitative analysis," *Case Studies on Transport Policy*, vol. 10, pp. 2025–2043, 12 2022.

[7] A. S. Bergantino and L. Madio, "Intermodal competition and substitution. hsr versus air transport: Understanding the socio-economic determinants of modal choice," *Research in Transportation Economics*, vol. 79, p. 100823, 3 2020.

[8] F. Zhang, D. J. Graham, and M. S. C. Wong, "Quantifying the substitutability and complementarity between high-speed rail and air transport," *Transportation Research Part A: Policy and Practice*, vol. 118, pp. 191–215, 12 2018.

[9] F. Savelberg and M. D. Lange, "Substitutie mogelijkheden van luchtvaart naar spoor," 2018.

[10] B. Donners, E. van Buuren, and R. Rijniers, "Steden verbonden, trein of vliegtuig, effecten van substitutie," *Colloquium Vervoersplanologisch Speurwerk*, 2018.

[11] E. Kroes and F. Savelberg, "Substitution from air to high-speed rail: The case of amsterdam airport," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, pp. 166–174, 4 2019. [Online]. Available: https://journals.sagepub.com/doi/10.1177/0361198119839952

[12] B. Donners and M. H. Kantelaar, "Emissies van korte afstandsvluchten op nederlandse luchthavens," 11 2019.

[13] Schiphol, "Maandelijkse verkeer & vervoer cijfers december 2019," 12 2019.

[14] F. Airport, "Frankfurt airport air traffic statistics 2019," 2019.

[15] G. ADP, "Traffic at paris aéroport up by 2.5

[16] Statista, "Share of passengers transferring and terminating at heathrow airport in the united kingdom (uk) between 2002 and 2019," 12 2020. [Online]. Available: https://www.statista.com/statistics/303920/flight-transfers-passengers-terminating-at-heathrow-airport-uk/

[17] J. G. Brida, J. C. Martín, . C. Román, R. Scuderi, C. Román, and C. Es, "Air and hst multimodal products. a segmentation analysis for policy makers," 2017.

[18] S. Wu and D. Han, "Accessibility of high-speed rail (hsr) stations and hsr–air competition: Evidence from china," *Transportation Research Part A: Policy and Practice*, vol. 166, pp. 262–284, 12 2022.

[19] F. Song, S. Hess, and T. Dekker, "Accounting for the impact of variety-seeking: Theory and application to hsr-air intermodality in china," *Journal of Air Transport Management*, vol. 69, pp. 99–111, 6 2018.

[20] Eurocontrol, "Care ii: The airport of the future: Central link of intermodal transport? - wp 1 - review of the current intermodality situation," 2004.

[21] J. D. Costa and J. A. de Abreu Silva, "Factors of air-rail passenger intermodality," 2012.

[22] I. R. S. Group, N. R. C. U. T. R. Board, A. C. R. Program, and U. S. F. A. Administration, *Integrating aviation and passenger rail planning*, 2015.

[23] R. voor de Leefomgeving en Infrastructuur, "Verzet de wissel: naar beter internationaal reizigersvervoer per trein." 7 2020. [Online]. Available: www.rli.nl

[24] Y. Wang, Q. Lu, X. Cao, X. Zhou, V. Latora, L. C. Tong, and W. Du, "Travel time analysis in the chinese coupled aviation and high-speed rail network," *Chaos, Solitons & Fractals*, vol. 139, p. 109973, 10 2020.

[25] F. Bruno, "The connectivity of the long-distance rail and air transport networks in europe," 2022. [Online]. Available: http://resolver.tudelft.nl/uuid:d2322bfd-248e-411c-8e75-c8ef295de9a5

[26] M. S. Ryerson and H. Kim, "Integrating airline operational practices into passenger airline hub definition," *Journal of Transport Geography*, vol. 31, pp. 84–93, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0966692313001099

[27] Y. Chen, K. Fuellhart, S. Zhang, and F. Witlox, "Airport classification in chinese multi-airport regions: An interaction network perspective between aviation and high-speed rail," *European Journal of Transport and Infrastructure Research*, vol. 22, pp. 1–21, 2022. [Online]. Available: https://journals.open.tudelft.nl/ejtir/article/view/6043/5264

[28] M. van Infrastructuur en Waterstaat, "Luchtvaartnota 2020-2050," 11 2020.

[29] M. Givoni and D. Banister, "Airline and railway integration," *Transport Policy*, vol. 13, pp. 386–397, 9 2006.

[30] W. Grimme and M. Jung, "Towards more sustainability? – the development of aviation emissions from germany between 1995 and 2016," *Proceedings of the 22nd Air Transport Research Society World Conference*, 7 2018. [Online]. Available: https://elib.dlr.de/121383/

[31] E. Commission, "Eurostat data browser," 2022. [Online]. Available: https://ec.europa.eu/eurostat/databrowser/view/DEMO_R_PJANGRP3__custom_2790186/bookmark/table?lang=en\&bookmarkId=168107c0-c56f-4adb-85b4-381126e7944b

[32] Z. Yuan, C. Dong, and X. Ou, "The substitution effect of high-speed rail on civil aviation in china," *Energy*, vol. 263, p. 125913, 1 2023.

[33] A. Maćkiewicz and W. Ratajczak, "Principal components analysis (pca)," *Computers & Geosciences*, vol. 19, pp. 303–342, 3 1993.

[34] I. D. Baruah, "Cheat sheet for implementing 7 methods for selecting the optimal number of clusters in python," 10 2020.

[35] ProRail, "Integrale mobiliteitsanalyse 2021," 2021. [Online]. Available: www.prorail.nl

[36] Jernbanedirektoratet, "Jernbanedirektoratets perspektivanalyse," 1 2020.

[37] Deutschlandtakt, "Deutschlandtakt," 2023. [Online]. Available: https://www.deutschlandtakt.de

[38] Railtech, "New european mobility strategy aims to triple high-speed rail traffic," 12 2020. [Online]. Available: https://www.railtech.com/policy/2020/12/10/new-european-mobility-strategy-aims-to-triple-high-speed-rail-traffic/

[39] E. Barcelona, "Más viajeros, más trenes, menos emisiones," 6 2021.

[40] Railtech, "Romania to study first high-speed railway and revamp rail infrastructure," 4 2022. [Online]. Available: https://www.railtech.com/infrastructure/2022/04/28/romania-plans-to-study-first-high-speed-railway-and-revamps-rail-infrastructure/

[41] C. P. Komunikacyjny, "Centralny port komunikacyjny," 2023. [Online]. Available: https://www.cpk.pl/en/

[42] T. E. R. R. A. Council, "Rail route 2050: The sustainable backbone of the single european transport area," 2011. [Online]. Available: www.errac.org

[43] Euronews, "It's official: France bans short-haul domestic flights in favour of train travel," 12 2022. [Online]. Available: https://www.euronews.com/green/2022/12/02/is-france-banning-private-jets-everything-we-know-from-a-week-of-green-transport-pr