Universidade do Porto
Faculdade de Engenharia
FEUP

DEVELOPMENT AND APPLICATION OF A STRATEGIC RAIL NETWORK OPTIMIZATION MODEL FOR FREIGHT TRANSPORT

Doctoral thesis

Thesis submitted to the Faculty of Engineering of the University of Porto in candidacy for the degree of Doctor of Philosophy in the field of Transportation Systems

Author

Luís Rocha Neves Couto Maia

Supervisors

António José Fidalgo do Couto (University of Porto, Portugal)

Paulo Fonseca Teixeira (University of Lisbon, Portugal)

Porto

October, 2013

Aos meus pais

FINANCIAL SUPPORT

This research work was financed by the Portuguese Science and Technology Foundation (FCT) through a PhD grant with the reference SFRH/BD/60449/2009.

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

AGRADECIMENTOS

A ideia de fazer um doutoramento foi-me sugerida pelo Professor António Couto, quando ainda estava a concluir a tese de mestrado em Engenharia Civil sob a sua orientação. Foi a partir daí que comecei a sonhar com este projecto, tendo iniciado o doutoramento cerca de um ano depois. Os conselhos e a ajuda do Professor António Couto foram essenciais para dar forma à ideia desta tese e, posteriormente, para me auxiliar durante o desenvolvimento da mesma, como meu orientador. A sua excelente orientação, apoio, incentivo e amizade foram um factor determinante para que concluísse esta tese com sucesso, pelo que lhe estou profundamente grato.

Agradeço também a todos os professores que me acompanharam ao longo deste processo por todos as ideias e conhecimentos que me transmitiram. Gostaria de agradecer em particular aos professores que mais influenciaram o meu trabalho com as suas ideias, sugestões e conselhos. Destaco o meu co-orientador, o Professor Paulo Teixeira, pelo apoio, conselhos e sugestões apresentadas, e os Professores Jorge Pinho de Sousa e António Pais Antunes, pelas suas críticas construtivas e sugestões de melhoria, que muito contribuíram para o aperfeiçoamento deste projecto.

Em termos institucionais, agradeço à Fundação para a Ciência e a Tecnologia pelo apoio financeiro dado, que foi fundamental para que eu pudesse realizar este doutoramento. Gostaria também de agradecer à REFER, em particular ao Engenheiro Francisco Ganhão, pela disponibilidade demonstrada no fornecimento de dados de tráfego que se revelaram fundamentais para a aplicação do modelo desenvolvido à rede da Península Ibérica.

Aos meus companheiros do LAT, com quem partilhei o local de trabalho nos últimos quatro anos, agradeço a companhia, amizade e boa disposição, que tornaram a nossa sala num ótimo local de trabalho. Este agradecimento estende-se tanto aos atuais ocupantes, Cristina, Lobo, Marco e André, como aos que foram passando por cá ou foram sempre visitas habituais, Juliana, Nuno, Luís e Giulio. Agradeço também o apoio e constante boa disposição dos Professores Carlos Rodrigues e José Pedro Tavares, que foram uma presença frequente na nossa sala.

Agradeço aos meus pais a educação que me deram e todas as oportunidades que me proporcionaram ao longo da vida, incluindo a oportunidade de obter os estudos que me levaram até este doutoramento. Agradeço-lhes também o apoio e encorajamento que sempre me souberam dar, assim como o facto de terem sido sempre um modelo para mim. Agradeço também ao resto da minha família pelo apoio e incentivo que me deram ao longo deste longo caminho. Não posso também deixar de referir os meus falecidos avós, a quem também devo muito da pessoa que sou.

Agradeço também a todos os meus amigos pela alegria que me transmitiram e o apoio que me deram ao longo destes anos. Estou particularmente grato aos meus amigos mais próximos pela sua amizade inabalável e por todos os bons momentos porque passamos. Foi nos nossos frequentes e inesquecíveis jantares que vivi alguns dos melhores e mais felizes momentos destes últimos anos.

Por último quero agradecer à minha mulher, Sara, por ter estado ao meu lado em todos os momentos deste longo caminho e por fazer de mim uma pessoa feliz. O seu incondicional apoio e confiança deram-me força e motivação para vencer as dificuldades, e a sua contagiante alegria e efusividade ensinaram-me a valorizar as minhas pequenas vitórias.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AoN - All or nothing

O/D - Origin/destination

CO2 - Carbon dioxide

TEU - Twenty foot equivalent unit

GDP - Gross domestic product

SCGE - Spatial computable general equilibrium

DNDP – Discrete network design problem

CNDP - Continuous network design problem

NUTS - Nomenclature of territorial units for statistics

REFER - Rede Ferroviária Nacional - Portuguese rail network administrator

ADIF - Administrador de Infraestructuras Ferroviarias - Spanish rail network administrator

ABSTRACT

This thesis presents an optimization approach to the long term planning of rail network investments at a strategic level, with the goal of improving the conditions for the transport of freight. The work that was developed along this thesis can be divided in four major areas: development of a strategic freight traffic assignment model; development of a rail network optimization model for freight transport; validation of the network optimization model through its application to the transport network of the Iberian Peninsula; application of the optimization model under different hypothetical future scenarios in order assess how these different circumstances will affect the impact of potential network investments.

The strategic freight traffic assignment model contemplates two different types of cargo, intermodal cargo and general cargo, using different assignment techniques for each of them. Its main innovative feature is the fact that it takes into account both capacity constraints and a variable perception of costs by users, while being much simpler than stochastic equilibrium model.

The rail network optimization model for freight transport is quite flexible and innovative, allowing for both upgrades in the quality of existing links as well as the construction of new ones. The assessment of the quality of each network improvement solution is based on the reduction of the total generalized transport costs and CO2 emissions. As for the optimization process, it uses a local search heuristic which tries to meet a balance between efficiency and effectiveness, by delivering good solutions in a reasonable computing time.

The validation of the network optimization model consisted in its application to the transport networks of the Iberian Peninsula in order to obtain an optimal solution for investing a specific amount of money in the region's rail network. This application served to test and evaluate the performance of the network optimization model, as well as to calibrate and empirically validate the traffic assignment process, by comparing the estimated rail traffic with the actual traffic in the network.

Lastly, the application of the optimization model under different scenarios considered different evolutions for the demand of freight and for the price of oil, in order to study the influence of these changes on the distribution of traffic and on the impact of the network improvements. This included a robustness analysis which helped to identify the network improvement solutions that can best cope with different possible future scenarios.

Keywords: network optimization; traffic assignment; freight transport; Iberian Peninsula.

RESUMO

Esta tese apresenta uma abordagem baseada num método de otimização para o planeamento a longo prazo de investimentos estratégicos em redes ferroviárias, com o objectivo de melhorar as condições para o transporte de mercadorias. O trabalho que foi desenvolvido ao longo desta tese pode ser dividido em quatro grandes áreas: desenvolvimento de um modelo estratégico de afetação de tráfego de mercadorias; desenvolvimento de um modelo de optimização de redes ferroviárias na ótica do transporte de mercadorias; validação do modelo de optimização de redes através da sua aplicação à rede de transporte da Península Ibérica; aplicação do modelo de otimização considerando vários cenários hipotéticos futuros, de modo a avaliar como é que essas diferentes circunstâncias irão afetar o impacto de potenciais investimentos na rede de transporte.

O modelo estratégico de afetação de tráfego de mercadorias contempla dois tipos de carga, carga intermodal e carga geral, usando diferentes técnicas de afectação de trafego para cada uma delas. A sua característica mais inovadora é o facto de considerar restrições de capacidade e uma perceção dos custos variável pelos utilizadores, sendo muito mais simples do que um modelo estocástico de equilíbrio.

O modelo de optimização de redes ferroviárias na ótica do transporte de mercadorias é bastante flexível e inovador, permitindo tanto melhorias na qualidade das ligações existentes como a construção de novas linhas. A avaliação da qualidade de cada solução de melhoria da rede é baseada na redução dos custos generalizados de transporte e das emissões de CO2. O processo de optimização usa uma heurística de pesquisa local que tenta atingir um equilíbrio entre eficiência e eficácia, obtendo boas soluções num período de computação razoável.

A validação do modelo de optimização de redes consistiu na sua aplicação à rede de transporte da Península Ibérica, de modo a obter a melhor solução para investir uma quantidade específica de dinheiro na rede ferroviária da região. Esta aplicação serviu para testar e avaliar a performance do modelo de optimização de redes, bem como para calibrar e validar empiricamente o modelo de afectação de tráfego, comparando o tráfego ferroviário estimado com o tráfego real.

Para finalizar, a aplicação do modelo de otimização considerando vários cenários simula várias evoluções para a procura de transporte de mercadorias e para o preço do petróleo, de modo a estudar a influência que essas alterações têm na distribuição do tráfego e no impacto das soluções de melhoria da rede. Esta aplicação incluiu uma análise de robustez que ajudou a identificar quais as soluções de melhoria da rede capazes de lidar melhor com diferentes cenários futuros.

Palavras-chave: otimização de redes; afetação de tráfego; transporte de mercadorias; Península Ibérica.

DEVELOPMENT AND APPLICATION OF A STRATEGIC RAIL NETWORK OPTIMIZATION MODEL FOR FREIGHT TRANSPORT

Chapter 1

INTRODUCTION

1. INTRODUCTION

1.1. Problem statement

Freight transport plays a crucial role in the day to day life of any modern society, being critical to a large part of the economy. However, it is a subject that has received much less attention by the academia than its passenger counterpart. This is probably due to the fact that freight transport is not as appealing to policy makers as passenger transport and due to the natural complexity of this subject. This complexity is mainly justified by the multiplicity of goods transported, the complexity of the freight supply chain and the difficulty in getting the needed data, which is due to the general lack of complete and up to date databases and the unwillingness of transport companies to share data due to confidentiality reasons. The economic importance of freight transport highlights the importance of thinking of it as a separate part of the transport spectrum, instead of being grouped with passenger transport, as it frequently happens. Therefore, it is important to study this type of transport using models specifically made for it, in order to account for its distinct characteristics.

Furthermore, when planning for improvements in the transport network, the main focus of attention tends to be put on their impact on passenger transport, rather than on freight. This is exacerbated by the fact that the decisions to make big infrastructure investments are always political, with politicians tending to favor passenger transport, which has a greater direct impact on the lives of people. This is a problem due to the fact that although freight transport is not as appealing to the ordinary person as passenger transport, it has an important economic impact which, in the end, will have an important impact on people’s lives. To overcome this problem, it is necessary to study network investments weighting their impact both on freight and on passenger transport, avoiding the common mistake of focusing on passengers and paying little attention to freight. This is justified not only by the importance of freight in itself, but also because the needs of freight transport are different from those of passenger transport, meaning that the type of network improvements that benefit freight transport can be considerably different from the improvements which benefit passenger transport.

In order to assess the impact of different network improvement solutions on freight transport it is necessary to define the parameters which will be used to assess the value of each improvement solution. There are several different parameters that can be used for this purpose, depending on the goals of the type of analysis being performed. In the case of large scale macro analysis, the parameters are generally related to generalized transport costs and the environmental impact caused by freight transport. These two key factors are able to measure the impact that the improvement operations have, both on the freight operators and on society as a whole: transport costs affect both freight operators and final costumers, and the environmental impact caused by freight transport affects all the people that are harmed by it in any way. Furthermore, there is the possibility that freight carriers may be forced in the future to pay for the pollution they cause, in an effort to internalize the external costs caused by them, making the environmental impact part of the generalized transport costs.

Planning for investments on transport infrastructure is something that inherently has a long term horizon and, as such, has to be carefully thought through. The long term nature of this kind of analysis means that some factors and parameters that today assume a certain value, may suffer significant changes in the long run. As such, it is important to take that into account that there may be fundamental changes in some key factors, such as the demand for freight transport and the price of oil. This means that the impact of the network investments will depend on exogenous factors that are not controlled by planners. Due to that, the choice of the best possible investment will likely vary according to the used values for different parameters, resulting that the most responsible investments will be those that deliver a robust solution, capable of copping with different possible future scenarios.

1.2. Research objectives

The goal of this thesis is to study the impact on freight transport of making significant investments on rail transport networks, namely by studding what kind of investments are most beneficial for this type of transport. Since there is no available network optimization model that is able to cope with the particular demands of this study, this will require the development of a traffic assignment model and a network optimization model made specifically for freight transport.

The first objective of this thesis is to develop a freight traffic assignment model for freight designed to model transport networks at a strategic planning level. The model should simulate land freight transport, by considering both road and rail transport modes, and it must consider two different types of cargo: general cargo and intermodal cargo.

The second objective is to develop a rail network optimization model for freight incorporating the previously developed traffic assignment model. The model should be able to determine how to invest a specific volume of capital on a rail network in order to improve the conditions for freight transport. The quality of the each network optimization solution should be assessed based on the minimization of the generalized costs and of the environmental impact.

The third objective is to apply the network optimization model to a real network, in order to properly validate and calibrate the traffic assignment model and to analyze the results from the network optimization.

The forth objective is apply the network optimization model on a real network considering possible modifications on some key variables that affect freight transport, such as the demand for freight and the price of oil. This will be done in order to study the influence of these changes on the distribution of traffic and on the impact of the network improvements.

Regarding the divulgation of the work, the objective is to write four scientific papers: a paper about the development of the freight traffic assignment model; a second paper covering the development of the network optimization model; a third paper on the application of the optimization model on a real transport network; a forth paper on the application of the model under various scenarios simulating possible modifications on some key variables.

1.3. Description of the thesis

This Phd thesis is structured in seven chapters. Chapters 2 to 5 are based on scientific papers and can be read separately, containing an introduction, a body and a conclusion. Although this structure can lead to the repetition of some concepts and ideas, there was an effort to avoid unnecessary repetitions.

Chapter 2 presents an innovative freight traffic assignment model for road and rail transport. It considers two different types of cargo, general cargo and intermodal cargo, and uses different assignment techniques for each type. Its main innovative feature is the fact that it takes into account both capacity constraints and a variable perception of costs by users, while being much simpler than stochastic equilibrium model. The model is designed to model macro networks with a high aggregation level and does not require very detailed data inputs, being a strategic planning model. The purpose of the model is to estimate the movement of freight at a national or international scale, being a useful tool for a variety of planning and policy decisions.

Chapter 3 is the core of this thesis, presenting a strategic rail network optimization model for freight. The model assesses the best way in which to invest a specific amount of money in a given rail network, with the quality of the improvement solutions being measured by the reduction of the total generalized costs and CO2 emissions. The optimization process is quite flexible and innovative, allowing for both upgrades in the quality of existing rail and intermodal terminal links as well as the construction of new ones, not having a limit on the number or variety of improvement solutions. This is achieved by defining a set of possible link levels for each link type, according to the users' preferences, including the mere possibility of building a link. The optimization process is based on a heuristic and tries to meet a balance between efficiency and effectiveness, by delivering good solutions in a reasonable computing time.

Chapter 4 is dedicated to the application of the network optimization model on the transport network of the Iberian Peninsula. This is done in order to validate the traffic assignment model and to analyse the results from the optimization process. Two different scenarios are considered: a scenario in which the only goal is the minimization of transport costs and another in which the goal of reducing CO2 emissions given an equal importance. The results validate the traffic assignment model and reveal some critical findings regarding the planning of investments in rail infrastructure.

Chapter 5 features the application of the network optimization model on the transport network of the Iberian Peninsula considering twelve different scenarios. These scenarios simulate possible modifications on some key variables, namely on demand for freight and the price of oil. This is performed in order to study the influence of these changes on the distribution of traffic and on the impact of the network improvements. A robustness analysis is performed on the results obtained for the various scenarios in order to identify robust solutions that can cope with the various possible scenarios.

Chapter 6 is dedicated to the description of the network optimization program that was developed in chapter 3. The main goal of this chapter is to describe how the program works, namely how to insert the input data and how to read the output files.

Chapter 7 is dedicated to the conclusions drawn from this thesis and to the possible future developments. It summarizes the research that was performed in the thesis and presents its main contributions, proposing some possibilities for future developments.

1.4. Scientific papers

The four scientific papers which serve as the base of chapters 2 to 5 of this thesis are the following:

- **Paper I – Described in Chapter 2:** Luís Couto Maia, António Fidalgo do Couto (2013) 'An Innovative Freight Traffic Assignment Model For Multimodal Networks', *Computers in Industry*, 64(2), pp. 121-127;
- **Paper II – Described in Chapter 3:** Luís Couto Maia, António Fidalgo do Couto (2013) 'A Strategic Rail Network Optimization Model For Freight Transport', *Transport Research Record (in press)*;
- **Paper III – Described in Chapter 4:** Luís Couto Maia, António Fidalgo do Couto (2013) 'Validating A Network Optimization Model For Freight: The Case Of The Iberian Peninsula', *Research paper*;
- **Paper IV – Described in Chapter 5:** Luís Couto Maia, António Fidalgo do Couto (2013) 'Assessing The Impact Of Rail Network Improvements On Freight Transport Under Different Scenarios', *Research paper*.

Papers I and II have already been published or accepted for publication in international journals and papers III and IV are research papers which we plan to submit for future publication.

Chapter 2

A STRATEGIC FREIGHT TRAFFIC ASSIGNMENT MODEL FOR MULTIMODAL NETWORKS

2. A STRATEGIC FREIGHT TRAFFIC ASSIGNMENT MODEL FOR MULTIMODAL NETWORKS

2.1. Introduction

Freight transport plays a very important role in the day to day life of any modern society, having a considerable impact in the lives of people and companies. In spite of that, it is a subject that has received considerably less attention by the academia than its passenger counterpart. This is justified, among other reasons, by the fact that freight is a very complex subject being considerably harder to model than passenger transport. Additionally, it is often hard to obtain the data needed to run freight models, due to the reluctance of freight transport companies to share data. Even so, more and more attention has been given to this subject over the years, which has resulted in the development of models that consider freight transport as a separate part of the transport spectrum, instead of being modelled together with passenger transport. This is justified not only by the importance of freight in itself, but also because the reality of freight transport is very different from that of passenger transport, which means that the assignment models need to be different, as well as the type of data that is considered.

The assignment model presented in this chapter contributes to the advance of freight traffic assignment models by including some innovative features, such as the use of different assignment techniques for each type of cargo, and by being relatively light and easy to run in a desktop computer. This is an essential feature of this traffic assignment model, as it will have to be run multiple times in the network optimization model. It is designed to model macro networks with a high aggregation level, namely national or international networks, being a strategic planning model (Crainic and Laporte, 1997). As such, it will not require very detailed data inputs, making use of macro inputs or average values, and the outcome of its application is the estimation of the movement of freight at a national or international scale. Therefore, it is designed to be a useful tool for a variety of strategic planning and policy decisions, which are the kind of tasks for which this type of models are more suited for (Wigan and Southworth, 2006).

The model is designed to model land freight transport, considering the road and rail transport modes, as well as the necessary connector links and intermodal terminals. It contemplates two different types of cargo: intermodal cargo, which represents the cargo that can be easily transferred between different transport modes at intermodal terminals, namely containerized cargo; and general cargo, which represents all the remaining cargo. The model uses different assignment techniques for each type of cargo, with its main innovative feature being the fact that it takes into account both capacity constraints and a variable perception of costs by users, while being much simpler than stochastic equilibrium model. This is a characteristic that distinguishes it from the commonly used all or nothing (AoN), equilibrium or stochastic (multi-flow) models, none of which considers these two factors simultaneously (Jourquin, 2005). This traffic assignment model was developed in C++ programing language, which later enabled it to be integrated with the network optimization model, which was developed using the same programing language.

This chapter includes an application of the model to a fictitious transport network that was developed for this purpose, where its results are analyzed and compared to those obtained by using an AoN technique. This application produces satisfactory results which clearly reflect the fact that the model takes into account both capacity constraints and a variable perception of costs, which distinguishes it from the most frequently used models.

This chapter is structured in five sections. The following section is dedicated to a brief literature overview on traffic assignment models. The subsequent section presents a description of the developed freight traffic assignment model. The fourth section is dedicated to the application of the model on a fictitious network, in order to show its applicability and potentialities. The final section is devoted to the concluding remarks, including suggestions for future research.

2.2. Literature overview

2.2.1.Demand estimation

The demand for freight transport is a necessary input to any freight traffic assignment model and its estimation is a research field that has received considerable attention by the scientific community due to its crucial importance to both planners and operators. Despite being a crucial part of freight transport planning, the scientific knowledge in this area is not consolidated and there are various scientific methods to approach the definition of origin/destination (O/D) tables for freight. The process to estimate the demand for freight transport is usually separated in two steps, following the traditional four step model frequently used to model passenger transport. Those are the production and attraction step and the distribution step (Jong et al., 2004). The production and attraction step is dedicated to the estimation of the amount of freight that is produced and consumed by each generating pole. As for the distribution step, it is the process of estimating where does the freight produced in each pole goes to, and where does the freight attracted by each pole comes from, converting the production and attraction data into an O/D matrix.

The various types of models that have been used to address the production and attraction step are shown in Table 1. These models are divided between those which are based on the observation and analysis of the past reality in order to predict the future, such as time series and system dynamics, and those which try to estimate the production and attraction of each zone based on its economic characteristics, like trip rates and input-output analysis. The latter type of models has the advantage of being usable when there is no data available on the past and present movement of goods in the area under study or when the available data is insufficient. While most demand models resort either to the observation and analysis of past data or to the study of the economic characteristics of each pole to estimate the demand for freight, some models use a combination of both methods (Vilain et al., 2010).

Table 1 – Summary of freight transport production and attraction models (Jong et al., 2004)

Type of model	Advantages	Disadvantages
Time series	Limited data requirements (but for many years)	Little insight into causality and limited scope for policy effects
System dynamics	Limited data requirements Can give land-use interactions External and policy effects variables can be included	No statistical tests on parameter values
Trip rates	Limited data requirements (zonal data)	Little insight into causality and limited scope for policy effects
Input–output	Link to the economy Can give land-use interactions Policy effects if elastic coefficients	Need input–output table, preferably multiregional Restrictive assumptions if fixed coefficients Need conversion from values to tonnes Need to identify import and export trade flows

When the demand estimation is based on past data, that data is frequently flawed, being incomplete or simply not vast enough to answer to the planners' needs. Thereby, the missing data has to be estimated using methods such as log-linear modeling or iterative proportional fitting (Peterson and Southworth, 2010). On the other hand, the models based on economic characteristics also face data challenges due to the fact that input-output tables for the desired kind of goods and for the economic conditions under analysis are frequently not available. Therefore, modelers frequently have to resort to models that relate freight demand with more aggregate economic data such as the gross domestic product (GDP) (Ma et al., 2012; Müller et al., 2012) or the volume of exports (Rao, 1978).

Apart from the more traditional models, there are also some innovative methods that have been used in recent years, including neural networks (Bilegan et al., 2007) or fuzzy methods (Wong et al., 2002). Although these methods have advantages, being well suited for complex and highly dynamic environments, they also have significant shortcomings, being difficult to implement for non-specialist end-users and not mature enough to be used as a widespread method. Another possible way to deal with the problem of production and attraction of freight is a micro approach, which is a suitable option for more detailed micro studies. There are methods which study the generation of freight trips by each firm by establishing the relation between economic factors such as the number of employers in a firm and the amount of freight generated (Iding et al., 2002). The problem with this type of micro approaches is that they are very specific to the type of industry under analysis and cannot be generalized due to the fact that there are large variations between different firms and different sectors of industry.

Table 2 – Summary of freight transport distribution models (Jong et al., 2004)

Type of model	Advantages	Disadvantages
Gravity	Limited data requirements Some policy effects through transport cost function	Limited scope for including explanatory factors and policy effects Limited number of calibration parameters
Input–output	Link to the economy Can give land-use interactions Policy effects if elastic coefficients	Need input–output table, preferably multiregional Restrictive assumptions if fixed coefficients Need conversion from values to tonnes

There are two main types of approaches for the distribution step: gravity models and input-output models. In gravity models (Ashtakala and Murthy, 1988; Levine et al., 2009) the distribution is based on a gravitational principle, where freight is attracted by the different poles according to their size, and that attraction is counterbalanced by the distance, or other impedance factor, between the origin and destination poles. In input-output models, the distribution is based on the economic characteristics of the generation poles, which makes this kind of models more solid and scientific based than gravity models. The drawback is that they need big amounts of data which makes them impractical and difficult to implement, meaning that they are not frequently used.

There is another class of freight demand estimation models worth mentioning, which are general equilibrium models, namely spatial computable general equilibrium (SCGE) models. These models describe the correlation between the economy and the demand for freight with an elevated degree of detail. This type of models has become more popular due to the progress in computational power over the last decades, having been used in some national freight transport studies (Tavasszy, 2006). Although these models present very good results, they need vast amounts of detailed data in order to be successfully applied. This makes them unfeasible for many situations, although they can be excellent tools in freight studies where there is a refined knowledge of the situation under analysis and large amounts of available data.

2.2.2.Network representation

One of the first aspects that have to be defined when creating a traffic assignment model is how to represent the transport network. Transport networks are usually represented by a graph composed by centroids, nodes and different types of links. The links make the connection between two different nodes which are the points where two or more links converge. Centroids are the points where traffic is generated and consumed, representing cities, regions or any other traffic production points. Links can represent different transport modes, such as road and rail, or be transfer or connector links. Transfer links are used to make the connection between two different modes of transport, representing points where the interchange of freight between different transport modes is possible, such as intermodal terminals. As for connector links, they are used to make the connection between the centroids and the nearest nodes, representing the connection between a centroid, such as a city, and the nearest network mode, such as a motorway exit or a rail yard.

It is possible to use different links to represent different vehicles that use the same mode, differentiating for example between large trucks and small trucks. A network of this kind is called a virtual network (Jourquin and Limbourg, 2006) and may use multiple parallel links to represent different types of vehicles that use the same transport infrastructure. These networks have to be carefully modeled due to their special particularities, such as the fact that parallel virtual links that use the same transport infrastructure *have* to share its traffic capacity.

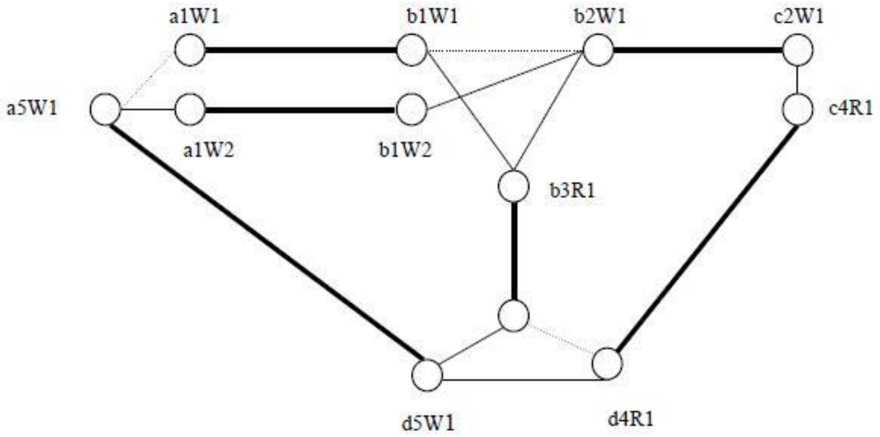


Figure 1 – Example of a virtual network (Jourquin and Limbourg 2006)

Figure 1 displays the representation of a virtual network in which it is possible to see four virtual nodes. Virtual nodes are the set of nodes represented by the letters *'a'* to *'d',* where it is possible to make transfers between the same vehicle and mode, or between different vehicles and modes, if there is that possibility. Each possible transfer is represented by a specific transfer link, which may have a cost equal to zero when there is no change of mode or vehicle. Although the use of virtual networks allows for the representation of different vehicles for each transport mode, conferring more realism and flexibility to the model, it adds to the complexity of the model. This added complexity can be a problem, as it can take a considerable amount of computing time to run a traffic assignment model on these networks.

Generally, each link type is divided into levels which share the same characteristics. The nature of the different levels refers to their general classification, such as a road link being separated between two lane roads and freeways of a rail link being divided between single track and double track lines. Apart from unique attributes like the links' length, all the links from a given category share the same attributes, avoiding the need to model each link individually. These attributes are generally fixed but it is possible to allow for some stochasticity in the definition of certain attributes, such as the link capacity (Xu et al., 2009).

2.2.3.Transport costs

Transport costs are generally estimated using a generic unit, usually the cost per tonne per kilometer. This way costs can be generalized for each link category and applied to all the links, regardless of their length. They usually account not only for the vehicle operation costs but also for time costs and other possible parameters. This definition of cost based on various factors that are deemed to have an impact on the decision of freight transporters is usually referred to as generalized costs.

The process of estimating the generalized vehicle costs involves several important choices and needs to be handled with caution, in order not to ignore important pieces of information and to deliver realistic generalized costs per tonne per km. This is especially true in the case of strategic broad scoped aggregated models, as their high level of aggregation makes it hard to make generalizations, giving that each link category typically encompasses a significant variety of possible realities. There are many different factors that affect the cost of transporting a tonne of freight for one kilometer of road or rail, namely the capacity of the vehicles circulating in the link, their loading factor, their fuel consumption, the purchasing cost of the vehicle and the salary of the driver, among other factors. Although the existence of different link types and link levels should limit those possibilities, by defining the type of vehicles that circulate on each class of links, there is always a considerable variability on the type of vehicles that use it, particularly in the case of strategic aggregate models. In the case of rail transport, the length of freight trains can vary considerably, even within the same link class, especially if there is no differentiation between the maximum allowed train lengths in the different types of rail links. This has an important impact in the cargo capacity of the vehicles, which is closely related with the trains' length, having a very significant impact on the cost per tonne (Janic, 2008). Other important factors that affect the transport costs are the frequency of service and the distance traveled, although these factors are hard to include in a strategic model, being difficult to account for the impact of these factors in the average costs. In freight services where the frequency is higher the costs per tonne per kilometer are lowered, the same happening in services where the traveled distance is longer, as it can be observed in Figure 2. Rail is the mode of transport where these economies of scale, scope and density are more clear (Bereskin, 2001).

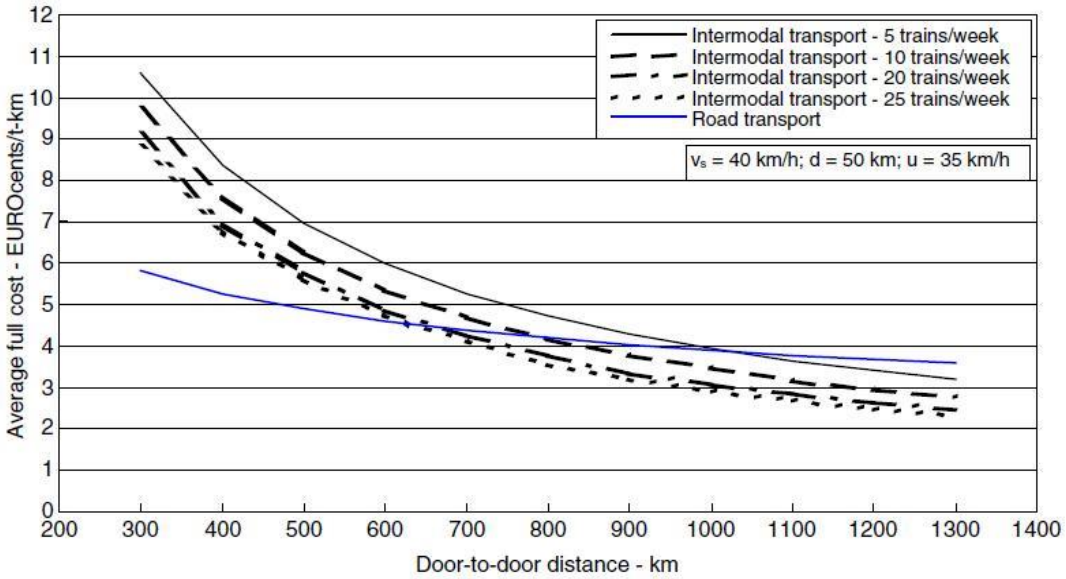


Figure 2 – Impact of the distance traveled and frequency of service on the transport cost per km (Janic, 2007)

There are several other factors that impact the cost of transporting freight, and they may or may not be explicitly considered in a traffic assignment modes, depending on the type of model being used and on the kind of study being performed. Reliability is an important factor in the transport of freight, as shippers are usually more concerned with the planned schedules being respected than with the speed of the transport. Although this aspect is usually taken into account using a generalist approach, by attributing time penalties to links where congestion is elevated and reliability problems are likely, some studies address the issue of reliability in an isolated fashion, particularly in order to understand what happens in situations where the normal behavior of the transport network is disturbed (Janic, 2009).

Apart from the regular road and rail links, it is also necessary to quantify the generalized costs of using other links, such as connectors and intermodal terminals. The cost of using connectors is sometimes assumed to be zero, although in other cases a fixed cost is assumed for this specific step, in order to account for the cost of the last mile. In the case of intermodal terminals, their costs have to be carefully considered, as the modal transfer operations cause a relative large portion of overall intermodal transport costs (Bontekoning et al., 2004; Hanssen et al., 2012; Racunica and Wynter, 2005).

Regarding the time costs, they may represent a significant part of the total generalized costs, as the time it takes to move cargo from one place to another is an important decision making variable. The impact of the travel time varies considerably for different types of freight, being crucial for the movement of some high valuable freight and not particularly relevant for other goods, such as raw minerals or cereals. In any case, the quantification and explicit modeling the cost of time is an indispensable part of any transport model. The time cost has to be quantified in all the operations that are time consuming, including the time that the cargo is on the move and the time that is spent in intermodal terminals and connectors. In the case of strategic models that do not differentiate between various types of cargo, there is a need to consider a reasonable average value. One of the most widely used method to quantify the value of freight transport time is the use of stated preference or revealed preference studies that measure the shippers' perception of the value of time (Bolis and Maggi, 2003) (Kang et al., 2010).

Another group of costs that may be included in traffic assignment models are external costs, such as the impact of freight transport in the environment caused by pollutant emissions or the noise pollution caused by it. The impact of externalities is hard to measure and quantify in monetary terms, but it is a subject that attracts ever more attention, with some studies including an explicit model of external costs (Forkenbrock, 2001). Another possible way to deal with the externalities is by not including them in the assignment model, but to analyze those impacts afterwards, in a post modeling stage.

2.2.4.Traffic assignment techniques

The previous discussed topics have to be interlinked in order to form a traffic assignment model capable of assigning the traffic from an O/D matrix into a transport network. While there is a considerable number and variety of traffic assignment models present in the literature, the traditional assignment techniques can be divided in just four big groups, as it can be seen in Table 3.

Table 3 – Traffic assignment techniques (Jourquin, 2005) – adapted

	Variable perception of costs	
Capacity constraint	No	Yes
No	All or Nothing (AoN)	Stochastic (multi-flow)
Yes	Equilibrium	Stochastic equilibrium

As it can be seen in Table 3, there are two major factors that are used to determine to which of the four traditional techniques does a model belong, namely if there are capacity constraints and if there is a variable perception of costs. Capacity constraint models are those in which the capacity of links is limited, often including time penalties due to congestion when certain limits of traffic are exceeded. As for the variable perception of costs, it reflects whether or not the mode and route choice decisions are made uniquely based on the lowest generalized cost (no variable perception), or if some stochasticity is included, spreading the traffic through different modes and routes. The strategic freight traffic assignment models that exist in the literature usually opt for the all or nothing or equilibrium techniques, with stochastic (multi-flow) models being seldom used. As for stochastic equilibrium models, their use in this area of study has been limited, probably due to their complexity.

In an all or nothing assignment, all the traffic that flows from an origin to a destination is assigned to the least costly route. In an equilibrium technique the traffic is also assigned to the least cost route, but there are capacity limits and penalties for congestion. This means that if capacity is reached on some links, creating problems of congestion, the traffic will tend to avoid these links, by using alternative routes. In a stochastic (multi-flow) assignment the traffic is distributed by more than one possible route according to their total generalized costs, but not considering any capacity limitations. As for the stochastic equilibrium technique, it includes the features of both stochastic (multi flow) and equilibrium assignment techniques, by distributing the traffic by more than one possible route, while also considering capacity limits and penalties for congestion.

An all or nothing technique is the most straightforward technique, which works by simply minimizing the total generalized costs in each O/D pair trip (least cost path) thus minimizing it for the whole network. It is best suited for cases where other assignment techniques are considered too complex, or simply not fit for the proposed approach (Beuthe et al., 2001; Jourquin and Beuthe, 1996). Also, if a modeler chooses not to consider a variable perception of costs, all or nothing can be preferred to an equilibrium approach when, due to the nature of the network being analyzed, it does not make sense to consider congestion.

Regarding the use of the equilibrium assignment technique, it is suited for networks with capacity limits where congestion effects are taken into account by admitting cost penalties when traffic is close to the capacity (Crainic et al., 1990; Guélat et al., 1990; Jourquin and Limbourg, 2006). Its solution is based on the Wardrop equilibrium (Wardrop, 1952), in which all vehicles choose their optimal route, so that no vehicle can improve its travel time by unilaterally changing routes. To solve this problem, the most commonly used algorithms are the incremental assignment, the method of successive averages and the algorithm of Frank-Wolfe (Jourquin and Limbourg, 2006).

As for the stochastic (multi-flow) approach, it differs from an all or nothing assignment in the fact that is does not assign all the traffic to the least cost path, distributing it by the different alternative routes. While this technique does not consider capacity constraints, it distributes the traffic by different routes, which makes it useful for models where congestion is not important, as it is the case in most intercity freight assignment models (Jourquin, 2005). Contrarily to all or nothing or equilibrium techniques, the use of stochastic assignment techniques ensures that the path with the least generalized costs never receives the totality of the traffic, meaning that other paths are also used. This is a valuable feature when dealing with strategic aggregated models, where the generalized transport costs are just an estimation of the average costs, in that it does not consider the mere minimization of the generalized costs as an undisputable deciding factor. This is additionally justified by the fact that there are many factors that are almost impossible to incorporate in the generalized costs of a strategic model, but have a decisive influence on the modal or route choice, such as the shipment size (Abdelwahab and Sargious, 1992) (Abdelwahab, 1998), the frequency of service (Shinghal and Fowkes, 2002), the service quality (Zlatoper and Austrian, 1990) (Andersen and Christiansen, 2009) and the existence of an integrated door to door logistic chain (Vanek and Smith, 2004). The Logit formulation has consistently been chosen to address the distribution of traffic, due to its versatility and convenience (Oum, 1979) (Southworth and Peterson, 2000) (Jourquin, 2005) (Tsamboulas and Moraitis, 2007). The multinomial Logit model is a function that calculates the percentage of trips that use each alternative route option (Tsamboulas and Moraitis, 2007):

$$P_{ij}^{k} = \frac{\exp(-\mu \cdot C_{ij}^{k})}{\sum_{k=1}^{n} \exp(-\mu \cdot C_{ij}^{k})} \qquad (1)$$

where:

P_{ij}^{k} - percentage of trips realized from origin *i* to destination *j* through route *k*

C_{ij}^{k} – Generalized cost of transport from origin *i* to destination *j* through route *k*

μ - parameter which determines the impact that the cost differentials have on the percentage of trips assigned to each alternative

The μ parameter affects the percentage of traffic that uses each route, as the use of the least costly route increases for higher values of the μ parameter. The Logit function is practical and appealing for its analytical convenience (Cook et al., 1999), and it can be used in any traffic assignment model to estimate the share of traffic that uses each one of the alternative route options.

Although there are many different traffic assignment models present in the literature, with many being built for just one specific work, there are at least two major freight traffic assignment models that are worth mentioning, due to their importance and extensive use. Those are STAN, which was developed in 1990 in Canada (Crainic et al., 1990; Guélat et al., 1990), making use of an equilibrium assignment technique, and the NODUS software, which was developed in Belgium a few years later (Beuthe et al., 2001; Jourquin, 2005; Jourquin and Beuthe, 1996; Jourquin and Limbourg, 2006) and that has been employed using all of the three most common assignment techniques: all or nothing, stochastic (multi flow) and equilibrium.

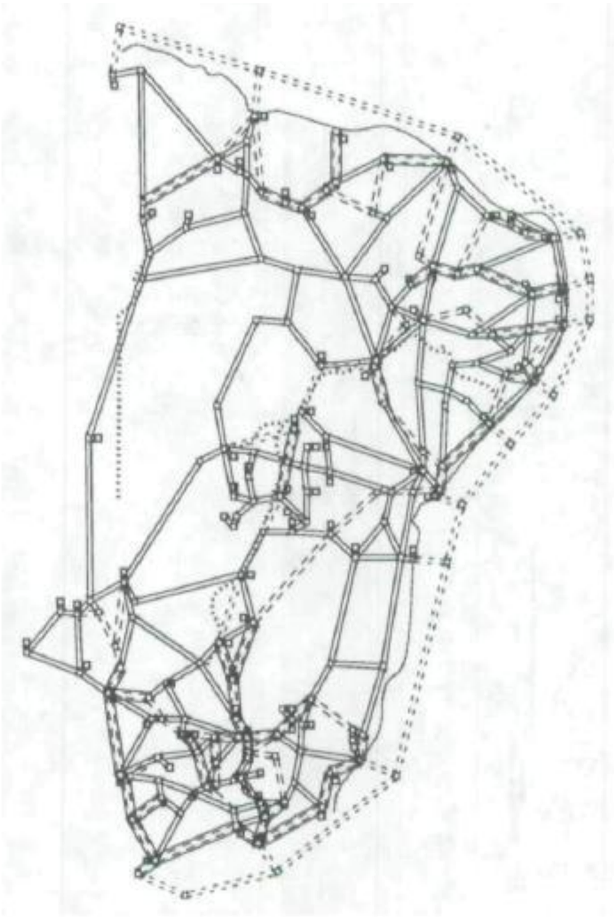


Figure 3 –Example of a network in STAN (Crainic et al., 1990)

While the most commonly used assignment procedure is the equilibrium, the fact that various techniques were used with the NODUS software shows that each one has its own advantages, which should be considered when choosing the type of technique to use. Apart from the strategic models, there are other types of models that address more specific problems and situations, using techniques such as micro simulation (Weidner et al., 2009). There are also some strategic models that incorporate a more detailed micro method to deal with a particular part of the modeling process (Southworth and Peterson, 2000).

Freight flows are frequently unbalanced, which causes problems in the transport chain, often causing freight vehicles to run empty on one way. These empty trips have a significant economic impact: the cost per tonne of transporting freight in only one way and returning empty is higher than in the case of loaded trips on both ways, which only happens is flows are balanced or if a different commodity can be transported on the way back. Also, it is important to know all the traffic flows (empty or loaded) in order to assess the total level of traffic that passes through each link. Although many traffic assignment models do not consider this problem explicitly, there is some relevant literature where this problem is comprehensively addressed (Holguín-Veras and Thorson, 2003), or where at least given some attention (Fernandez et al., 2004). One way to estimate the share of empty trips is to assess it based on the flow unbalance and on the distance between the O/D pair under analysis (Holguín-Veras and Thorson, 2003).

To finalize, it is important to remember that while traffic assignment models are very useful tools, being well suited for several planning tasks, they have important limitations. There is evidence that strategic freight traffic assignment models are usually not up to the task of forecasting future freight activity, although they can be very useful tools for a variety of planning and policy decisions where their limited forecasting capacity is not needed (Wigan and Southworth, 2006). Even so, most of the errors derived from the use of traffic assignment models are caused by inadequate or too scarce data or to an incorrect interpretation and modeling of the system.

2.3. Description of the traffic assignment model

2.3.1. General characteristics

The developed model is a strategic freight traffic assignment model, designed to model macro networks at a national or international level of planning. It considers a basic differentiation between two types of cargo: intermodal cargo and general cargo. The former represents the containerized cargo that can be easily transferred between different transport modes at intermodal terminals, while the later represents all the remaining cargo. It is designed to model inland freight transport, considering road and rail transport. Apart from those two major transport modes, the model also includes other types of links such as connector links that link the road and rail network to the centroids representing each region, and intermodal terminals that establish the connection between the road and rail modes. The connector links are very flexible, with its characteristics being freely defined by the modeler. This enables them to be used in situations in which it is necessary to introduce a very specific link, such as the representation of a congested rail node in the middle of a rail network.

Table 4 –Link attributes

	Variable	Definition
1	**ID**	Link identification number
2	**Type**	Link type
3	**Improvable**	Variable that defines if the link is improvable or not
4	**LinkStatus**	Link level of quality
5	**Startnode**	Identification number of the node where the link originates
6	**Endnode**	Identification number of the node where the link ends
7	**Length**	Link length
8	**Speed_gen**	Vehicle speed for general cargo
9	**Veicap_gen**	Vehicle cargo capacity for general cargo
10	**Veicost_gen**	Total vehicle cost per km for general cargo
11	**VOT_gen**	Value of time for general cargo
12	**Veicost_int**	Total vehicle cost per km for intermodal cargo
13	**Capacity_each_way**	Link traffic capacity in each direction
14	**CO2_km_gen**	Vehicle CO2 emission per km for general cargo
15	**Veicap_int**	Vehicle cargo capacity for intermodal cargo
16	**Bidirectio**	Variable that defines if the link is bidirectional or not
17	**Speed_int**	Vehicle speed for intermodal cargo
18	**CO2_km_int**	Vehicle CO2 emission per km for intermodal cargo
19	**VOT_int**	Value of time for intermodal cargo

Each link has a specific set of attributes that define all of its characteristics, as it can be seen in Table 4. Some of these attributes are relative to the physical characteristics of the link, namely its type and level of quality, the definition of the nodes from where it originates and where it ends, its length, and its traffic capacity in each direction. There are also two binary variables that define if the link may be improved to a better level and if it is unidirectional or bidirectional. The remaining attributes are related to the vehicles and type of cargo that uses the link, being separated for general cargo and intermodal cargo. This enables the use of different vehicles for each type of cargo and even to the use of different cargo units, such as the use of tonnes for general cargo and twenty foot equivalent units (TEU's) for intermodal cargo. These attributes include the speed, cargo capacity, total cost per km and CO2 emissions per km of the vehicles in that link, as well as the monetary value attributed to the time. Each of these attributes reflects the average conditions for each link, and should comprise all the factors that may affect that particular attribute. Thereby, the vehicle costs should include all the costs associated with operating a vehicle, including fuel and maintenance costs, the cost of the driver, and the cost of using the infrastructure, namely tolls or rail infrastructure charges.

It is important to stress that intermodal terminal and connector links are only representations of different types of transport connections and do not represent real links with real vehicles using them. Due to that, they display artificial characteristics which are meant to represent the time and costs of using those links. This may be accomplished by considering that the vehicle capacity is equal to 1 and adjusting the remaining link attributes in order to achieve the desired values for the time and cost for those links. This is a continuous model which does not consider individual vehicles, but rather undifferentiated amounts of cargo that use the different links. Thereby, the freight flows are represented as the amount of typical vehicles, and may not be a whole number.

The computation of the generalized transport costs per tonne on each link is calculated based on the link attributes, being constituted by a vehicle cost component and a time cost component:

$$p_i^k = \frac{l_i * q_i^k}{r_i^k} + \frac{l_i}{s_i^k} * v_i^k, \quad \forall i \in I, \quad k \in J \qquad (2)$$

with:

$I = set \quad of \quad all \quad links$

$J = \{Gen, Int\}$

where:

Gen = general cargo

Int = intermodal cargo

p_i^k = generalized cost per unit of cargo on link i, for cargo type k

l_i = length of link i

q_i^k = vehicle cost per distance on link i, for cargo type k

r_i^k = vehicle capacity on link i, for cargo type k
s_i^k = vehicle average speed on link i, for cargo type k
v_i^k = value of time on link i, for cargo type k

This cost is given in units of cost per tonne, and corresponds to the generalized costs incurred by a tonne of freight when passing through a given link. Based on the defined generalized costs for each link, it is possible to calculate the least expensive path between any given pair of nodes, which can be computed using a shortest path algorithm. The algorithm that is employed in this model is the Floyd-Warshall algorithm (Floyd, 1962) with path reconstruction, which computes the value of the least expensive path (shortest paths) between all the nodes, as well as the path in itself (the links used in each path). This algorithm was chosen due to the fact that it is a simple and efficient algorithm, being frequently used to tackle this type of problems.

2.3.2. Assignment techniques

The main innovative feature of this traffic assignment model is the use of two different assignment techniques for the two types of cargo. This is justified by the fact that intermodal cargo is flexible, being able to change from road to rail transport with relative ease, by using intermodal terminals. This does not happen with general cargo, which is thereby limited to using the same transport mode in each trip. The main features of the assignment process are resumed in Figure 4.

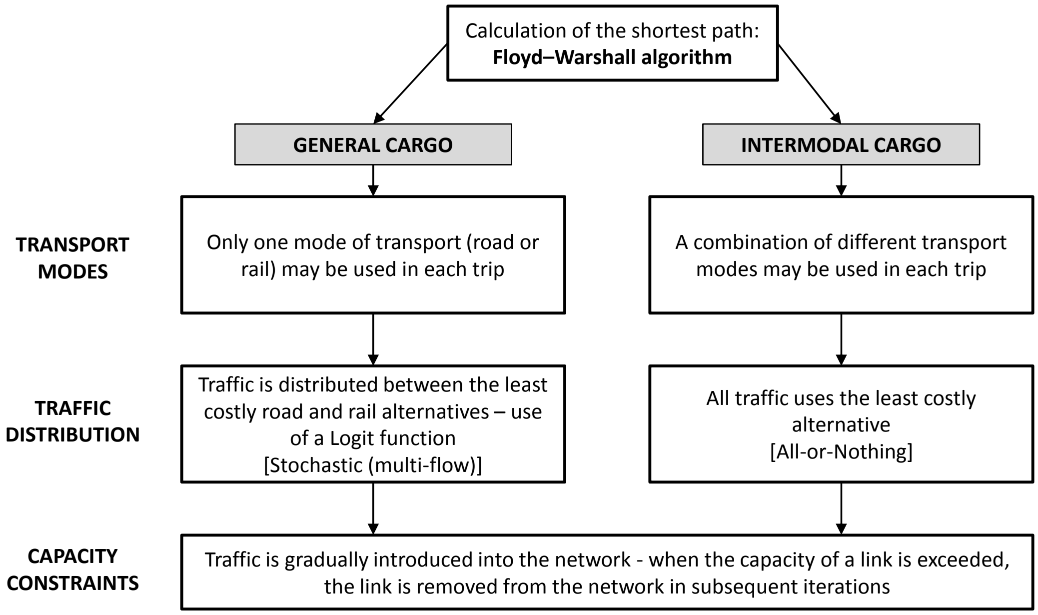


Figure 4 – Assignment process

In the case of general cargo, as each trip may only use one mode of transport, there is a clear mode choice decision between road and rail transport. This feature enables the use of an assignment technique that distributes the traffic by these two transport modes, which is coherent with the nature of this model, given that there is always some traffic distribution between concurring transport modes at this macro level of planning. Due to that, a stochastic (multi-flow) technique is used for the assignment of general cargo traffic, distributing the traffic between the two concurring transport modes based on the comparison between the least costly paths using road and rail transport. By distributing the traffic by the two available transport modes, the stochastic technique implicitly considers that the simple minimization of the total generalized costs is an imperfect measure of the quality of each transport alternative, which is particularly true in the case of strategic models. The imperfection of the generalized costs as a way to assess the best route is due to the existence of various issues other than the cost that affect the choice of transport mode, such as the suitability of each transport mode to each shipment or the routine of the shippers. These factors often cause a significant part of the traffic not to flow through the least costly route. The distribution of traffic through the road and rail modes is calculated using a logit function, as seen in Equation (1). It gives the percentage of traffic that uses each mode, being dependent on the generalized cost of each alternative and on the value of the μ parameter, which has to be calibrated in a case by case basis. In case there is only one available transport mode for a given O/D pair, all the traffic will flow through the least costly path of the available transport mode.

In the case of intermodal cargo, as it can use more than one transport mode per trip, there is no clear mode choice decision between road and rail transport. Also, due to the strategic nature of the model and to the diverse nature of transport networks, there is usually no clear route choice decision between two or more alternative routes. Thereby the model uses a technique that assigns the traffic to the least costly path between the origin and the destination, allowing for the combination of various transport modes. This is performed using an AoN technique, allowing for the indifferent use of road, rail and intermodal terminal links, as long as they are part of the absolute least expensive path.

It should be stressed that neither the AoN nor the Stochastic (multi-flow) consider any traffic capacity limitations. Given that this is a strategic model that only considers the main traffic generation regions and long distance links, it would be reasonable not to consider any capacity constraints, given that congestion is mostly observed in and around urban areas (Jourquin, 2005) and not on intercity links. There are nonetheless long distance transport links that have congestion problems, particularly rail links, due to the fact that their capacity is relatively rigid. Intermodal terminals may also be affected by capacity problems, affecting intermodal cargo. Due to that, the model includes a mechanism that allows for the imposition of capacity limits on the different links, according to needs of each network. These capacity constraints are not imposing using an equilibrium technique, which would be quite demanding in terms of computational power, slowing the model.

Since this model has to be as fast and light as possible, due to the fact that it will have to be run multiple times in the network optimization model, a much simpler and lighter technique was adopted.

As it can be seen in Figure 4, the imposition of capacity limits is performed using a method based on the gradual introduction of traffic in the network. The total freight flow is gradually inserted into the network in as much iterations as chosen by the user. After the introduction of traffic in each iteration, the model verifies if there are any links which have exceeded their capacity limits, removing them from the network in subsequent iterations. If any new link has been removed from the network, the traffic has to be assigned to the modified network for subsequent iterations. This process is repeated until all the traffic is assigned to the network. This is a relatively simple process which satisfactorily manages to impose traffic limits on transport links while being faster than an equilibrium technique.

It is important to notice that although this is a freight model, the total traffic capacity of each rail link has to be shared between the passenger and freight trains that use it. Thereby the procedure that is adopted in the model is to consider that the rail capacity available for freight trains is the total link capacity minus the traffic of passenger trains, reflecting the fact that passenger trains usually have priority over freight trains. This is achieved by assigning the O/D matrix of all passenger trains to the network, using an AoN assignment technique based on the minimization of the total distance for each trip. Apart from the issue of passenger trains, there is another matter that affects the capacity of rail links, which is the existence of empty freight trips. This is a complex subject that is difficult to contemplate in a macro model like this one. Even so, it is possible to consider part of the unbalancing effect caused by empty trips, as empirical evidence shows that even in cases of extreme unbalances in commodity flows, the vehicle flows in both directions tend to be equal (Hautzinger, 1984). This is achieved in the model by considering that if the traffic capacity of a link has been reached in one direction, the other direction is also considered as being full. This is done in order to simulate the empty vehicles that should be balancing the flow and which are not explicitly considered in the model, avoiding the occurrence of situations where a link is full in one direction but still has free capacity in the other, which is not realistic.

2.3.3.Innovative aspects of the model

This traffic assignment model for freight transport is innovative in the fact that it uses two different assignment techniques for the two types of cargo, and that it considers both capacity constraints and a variable perception of costs. This mix of techniques, combined with the fact that the capacity limits are imposed using a gradual introduction of the traffic in the network and not a congestion effect means that the model does not use a traditional assignment technique. The consideration of both capacity constraints and a variable perception of costs while being much simpler and faster to run than a stochastic equilibrium model is the most valuable asset of the model. This is particularly important, given that this simplicity makes the model relatively light and fast to run, which is very important in the context of a network optimization model, where the assignment model has to be run multiple times.

2.4. Application of the model on a fictional network

2.4.1.Description of the fictional network and considered scenarios

The developed traffic assignment model was applied to a fictional network in order to test and evaluate its performance. The use of a fictional network was justified by the fact that it is much easier to study the behavior of the model in a compact network that in a large real world network. Also, as the network was developed specifically for this purpose, its design allows for the model to display its full potentialities.

The network is relatively simple, containing six traffic generating poles (centroids), several road, rail and connector links and one intermodal terminal, which connects nodes 11 and 19. All the rail and road links share the same characteristics and there are only two types of connector links: centroid to road and centroid to rail. The values that were used for the various link attributes were defined by us for this specific application, being merely indicative reasonable values.

Table 5 – Adopted link characteristics

	Symbol	Rail connector	Road connector	Road links	Rail links	Intermodal terminal
1	**ID**	-	-	-	-	-
2	**Type**	0	0	1	2	3
3	**Improvable**	-	-	-	-	-
4	**LinkStatus**	1	2	1	1	1
5	**Startnode**	-	-	-	-	-
6	**Endnode**	-	-	-	-	-
7	**Length**	4	1.5	-	-	2.5
8	**Speed_gen**	10	10	80	40	10
9	**Veicap_gen**	1	1	20	1000/1200	1
10	**Veicost_gen**	1	1	1.5	35	1
11	**VOT_gen**	1	1	1	1	1
12	**Veicost_int**	1	1	1.5	35	1
13	**Capacity_each_way**	-	-	-	-	-
14	**CO2_km_gen**	-	-	-	-	-
15	**Veicap_int**	1	1	20	1000/1200	1
16	**Bidirectio**	1	1	1	1	1
17	**Speed_int**	10	10	80	40	10
18	**CO2_km_int**	-	-	-	-	-
19	**VOT_int**	1	1	1	1	1

As it can be seen in Figure 5, the network consists of two groups of centroids which are located to the north and to the south of a division in the middle which is crossed by a road link and two rail links. One of the rail links represents a rail tunnel while the other is an old rail line that bypasses the tunnel but is considerably longer. Although the rail tunnel represents a better alternative, it may have a limited capacity, leaving the old line as only rail alternative with no capacity problems. The total generalized transport costs per tonne for each link were calculated based on the adopted link characteristics displayed on Table 5.

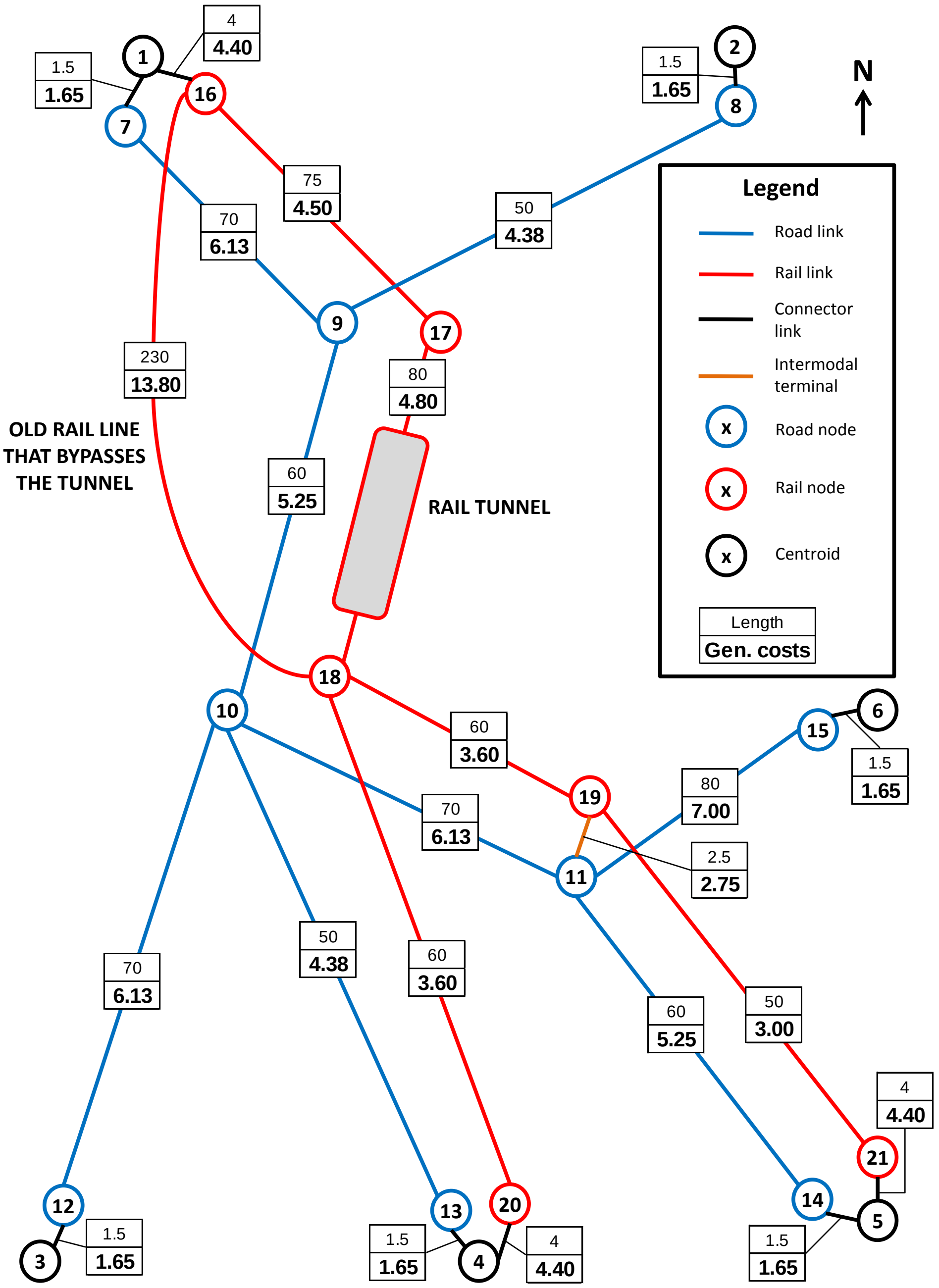


Figure 5 – Fictional network configured for scenarios 1 and 2

The traffic assignment model will be applied to the network considering four different scenarios where the capacity of freight trains as well as the traffic capacity of the rail tunnel will vary.

Table 6 – Considered scenarios

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Traffic capacity of the rail tunnel (vehicles)	-	35	-	35
Vehicle capacity on rail links (ton)	1000	1000	1200	1200

The values for the freight demand, traffic of passenger trains and the tunnel capacity were freely defined by us and are meant to represent daily values. The higher vehicle capacity on scenarios 3 and 4 means that the freight trains will be able to carry more cargo for the same transport cost, as the vehicle costs are the same for all the scenarios. Thus, rail freight will be more competitive in those two scenarios. For the sake of simplification, only one direction of traffic flow was considered (north to south), as it can be seen in Tables 7, 8 and 9.

Table 7 – General cargo O/D matrix

O/D	General cargo [ton]					
	1	**2**	**3**	**4**	**5**	**6**
1	-	-	-	50000	30000	25000
2	-	-	15000	10000	20000	15000
3	-	-	-	-	-	-
4	-	-	-	-	-	-
5	-	-	-	-	-	-
6	-	-	-	-	-	-

Table 8 – Intermodal cargo O/D matrix

O/D	Intermodal cargo [ton]					
	1	**2**	**3**	**4**	**5**	**6**
1	-	-	-	20000	-	10000
2	-	-	-	-	-	-
3	-	-	-	-	-	-
4	-	-	-	-	-	-
5	-	-	-	-	-	-
6	-	-	-	-	-	-

Table 9 – Passenger trains O/D matrix

O/D	Traffic of passenger trains [vehicles]					
	1	**2**	**3**	**4**	**5**	**6**
1	-	-	-	7	3	-
2	-	-	-	-	-	-
3	-	-	-	-	-	-
4	-	-	-	-	-	-
5	-	-	-	-	-	-
6	-	-	-	-	-	-

2.4.2. Application results and discussion

The model was run on the fictitious network for all the four scenarios, considering all the input conditions that were described above. The traffic was introduced in the network in 20 iterations and two different values were used for the μ parameter in the Logit Equation (1): *0,1* and *0,5*. This was done in order to simulate a smaller or bigger impact of the generalized costs in the modal distribution for general cargo, respectively. The values *0,1* and *0,5* were used due to the fact that they present an ample variation of this parameter while delivering reasonably balanced distributions of traffic.

In order to have some point of comparison for the developed model, the same network was also run in a commercially available traffic assignment software (STAN software) using an AoN technique which makes no distinction between the two types of cargo. Also, as the AoN technique does not consider capacity limits, the only scenarios that could be run using that technique where those where there are no capacity limits on the rail tunnel, which are scenarios 1 and 3. The obtained results are presented in Table 10, containing the values of the freight flow on some key links. The detailed results from the application of the assignment model to all the scenarios can be consulted in annexes A1 to A8.

Table 10 – Obtained results

	Scenario 1			**Scenario 2**			**Scenario 3**			**Scenario 4**		
Total Flows [vehicles]	**AoN**	**Model** μ**=0,1**	μ**=0,5**	**AoN**	**Model** μ**=0,1**	μ**=0,5**	**AoN**	**Model** μ**=0,1**	μ**=0,5**	**AoN**	**Model** μ**=0,1**	μ**=0,5**
Flow on link 9-10: **Road link**	8250	7864	8231	-	7994	8410	6500	7229	7205	-	7588	7947
Flow on link 17-18: **Rail Tunnel**	30.0	37.7	30.4	-	26.4	25.8	54.2	42.0	42.4	-	25.2	25.5
Flow on link 16-18: **Old rail line that bypasses the tunnel**	0	0	0	-	8.7	1	0	0	0	-	10.8	4.6
Flow on link 19-11: ***Intermodal terminal***	*0*	*0*	*0*	-	*0*	*0*	*35000*	*10000*	*10000*	-	*6000*	*6000*

In scenario 1, road transport offers the least costly alternative for all the O/D pairs, except from centroid 1 to centroid 5. Due to that, intermodal transport will not use rail transport, and the intermodal terminal will not be used. The results obtained using an AoN technique are almost identical to those of the presented model for a value of μ equal to 0,5, which is explained by the fact that for a higher value of μ there is less distribution of traffic, with the results being closer to those obtained using an AoN assignment. With a lower value of μ, the traffic is more evenly spread between the road and rail modes, which seems more plausible and reflects the fact that this model does not consider the minimization of the generalized costs as the undisputable deciding factor for the mode and route choice.

As for scenario 2, it reflects the same situation as before, but with the difference that there is a capacity limit of 25 vehicles on the rail tunnel, which reflects the total capacity of 35 vehicles minus the 10 passenger trains that use it. This limitation forces a decrease on the traffic of the rail tunnel and an increase on the alternative road link. The old rail line that bypasses the tunnel is also used by some trains, which is justified by the distribution of general cargo traffic by the available road and rail alternatives. Although the rail route that bypasses the tunnel is never the least costly path, it becomes part of the least expensive rail route when the rail tunnel is used to capacity, and is thereby used by some traffic, especially for lower values of μ. The fact that the traffic on the rail tunnel is slightly above the limit of 25 vehicles is justified by iterative introduction of the traffic in the network, as a link that is used to capacity is only removed from the network in the subsequent iteration after it reaches its capacity.

In scenarios 3 and 4 there is a greater use of rail transport, as the higher train capacity makes this mode of transport more competitive. Regarding scenario 3, the first noticeable aspect is that the AoN assignment makes a greater use of the rail tunnel link, due to the fact that all the cargo that goes from centroids 1 to 6 (35000 tonne) may use the rail tunnel in combination with the intermodal terminal, which is the least costly path. In the case of the developed model, as it prevents general cargo from using intermodal terminals, the only cargo traveling between centroids 1 and 6 that can use the intermodal terminal is intermodal cargo (10000 tonne). This implies that all the general cargo (25000 tonne) has to use road transport, as there is no rail connection to centroid 6. The fact that the results obtained with different values of μ are much closer than in scenario 1 is justified by the fact that the transport costs between the road and rail alternatives are not very different, with road transport being less costly on some routes and rail on others.

Finally, in regard to scenario 4, the existence of a capacity limit on the rail tunnel benefits road transport, as it happened in scenario 2. Even so, as rail is more competitive than in scenario 2, there is a much greater use of the old rail line that bypasses the tunnel, as the distribution of generalized cargo by the possible road and rail alternatives directs a considerable amount of cargo to the old rail line. The rail capacity in the rail tunnel also limits the use of the intermodal terminal, as the intermodal cargo that is unable to use the tunnel being rerouted via road transport. The use of a higher value of μ has the same effect as in scenario 2, by considerably reducing the traffic on the old rail line, which is due to the fact that that route is still significantly more expensive than the road alternative.

The application of the model on this fictional network delivered good results which highlight the innovative characteristics of the model. Also, the comparison with an AoN assignment technique provides a term of comparison and contributes to the validation of the model. The fact that the old rail line that bypasses the tunnel was used in scenarios 2 and 4 is justified by the combination of assignment techniques used in the model, which take into account for both capacity constraints and a variable perception of costs. This would not happen in the traditional AoN, equilibrium or stochastic (multi-flow) assignments techniques under the same conditions, namely a capacity limit on the rail tunnel for the equilibrium model and a traffic distribution between the least costly road and rail alternatives for the stochastic technique. The equilibrium model would always use the road alternative after the tunnel reached full capacity and the stochastic (multi-flow) technique would simply distribute the traffic between the rail tunnel and the road alternative, being unable to impose capacity limits. In our opinion, the outcome obtained by the presented model is more realistic, as it considers that the old rail line is used by some trains. Those trains represent the cargo that did not obtain one of the available tunnel slots but continued to opt for rail transport, which may be justified by various reasons.

2.5. Conclusions

This chapter presents a strategic freight traffic assignment model designed to model road and rail transport at a national or international level. It considers two different types of cargo: intermodal cargo, which represents the containerized cargo which can be easily handled at intermodal terminals and general cargo, which represents all the other cargo. It is designed to model surface transport, considering road and rail links, intermodal terminals and different types of connector links. The model uses different traffic assignment techniques for each type of cargo, resulting in a global assignment process which is innovative in the fact that it considers both capacity constraints and a variable perception of costs. This distinguishes it from the traditionally used all or nothing, equilibrium or stochastic (multi-flow) techniques, none of which considers these two factors simultaneously. While stochastic equilibrium models also consider both capacity constraints and a variable perception of costs, they are much more complex than the presented model and are very demanding in terms of computational power. The relative simplicity of this model is one of its most valuable assets, as it makes it relatively light and fast to run, which is very important in the context of a network optimization model, where the assignment model has to be run multiple times.

The application of the model on a fictional network produced satisfactory results, which clearly reflected the fact that the model takes into account both capacity constraints and a variable perception of costs. In our opinion, the way in which the model assigns the traffic to the network is more realistic than in the traditionally used equilibrium and stochastic (multi-flow) models, as it considers the use of inferior quality links of the same transport mode after there is no more capacity left in the main links.

Although the developed model is a useful tool for the strategic planning of freight transport, as it was shown by its application, we consider that there is still room for future improvement on four significant aspects. The first one is the assignment technique used for intermodal cargo, which may be improved by considering some sort of traffic distribution between various possible routes. The most critical aspect of this improvement would be the development of a system for the creation of one or more feasible alternative routes. The second aspect which may be improved in the future is the number of different cargo types that are contemplated in the model. It would be interesting to separate general cargo into various sub-categories, in order to have a more refined model. Also, different assignment techniques could be considered for the various sub-categories of cargo, according to their characteristics. The third improvable aspect is the estimation of empty trips, which is something that is not contemplated in this model. The explicit estimation of empty freight trips would entail the development of a method to estimate the amount of empty trips based on the flow imbalances between the different O/D pairs. This would make the model more realistic, improving the accuracy of its traffic estimations. To finalize, the fourth improvement possibility for the future is the integration of the model with a passenger traffic assignment model. This would lead to the creation of a comprehensive strategic surface transport model for both passenger and freight transport, which would deliver a comprehensive picture of all the long distance traffic using the road and rail networks, being a valuable strategic planning tool.

Chapter 3

A STRATEGIC RAIL NETWORK OPTIMIZATION MODEL FOR FREIGHT TRANSPORT

3. A STRATEGIC RAIL NETWORK OPTIMIZATION MODEL FOR FREIGHT TRANSPORT

3.1. Introduction

Freight transport is an activity with a significant economic impact, being considerably more complex than passenger transport. This added complexity is mainly justified by the multiplicity of goods transported and the intricacy of the supply chains involved in the transport of freight. Thus, it is important to study freight transport using models specifically made for it, in order to account for its distinct characteristics. Also, the type of network investments needed to improve the conditions for freight transport can be considerably different from those aimed at improving the conditions for passenger transport. Due to that, the network optimization model that is presented in this chapter has been developed specifically for this type of transport.

The presented network optimization model assigns traffic to the networks under study using the strategic traffic assignment model that was described in chapter 2, being designed to model macro networks at a national or international level of planning. Due to its macro nature, the model contemplates only the main road and rail transport links, as well as connectors and the main intermodal terminals. The strategic nature of the previously developed traffic assignment model make it particularly suited for the planning and policy decisions that will be addressed by the network optimization model.

The model is quite flexible and innovative, allowing for both upgrades in the quality of existing links, as well as the construction of new ones. Also, it allows for the planner to define the type of improvements it wants to test, not having a limit on the number or variety of improvement solutions. This is achieved by defining a set of possible link levels for each link type, according to the preferences of the planer, including the mere possibility of building a link. As for the assessment of the quality of each network improvement solution, it is based on the reduction of the total generalized costs and CO2 emissions, with the weight given to each of those parameters being defined by the planner according to the conditions of the situation under study. As for the optimization process, it uses a local search heuristic which tries to meet a balance between efficiency and effectiveness, by delivering good solutions in a reasonable computing time.

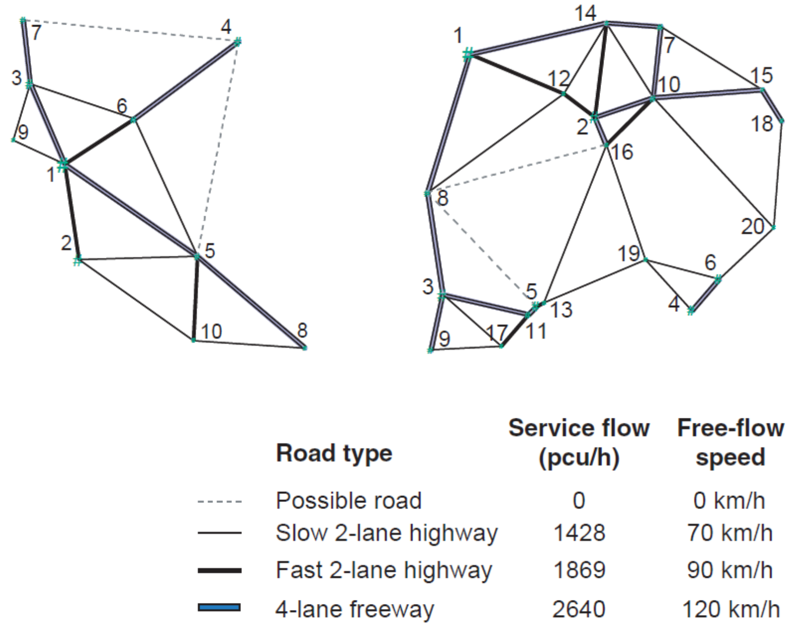
The network optimization model is applied to two transport networks under different scenarios, as a test to the applicability and flexibility of the model. This application produces satisfactory results which highlight the versatility of the model.

This chapter is structured in five sections. The following section is dedicated to a brief literature overview on network optimization models. The third section presents a description of the developed network optimization model. The fourth section is dedicated to the application of the model on two different networks, under different scenarios. The final section is dedicated to the concluding remarks, including suggestions for future research.

3.2. Literature overview

3.2.1. Generation of network improvement scenarios

The first step in the development of a network optimization model is to define a method for the creation of feasible network improvement solutions, combining various individual improvement operations in order to form a comprehensive improvement solution. A network improvement operation is an investment in the transport infrastructure that improves the quality of a given link. It may represent an upgrade in the quality of an existing link or the construction of a new link from scratch. Most of the existing network optimization models are focused on either building new transport links or improving existing infrastructures, although there are some models which are more flexible, allowing for both improvements to existing links as well as the construction of new ones from scratch (Santos et al., 2008). Each network improvement operation must have an associated cost, in order to quantify the investment needed to perform a proposed network improvement scenario. The definition of the possible network improvement operations and their associated costs are the base conditions for the creation of all network improvement solutions. After having those base conditions well defined, it is possible to create a method for the generation of improvement solutions, which can be made at random, or following a set of rules defined by the planner. The amount of investment needed for each solution is calculated based on the sum of the cost of all the performed improvement operations.



	Road type	Service flow (pcu/h)	Free-flow speed
-----	Possible road	0	0 km/h
—	Slow 2-lane highway	1428	70 km/h
━	Fast 2-lane highway	1869	90 km/h
▬	4-lane freeway	2640	120 km/h

Figure 6 – Network improvement possibilities (Santos et al., 2008) - adapted

Figure 6 display a good example of a network with clearly defined network improvement operations, which allows for both the improvement of existing links as well as the construction of new ones. As it can be seen in the legend, there are four possible discrete link levels, starting at zero, which corresponds to the mere possibility of building a link. The three remaining link levels correspond to different types of roads, with diverse technical characteristics. The associated costs per kilometer for each possible improvement operation are displayed in Table 11.

Table 11 – Link construction and upgrading cost (Santos et al., 2008)

	To		
From	*Slow Two-Lane Highway*	*Fast Two-Lane Highway*	*Four-Lane Freeway*
Possible road	1	2	3
Slow two-lane highway	-	1.5	2.5
Fast two-lane highway	-	-	2

Based on the exemplified network conditions and improvement costs, it is possible to generate different network improvement solutions and to calculate the investment needed to perform them.

3.2.2. Optimization process

The aim of a transport network optimization model is to find the best way in which to allocate investment for the improvement of a transport network. The goal of the optimization process is to improve the parameters that are used to assess the quality of each solution. Many different parameters can be used in order to assess the quality of each solution, with the most widely used being the minimization of the total generalized transport costs. Other possible parameters include the minimization of the environmental impact caused by transport vehicles, the robustness of the network under uncertain conditions (Santos et al., 2009; Yang et al., 2011) and the equity of the territorial accessibility (Santos et al., 2008). There are various possible methods to quantify each of those parameters, using more or less complex formulations. The environmental impact can be measured based on estimations of the average amount of pollutants emitted by transport vehicles, or on more complex methods that also take into account for the local impact of transport, considering factors such as the impact of the noise pollution. As for the robustness of the network, it can be quantified as the existing spare capacity of the network, or using more complex methods such as the network vulnerability to unpredictable scenarios. Regarding the equity of the territorial distribution, it is usually assessed using the Gini coefficient, although other methods such as the definition of minimum accessibility conditions can also be used.

Most of the research found in the literature on the subject of network optimization was performed using two types of models: the discrete network design problem (DNDP) (Arnold et al., 2004; Chen and Alfa, 1991; Santos et al., 2010; Yamada et al., 2009) and the continuous network design problem (CNDP) (Zhang and Lu, 2008). The DNDP approach tends to concentrate on the addition of new links, with the discrete decision being whether or not to construct a new link. As for CNDP problems, they are generally used for the improvement of existing links, using a continuous approach to measure the quality of links. It is also possible to use a discrete approach allowing for both the addition of new links and the improvement of existing links, as seen in the example presented in Figure 6 (Santos et al., 2010). Although most of the existing transport network optimization models are focused on passenger transport, the optimization process is the same for freight transport.

Due to the considerable complexity of the transport networks and to the discrete nature of most models, there is no practical analytical solution for this problem, which leads to the adoption of heuristic techniques. Several techniques have been successfully used to address this kind of problems, predominantly metaheuristics such as tabu search, simulated annealing, ant colony optimization and genetic algorithms (Arnold et al., 2004; Crainic, 2000; Gallo et al., 2012; Santos et al., 2010; Yaghini et al., 2011; Yamada et al., 2009). Figure 7 shows an example of a comprehensive network optimization heuristic which starts with the generation of an initial solution and delivers a final optimized solution.

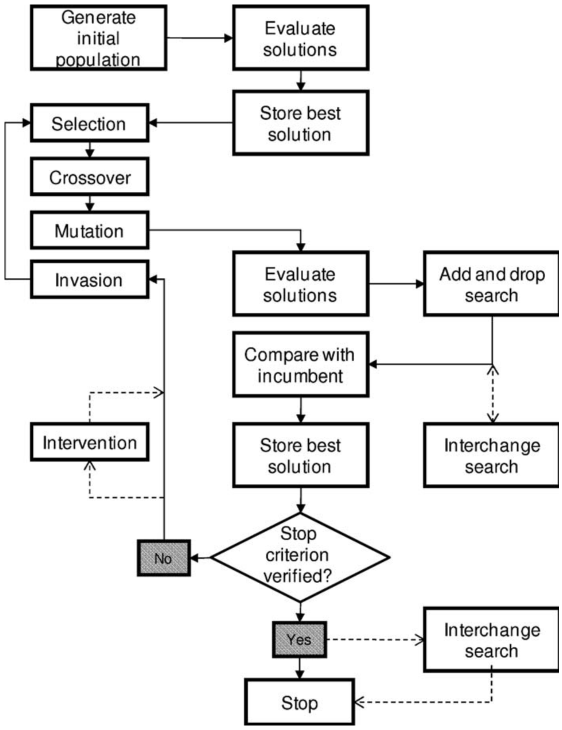


Figure 7 – Enhanced genetic algorithm for a network optimization problem (Santos et al., 2010)

Although the use of an heuristic technique does not guaranty that the obtained result is the best possible improvement solution, a good heuristic strives to find good solutions that are as close to the optimum as possible. Also, their relative swiftness makes them the only feasible alternative for complex transport networks with a vast range of improvement possibilities.

While there are some network optimization models made specifically for freight transport, they are limited in the type of improvement operations that can be performed. These limitations are due to various factors, namely by only allowing for either the creation of new links (Chen and Alfa, 1991) or the improvement of existing ones, having a limited search space with a small number of improvement possibilities (Yamada et al., 2009), or focusing on the optimization of specific infrastructures such as intermodal terminals (Arnold et al., 2004).

3.3. Description of the network optimization model

3.3.1. Freight Traffic Assignment Model

The presented network optimization model uses the traffic assignment model that was presented in chapter 2 to assign traffic to the networks under study. The strategic nature of the assignment model is perfectly suited for the type of strategic planning and policy decisions that this network optimization problem is designed to address. Although the assignment model considers only the main road and rail links, as well as the main intermodal terminals, it is quite flexible in the definition of the characteristics of the connector links. This feature can be used to simulate specific links, according to the needs of each network. This possibility is used in one of the networks on which the network optimization is tested, which includes a link representing a congested rail node in the middle of a rail network.

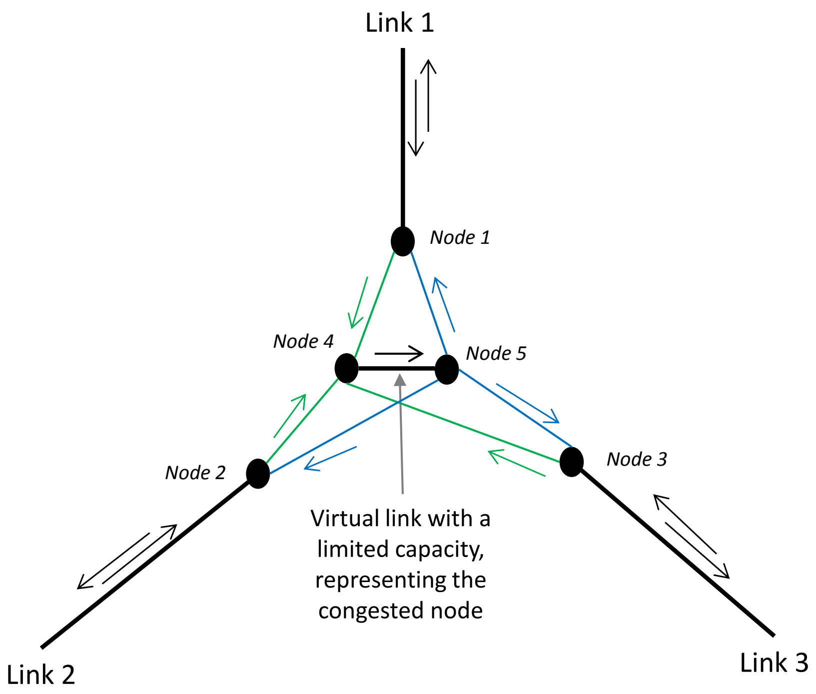


Figure 8 – Example of a congested rail node with 3 rail links converging on it

The introduction of this type of links may be justified on some networks, as most of the capacity problems in rail networks are due to specific points in the network, such as congested rail junctions As it can be seen in Figure 8, which displays the technique used to model congested rail nodes (Crainic et al., 1990), the model considers a virtual link that represents the total capacity of the rail node. The virtual link is basically a connector link which was adapted in order to simulate a congested rail node, becoming unidirectional and considering capacity limits. The capacity of the virtual link represents the maximum amount of trains that can pass through the congested rail node

3.3.2. Overview of the Optimization Process

The first step in the development of a network optimization model is the definition of the network structure, which will condition the type of network improvement operations that can be performed. The adopted solution for this model is a network structure where the links have a limited number of discrete quality levels, which each level corresponding to a different type of link, as exemplified in Figure 6. Link's quality levels can vary from zero, which corresponds to the mere possibility of building a link, to the highest level, corresponding to the best possible link quality. Each link level has an associated set of characteristics for each type of cargo, which may be freely defined by the user. This network structure allows for both the improvement and the construction of new links, permitting unlimited improvement possibilities. The model allows for the improvement of rail, intermodal terminal and virtual links representing rail nodes, meaning that all the links which are related to rail transport may be improved, as the model is designed to optimize rail networks. All the possible improvement operations have to have an associated cost, in order to quantify the investment necessary to perform each improvement scenario.

The assessment of the quality of each network improvement solution is based on two factors: the total generalized transport costs, and the total emissions of CO2. The total generalized transport costs reflect the economic costs that are supported by the freight carriers, and according to which they make their transport decisions. These costs are ultimately reflected in the price that is paid by the final users, reflecting the total amount of money that is spent on long distance freight transport. As such, any network improvement solution should try to minimize this parameter, in order to make freight transport less expensive. As for the total emissions of CO2, they quantify the total CO2 emitted by all the vehicles transporting freight, serving as a measure of the environmental impact caused by freight transport. Although freight transport contributes to the emission of various other pollutants, CO2 serves as an estimate for the total amount of pollutants, as it is highly correlated with the emission of those other pollutants. Although this is a parameter which does not have a direct impact on the costs of freight transport, it is an important externality which has a significant impact on the environment, and should thereby be taken into account when planning for transport investments. Also, there is a growing political will to internalize the external impacts produced by transport, via pollution charges or other instruments, which may lead to the effective incorporation of CO2 emissions in the total transport costs. Thereby, the network improvement solutions should try to minimize this parameter.

The assessment of the global quality of each improvement solution is calculated considering the minimization of the two considered parameters. Given the existence of more than one optimization parameter, it is necessary to define the weight that is given to each one of them. These weights are freely defined by the planner according to the priorities of the case under study, allowing for the necessary flexibility to deal with different situations and diverse politics.

The mathematical formulation of the network optimization problem can be described as follows:

Objective:

$$min\ N = \frac{w_{GC}}{w_{GC} + w_{EI}} * \frac{GC - GC_0}{GC_0} + \frac{w_{EI}}{w_{GC} + w_{EI}} * \frac{EI - EI_0}{EI_0} \qquad (3)$$

subject to:

$$\sum_{i \in I} m_i \leq B, \quad i \in I \qquad (4)$$

with:

$$GC = \sum_{i \in I} \sum_{k \in J} f_i^k * r_i^k * p_i^k, \quad \forall i \in I, \quad k \in J \qquad (5)$$

$$EI = \sum_{i \in I} \sum_{k \in J} f_i^k * c_i^k, \quad \forall i \in I, \quad k \in J \qquad (6)$$

$$m_i = \begin{cases} 0, \quad if \quad a_i \in \{0,1\}, \quad \forall i \in I & (7) \\ l_i * d_{a_i u_i z_i}, \quad if \quad a_i \in \{2\}, \quad \forall i \in I & (8) \\ d_{a_i u_i z_i}, \quad If \quad a_i \in \{3,4\}, \quad \forall i \in I & (9) \end{cases}$$

potential link status for each different link type:

$$u_i \in \{1,2,3,4,5\} \quad and \quad z_i \in \{1,2,3,4,5\}, \quad if \quad a_i \in \{0\}, \quad \forall i \in I \qquad (10)$$

$$u_i \in \{1\} \quad and \quad z_i \in \{1\}, \quad if \quad a_i \in \{1\}, \quad \forall i \in I \qquad (11)$$

$$u_i \in \{0,1,2,3,4,5\} \quad and \quad z_i \in \{0,1,2,3,4,5\}, \quad if \quad a_i \in \{2\}, \quad \forall i \in I \qquad (12)$$

$$u_i \in \{0,1,2,3,4\} \quad and \quad z_i \in \{0,1,2,3,4\}, \quad if \quad a_i \in \{3\}, \quad \forall i \in I \qquad (13)$$

$$u_i \in \{1,2,3,4,5,6\} \quad and \quad z_i \in \{1,2,3,4,5,6\}, \quad if \quad a_i \in \{4\}, \quad \forall i \in I \qquad (14)$$

possible link improvement operations for each different link type:

$$z_i = u_i, \quad if \quad a_i \in \{0,1\}, \quad \forall i \in I \qquad (15)$$

$$z_i \geq u_i, \quad if \quad a_i \in \{2\} \quad and \quad u_i \in \{0,1,2,4,5\}, \quad \forall i \in I \qquad (16)$$

$$z_i \in \{3,5\}, \quad if \quad a_i \in \{2\} \quad and \quad u_i \in \{3\}, \quad \forall i \in I \qquad (17)$$

$$z_i \geq u_i, \quad if \quad a_i \in \{3,4\}, \quad \forall i \in I \qquad (18)$$

with:

$$I = set \quad of \quad all \quad links$$

$$J = \{Gen, Int\}$$

where:

N - weighted change in total generalized cost and environmental impact

GC - total generalized cost

GC_0 – initial total generalized cost

EI - total environmental impact

EI_0 – initial total environmental impact

w_{GC} - weight given to generalized costs

w_{EI} - weight given to environmental impact

f_i^k – traffic flow (both ways) on link i, of cargo type k

r_i^k = vehicle capacity on link i, for cargo type k

p_i^k = generalized cost per unit of cargo on link i, for cargo type k

l_i = length of link i

c_i^k - CO2 emissions per vehicle per km on link i, for cargo type k

B - total available budget

m_i - investment on link i

a_i - link type of link i

u_i - original link status of link i

z_i - new link status of link i

d_{auz} - unit investment to upgrade link type a from original link status u to new link status z

This mathematical formulation reflects the network structure that is shown in Table 12, and allows for improvements on rail links, intermodal terminals and rail nodes.

Table 12 – Network structure

	LINK TYPE				
	Type 0 - Connector	**Type 1 - Road link**	**Type 2 - Rail link**	**Type 3 - Intermodal terminal**	**Type 4 - Rail node virtual Link**
POSSIBLE LINK LEVEL	-	-	**0**	**0**	-
	1	**1**	**1**	**1**	**1**
	2	-	**2**	**2**	**2**
	3	-	**3**	**3**	**3**
	4	-	**4**	**4**	**4**
	5	-	**5**	-	**5**
	-	-	-	-	**6**

This network optimization process is extremely difficult to be solved to exact optimality using an analytical process, which justifies the use of a heuristic. Due to its nature, a heuristic process may not guarantee the absolute best possible solution, but it will scrutinize the search space in order to find the most satisfactory solution possible, with the quality of the obtained solutions being dependent on the quality of the heuristic. The solution adopted for this model was to employ a local search heuristic. The decision to use this type of heuristic was made after studying various alternatives, and having consulted with specialists in the area. This is a relatively straightforward heuristic which delivers good results for this type of problem.

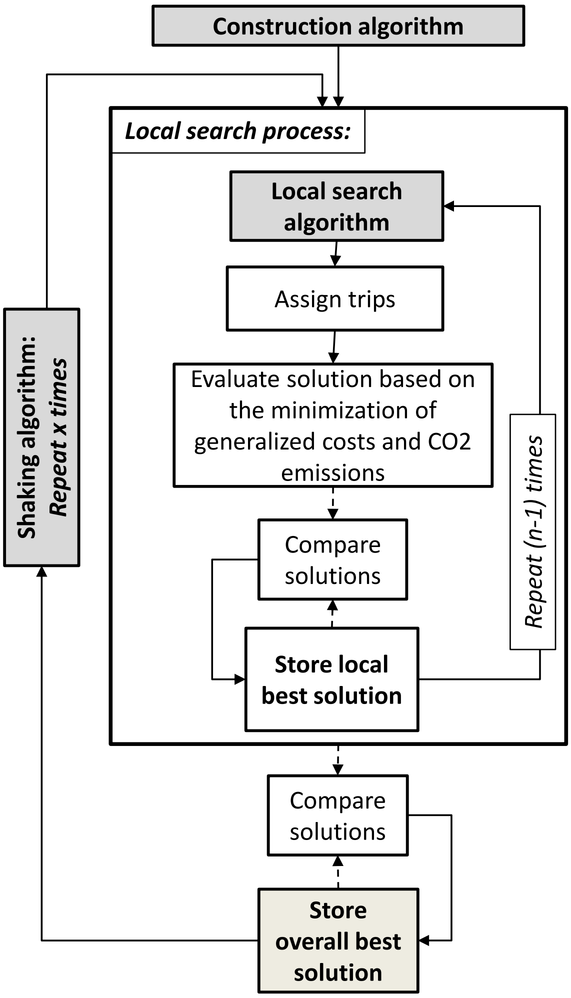


Figure 9 – Optimization process

As it can be seen in Figure 9, the network optimization process contains three main independent processes: an initial construction algorithm, a local search algorithm and a shaking algorithm. The construction algorithm is employed to create a feasible initial solution, through the use of a greedy algorithm. Based on that initial solution, the model runs two different cycles: an inner local search process, and an outer shaking process. The local search algorithm tries to optimize the solution by searching for better solutions on the search space vicinity of the initial solution. As for the shaking algorithm, it is used to make the solutions that come out of the local search process "jump" to a different point in the search space, in order to avoid being stuck in a local optimum.

3.3.3. Construction Algorithm

The construction process is based on a greedy algorithm which iteratively improves the links with the highest perceived improvement benefit to their best possible level, until there is no more available budget for improvements, as displayed in Figure 10.

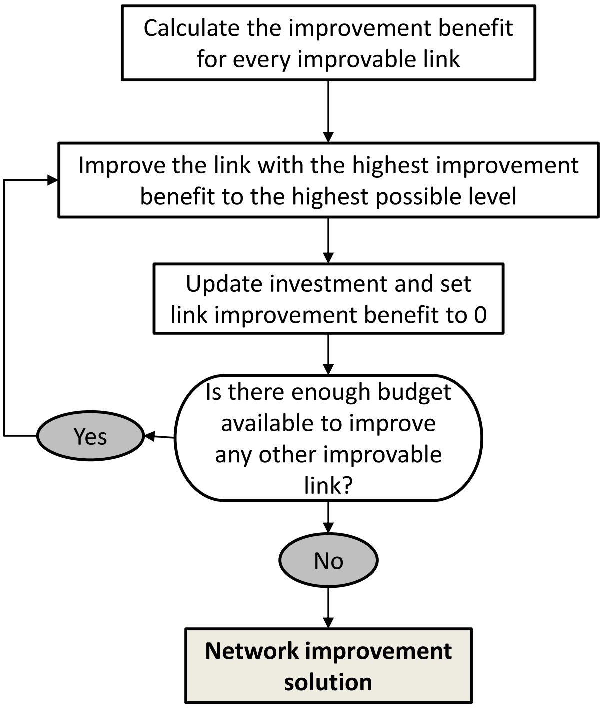


Figure 10 – Construction algorithm

The formula that is used to measure the perceived improvement benefit of each link, which was freely defined by us, is based on the amount of traffic that uses each link, as the links which are used by the biggest amount of traffic tend to be the most critical ones:

$$IB_i = \left(1 + \frac{h_i}{cap_i}\right)^2 * \sum_{k \in J} f_i^k \qquad \forall i \in I, \quad \forall k \in J \qquad (19)$$

where:
IB_i – improvement benefit of link i
h_i – maximum one way traffic flow on link i, of all cargo
cap_i – capacity (each way) of link i
f_i^k – traffic flow (both ways) on link i, of cargo type k

As it can be seen in Equation (19), the improvement benefit for each link is proportional to the volume of traffic that uses it and to its relative utilization. The reasoning behind the equation is that the links with higher volumes of traffic should be improved first, as their improvement benefits a higher share of the total traffic. But the relative utilization of each link is also important, as the most congested links will benefit the most with a capacity increase, allowing them to be used by more traffic. Therefore, the formula attributes a considerable bonus to the links with a high use relative to their capacity. Also, if an improvable link has no traffic passing throw it, which is the case in links of level 0, which represent the mere possibility of building a link, the model attributes it an improvement benefit marginally bigger than zero. This is done in order to guarantee that those links will be improved in case there are no other improvable links, as an improvement in those links may be beneficial.

The algorithm iteratively improves the links with the higher value of improvement benefit to their maximum level until there is no more budget available to make new improvements or until there are no more improvable links. This rational optimization process delivers an initial network optimization solution that is a good starting point for the local search algorithm.

3.3.4. Local Search Algorithm

The algorithm that is used for the local search process, which is the core of the whole optimization process, is schematized in Figure 11.

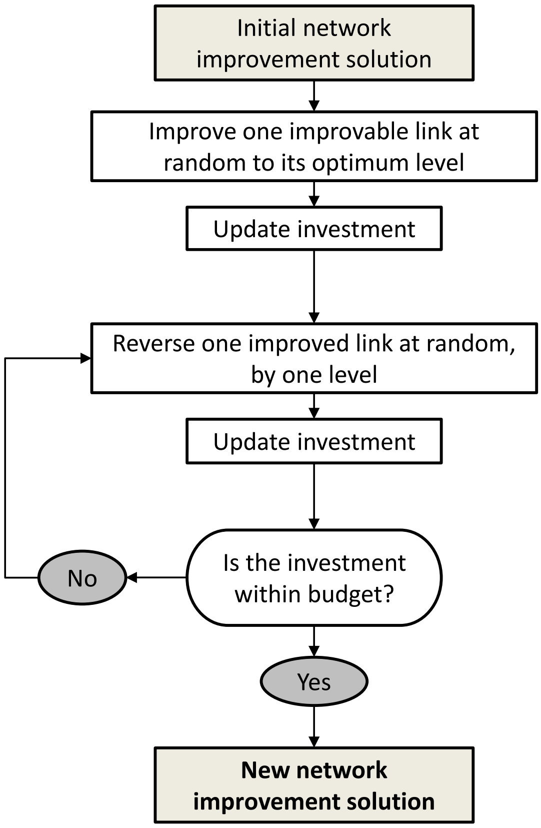


Figure 11 – Local search algorithm

The local search process takes an initial network improvement solution and makes a small change in it, creating a new solution in the search space vicinity of the initial solution. This is done by improving an improvable link at random to its optimum level, and then iteratively reversing a link at random by one level, until the investment is within budget. Each link's optimum level corresponds to the best possible level that a link can reach, without unnecessarily improve its capacity. This means that the optimum level for a link which does not need a capacity improvement is the best possible link level within its capacity bracket, and the optimum level for a link which needs a capacity improvement is the best possible link level in the next capacity bracket. The model considers that a link only needs a capacity improvement if the amount of traffic that is using it is greater than its capacity.

The goal of the local search algorithm is to search for better solutions in the search space vicinity of the incumbent solution. This is achieved by making a small change in the incumbent network optimization solution in order to test if that change has had a positive impact in the improvement solution, in which case it becomes the new incumbent solution. As it was shown in Figure 9, the new solution that comes out of the local search process is tested to see if it is better than the incumbent solution, in which case it becomes the new solution. This cycle is repeated by as many times as defined by the planner, delivering a local best solution.

3.3.5.Shaking Algorithm

The purpose of the shaking algorithm is to avoid the possibility of the model being stuck in a local optimum solution, as this is a problem that may arise from the application of a local search process. A local optimum is a situation in which the incumbent solution is the best solution in its search space vicinity but is not the best overall solution. In order to reach the best overall solution, it is necessary to leave the search space vicinity of the incumbent solution, but the local search process is incapable of searching for solutions that are not in that vicinity. Due to that, the purpose of the shaking algorithm is to make the solutions “jump” to a different place in the search space, leaving the vicinity of the incumbent solution.

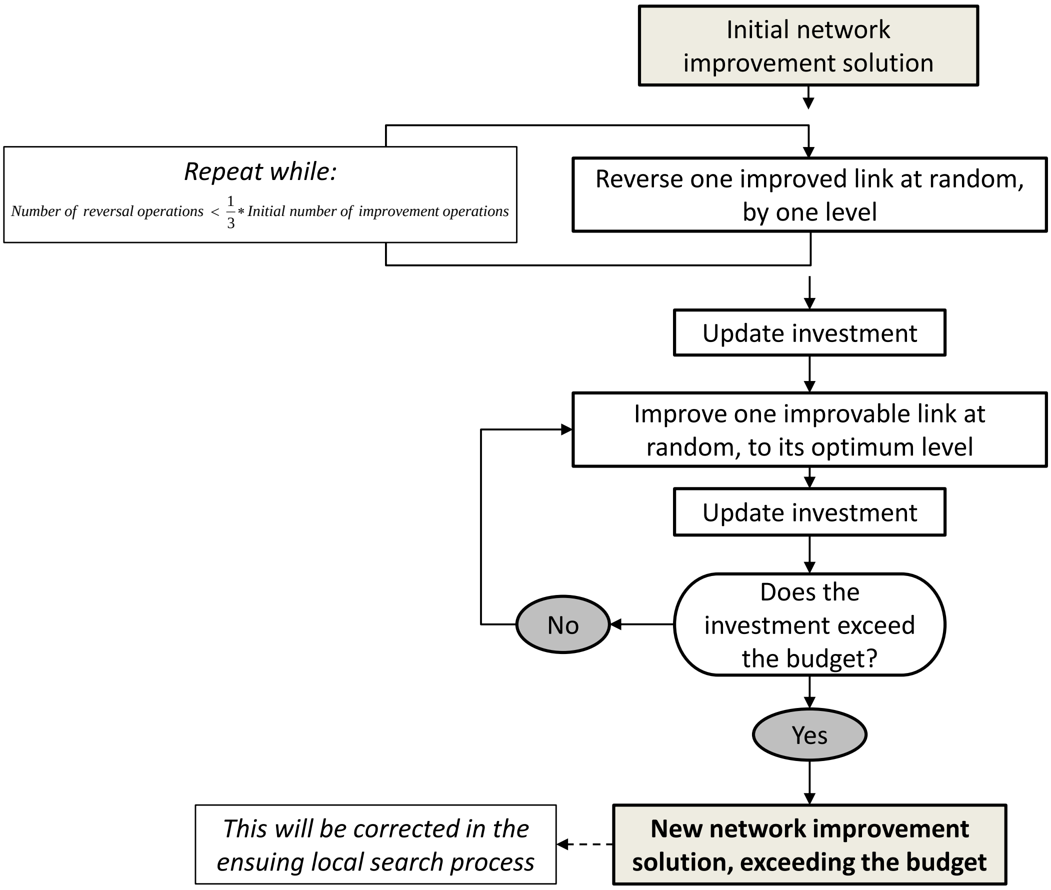


Figure 12 – Shaking algorithm

As it can be seen in Figure 12, the shaking algorithm consists in the reversal of a part of the improvement operations that were originally done, followed by the iterative improvement of random links to their optimum level, until the budget is reached or exceeded. An improvement/reversal operation is a one level improvement/reversal in the quality of a link, and the total number of improvement operations is the difference between the sum of the quality level of all the links and the sum of the original quality level of those same links. This process creates a new solution which is significantly different from the initial solution, while still retaining most of its features. The new solution will probably have left the vicinity of the search space of the initial solution, which is the goal of this shaking process. Due to the fact that the shaking algorithm only stops when the budget is reached or exceeded, most of the improvement solutions that will come out of this shaking algorithm will exceed the available budget, at least slightly. However, the ensuing local search process will correct this, by reversing as many links as needed until the total network investment is within budget.

The development of the shaking algorithm involved the study of various options on how to address this problem, which a special emphasis on two main aspects: what should be the starting point of the shaking algorithm and what should be the magnitude of change caused by this process, which is determined by the number of reversal operations.

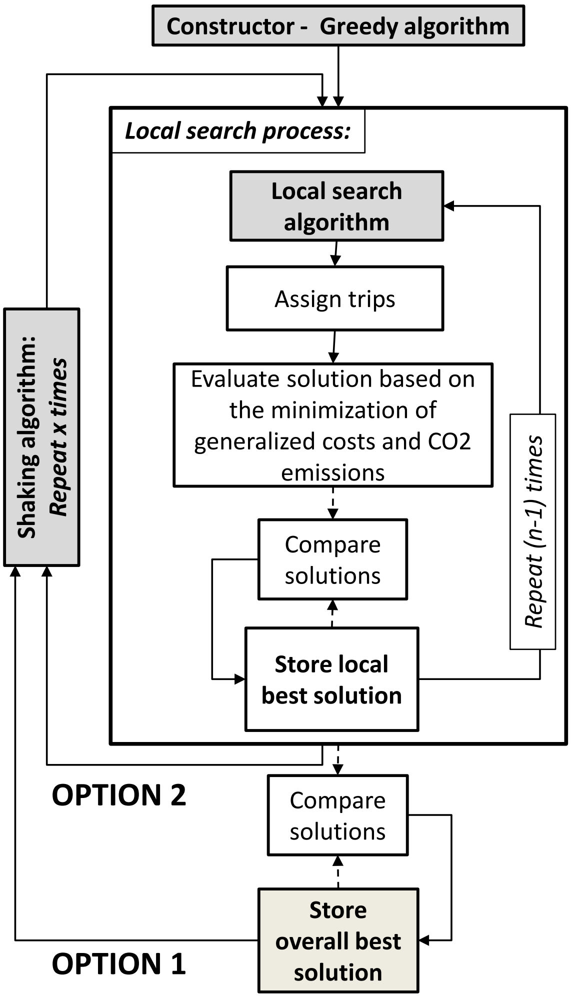


Figure 13 – Considered options for starting point of each shaking cycle

As it can be seen in Figure 13, two possible options were considered for the starting point of each shaking cycle: either to use the overall best solution obtained so far (Option 1) or to use the solution obtained in the previous local search process (Option 2). While the first option has the advantage of working with the best possible initial solutions, the second option allows for a greater diversity in the starting solutions, as each local search process creates a new solution. Two possible options were also considered for the magnitude of change caused by the shaking algorithm: either to reverse one third or two thirds of the initial improvement operations of each solution.

In light of the possible alternatives for the starting point of the shaking cycle and for the number of reversal operations, two different methods were tested: one that considers Option 2 for the starting point and reverses 2/3 of the improved links and another that considers Option 1 and reverses 1/3 if the improved links, which ended up being the adopted method. The former method should deliver considerably diverse solutions, as they are based on the solution created by the previous local search process and suffer an important change. The solutions delivered by the latter method should be more consistent, as it starts from the overall best solution and makes a less drastic change to it. The comparison of results between these two alternatives revealed that the latter method is clearly better, which lead to its adoption in the model.

Table 13 – Comparison of different methods for the shaking process

	Alternative method	**Adopted method**	
	Option 2 for initial solution of each shaking cycle and reversing 2/3 of the improved links	Option 1 for initial solution of each shaking cycle and reversing 1/3 of the improved links	
Percentage weighted change in total generalized cost and environmental impact	**Number of occurrences**		**% change**
0% < change < +2%	20	15	-25.0%
-2% < change < 0%	537	254	-52.7%
-4% < change < -2%	970	695	-28.4%
-6% < change < -4%	264	271	2.7%
-8% < change < -6%	509	819	60.9%
-10% < change < -8%	178	348	95.5%
-12% < change < -10%	22	98	345.5%
Total	**2500**	**2500**	

The superiority of the adopted method can be perceived in Table 13, which displays a comparison between the results obtained by the network optimization program using different methods for the shaking algorithm. These values are the result of the model’s application on a fictional network (network A, configured for scenario 2, from section *3.4*) comprising 50 shaking cycles with 50 local search cycles, totaling 2500 iterations. The various tests that were performed revealed that both the use of Option 1 for the starting point of the shaking cycle and the reversal of 1/3 of the improved links are superior to the alternatives, in individual terms, and the combination of both is clearly superior to the combination of the alternatives. This can be appreciated by the higher number of iterations with good results, which are those with a greater reduction in the weighted total generalized cost and environmental impact, and consequent lower number of iterations with less satisfactory results.

3.4. Application of the model on two fictional networks

3.4.1. Description of the networks and considered scenarios

The developed network optimization model was applied on two different fictional networks, in order to test and evaluate its performance and flexibility. These networks were created for this specific purpose, being designed to allow the model to display its potentialities. The two networks are relatively simple and share the same basic structure, featuring six generating poles (centroids) and a simple road network. While they have the same road structure, they feature considerable differences in the rail networks, in order to test the adaptability of the model to different network improvement possibilities. Network A presents a more complex rail network, which includes a congested rail node, as it can be seen in Figure 14.

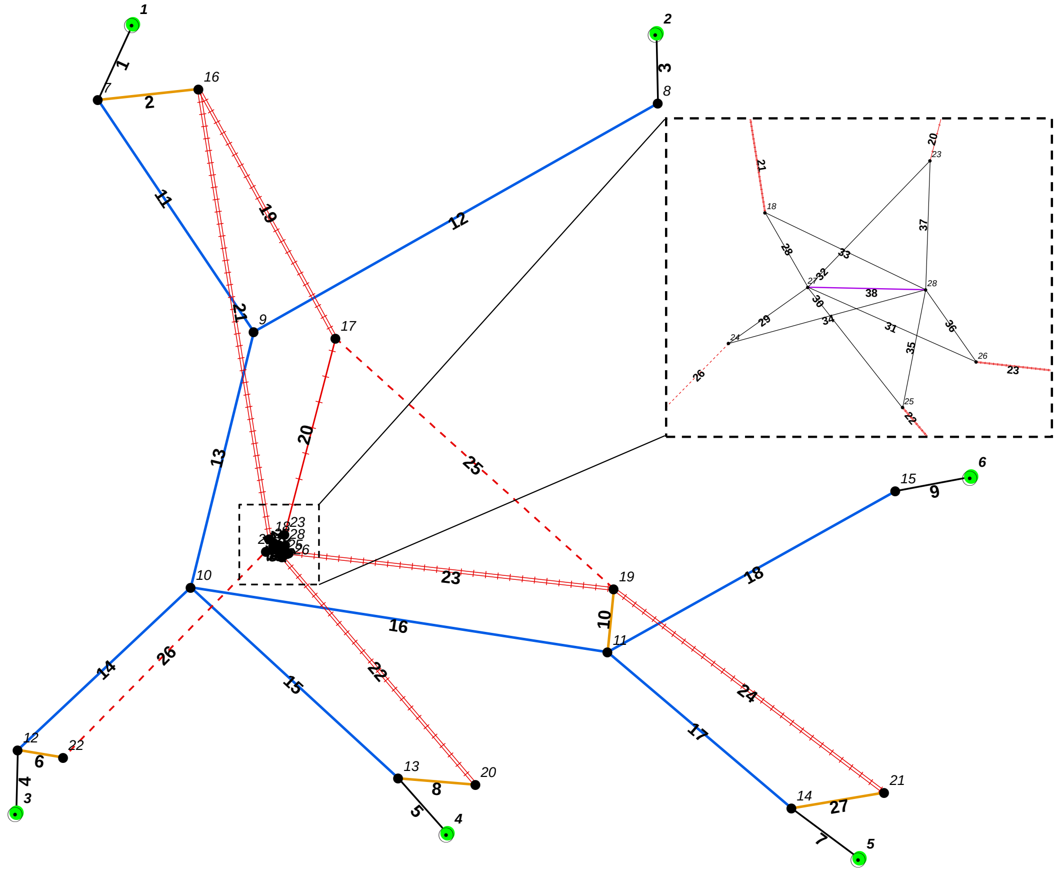


Figure 14 – Map of network A

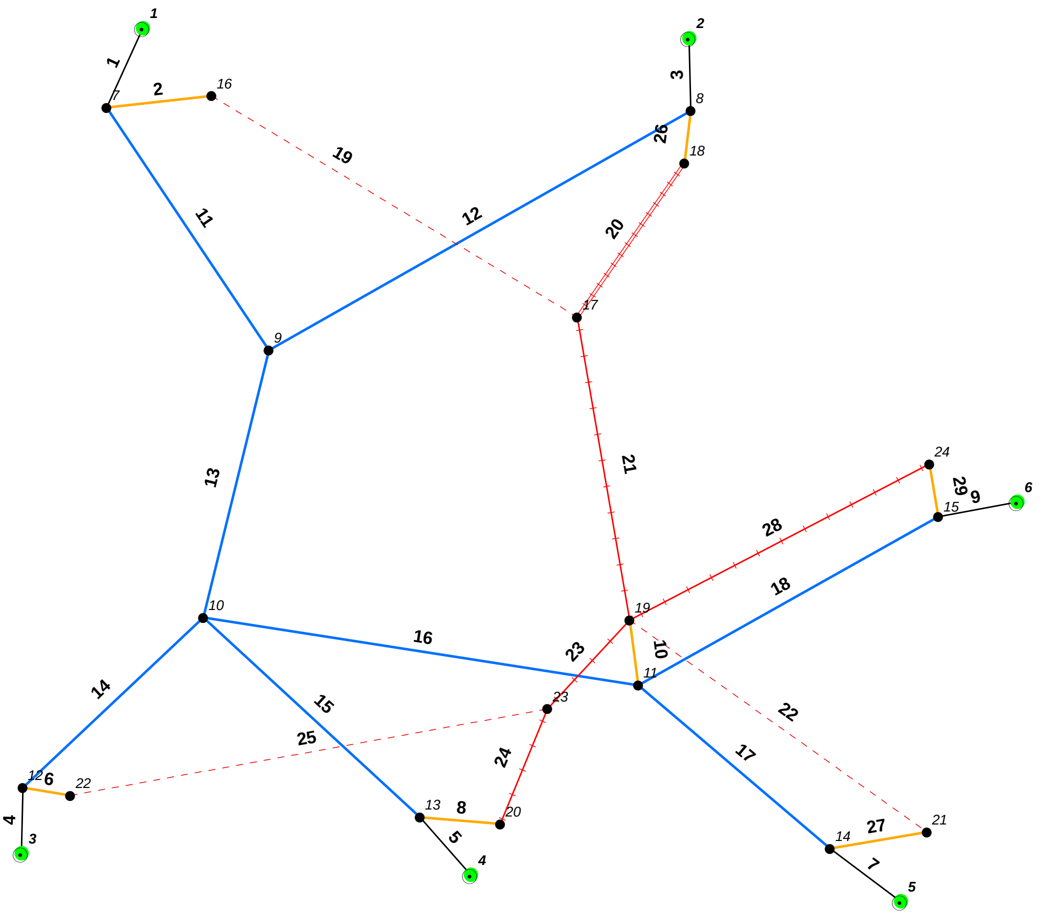


Figure 15 – Map of network B

The centroids are represented as large green dots and the nodes as small black dots. Connectors, road links, intermodal terminals and and virtual links are represented as black, blue, orange and purple lines, respectively. As for rail links, they are represented as various types of red lines: possible rail links are represented as dashed lines and single and double track lines are represented as single or double crossed lines, respectively. No centroids to rail connector links were used, being substituted by the introduction of intermodal terminals next to the centroids, making the connection between the centroid to road connectors and the rail network. These intermodal terminals can be used by general cargo, as they only allow for the transfer between rail and connector links, and not between rail and road links. This modeling approach is more realistic, as rail cargo usually resorts to some sort of local transport for the last mile of its journey.

In network A, there is a concentration of nodes in the convergence of links 20, 21, 22, 23 and 26 that represents a congested rail node, as previously exemplified in Figure 8, where the virtual link is link 38. As for network B, it features a simpler rail network, without congested nodes, but it offers the possibility of having all the centroids connected by rail. There are various possible link levels for each link type, as it can be seen in Table 14.

Table 14 – Summary of links characteristics

		LINK TYPE				
		Type 0 - Connector	**Type 1 - Road link**	**Type 2 - Rail link**	**Type 3 - Intermodal terminal**	**Type 4 - Rail node virtual Link**
LINK LEVEL	**0**	-	-	Possible link	Possible link	-
	1	Centroid to rail	Road link	Non-electrified single line; Max. train length = 450m	Intermodal terminal - capacity level 1	Rail node - capacity level 1
	2	Centroid to road	-	Electrified single line; Max. train length = 450m	Intermodal terminal - capacity level 2	Rail node - capacity level 2
	3	Port to rail	-	Electrified single line; Max. train length = 750m	Intermodal terminal - capacity level 3	Rail node - capacity level 3
	4	Port to road	-	Electrified double line; Max. train length = 450m	Intermodal terminal - capacity level 4	Rail node - capacity level 4
	5	Zero cost connector - for rail nodes	-	Electrified double line; Max. train length = 750m	-	Rail node - capacity level 5
	6	-	-	-	-	Rail node - capacity level 6

By consulting Figures 14 and 15, it is possible to see that some rail links are represented as dashed lines, which means they are level 0, while all the other rail links are either level 4 double lines or level 2 single lines. As for the intermodal terminals and the virtual link representing the congested rail node, they are all level 1. Since both networks share the same basic structure, it is possible to use the same O/D matrices, which facilitates the comparison between the results of the network optimization model.

Table 15 – General cargo O/D matrix

O/D	General cargo [ton]					
	1	2	3	4	5	6
1	0	0	40000	25000	22500	35000
2	0	0	30000	20000	27500	15000
3	40000	30000	0	0	0	0
4	25000	20000	0	0	0	0
5	22500	27500	0	0	0	0
6	35000	15000	0	0	0	0

Table 16 – Intermodal cargo O/D matrix

O/D	Intermodal cargo [ton]					
	1	2	3	4	5	6
1	0	0	0	0	55000	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	55000	0	0	0	0	0
6	0	0	0	0	0	0

Table 17 – Passenger trains O/D matrix

O/D	Passenger trains [vehicles]					
	1	2	3	4	5	6
1	0	0	0	7	5	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	7	0	0	0	0	0
5	5	0	0	0	0	0
6	0	0	0	0	0	0

The values that were considered for the network improvement costs, attributes of the different links levels and O/D matrices were chosen by us for this specific application, being realistic indicative values. For the sake of simplification, the link attributes for both intermodal cargo and generalized cargo were considered equal, with both types of cargo being measured in tonnes. Regarding the definition of CO2 emissions, they were quantified as grams per km in the road and rail links, and as grams per moved tonne of cargo in intermodal terminals. As for the links capacity, in the case of intermodal terminals it is measured in tonnes of moved cargo, while in rail links and congested rail nodes it is measured in number of trains.

3.4.2. Application results and discussion

The rail network optimization model was applied to both networks under the following conditions: regarding the assignment process, the traffic was introduced into the network in 20 interactions; as for the optimization process, each local search process consisted of 50 cycles, and the shaking process considered 50 shaking cycles. The optimization program was run in a dual core 2.5GHz processor and took approximately 17 minutes in the case of network A and 10 minutes in the case of network B. These are satisfactory running times for a strategic planning model, considering the size of the networks and the rather high number of improvement possibilities. The relative weights that were given to generalized costs and CO2 emissions minimization were 2 and 1, respectively, reflecting the critical relevance of transport costs but giving a significant weight to the environmental impacts. Two different scenarios were considered: scenario 1, with a total available budget of 250 million monetary units; and scenario 2, with a total budget of 500 million monetary units.

Table 18 – Results of the optimization process for network A

	NETWORK A													
	SCENARIO 1							SCENARIO 2						
Total investment (million monetary units)	247							484						
Percentage of reduction in total generalized cost	-0.1390%							-3.5632%						
Percentage of change in total CO2 emissions	-1.9433%							-23.8656%						
Percentage weighted change in total generalized cost and environmental impact	**-0.7404%**							**-10.3306%**						
	Id	Link type	Original link status	Improved link status	**Levels of improvment**	Initial traffic [vehicles]	Traffic after improvements [vehicles]	Id	Link type	Original link status	Improved link status	**Levels of improvment**	Initial traffic [vehicles]	Traffic after improvements [vehicles]
IMPROVEMENT SOLUTION	1	0	2	2	**0**	177500	177500	1	0	2	2	**0**	177500	177500
	2	3	1	1	**0**	13448	19297	2	3	1	3	**2**	13448	76104
	3	0	2	2	**0**	92500	92500	3	0	2	2	**0**	92500	92500
	4	0	2	2	**0**	70000	70000	4	0	2	2	**0**	70000	70000
	5	0	2	2	**0**	45000	45000	5	0	2	2	**0**	45000	45000
	6	3	1	1	**0**	0	0	6	3	1	1	**0**	0	0
	7	0	2	2	**0**	105000	105000	7	0	2	2	**0**	105000	105000
	8	3	1	1	**0**	6236	9047	8	3	1	1	**0**	6236	8152
	9	0	2	2	**0**	50000	50000	9	0	2	2	**0**	50000	50000
	10	3	1	1	**0**	0	0	10	3	1	1	**0**	0	0
	11	1	1	1	**0**	8203	7910	11	1	1	1	**0**	8203	5070
	12	1	1	1	**0**	4625	4625	12	1	1	1	**0**	4625	4625
	13	1	1	1	**0**	12828	12535	13	1	1	1	**0**	12828	9695
	14	1	1	1	**0**	3500	3500	14	1	1	1	**0**	3500	3500
	15	1	1	1	**0**	1938	1798	15	1	1	1	**0**	1938	1842
	16	1	1	1	**0**	7389	7238	16	1	1	1	**0**	7389	4352
	17	1	1	1	**0**	4889	4738	17	1	1	1	**0**	4889	1852
	18	1	1	1	**0**	2500	2500	18	1	1	1	**0**	2500	2500
	19	2	4	5	**1**	13.4	12.1	19	2	4	5	**1**	13.4	47.6
	20	2	2	5	**3**	13.4	12.1	20	2	2	2	**0**	13.4	8.2
	21	2	4	4	**0**	0	0	21	2	4	4	**0**	0	0
	22	2	4	4	**0**	6.2	9	22	2	4	4	**0**	6.2	8.2
	23	2	4	4	**0**	7.2	10.2	23	2	4	4	**0**	7.2	0
	24	2	4	4	**0**	7.2	10.2	24	2	4	5	**1**	7.2	42.5
	25	2	0	0	**0**	0	0	25	2	0	5	**5**	0	42.5
	26	2	0	0	**0**	0	0	26	2	0	0	**0**	0	0
	27	3	1	1	**0**	7212	10250	27	3	1	3	**2**	7212	67952
	28	0	5	5	**0**	0	0	28	0	5	5	**0**	0	0
	29	0	5	5	**0**	0	0	29	0	5	5	**0**	0	0
	30	0	5	5	**0**	6236	9047	30	0	5	5	**0**	6236	8152
	31	0	5	5	**0**	7212	10250	31	0	5	5	**0**	7212	0
	32	0	5	5	**0**	13448	19297	32	0	5	5	**0**	13448	8152
	33	0	5	5	**0**	0	0	33	0	5	5	**0**	0	0
	34	0	5	5	**0**	0	0	34	0	5	5	**0**	0	0
	35	0	5	5	**0**	6236	9047	35	0	5	5	**0**	6236	8152
	36	0	5	5	**0**	7212	10250	36	0	5	5	**0**	7212	0
	37	0	5	5	**0**	13448	19297	37	0	5	5	**0**	13448	8152
	38	4	1	2	**1**	26.9	38.6	38	4	1	1	**0**	26.9	16.3

Table 19 – Results of the optimization process for network B

	NETWORK B													
	SCENARIO 1							**SCENARIO 2**						
Total investment (million monetary units)	249							496						
Percentage of reduction in total generalized cost	-0.7633%							-0.3532%						
Percentage of change in total CO2 emissions	-5.9569%							-26.5079%						
Percentage weighted change in total generalized cost and environmental impact	**-2.4945%**							**-9.0714%**						
IMPROVEMENT SOLUTION	Id	Link type	Original link status	Improved link status	**Levels of improvment**	Initial traffic [vehicles]	Traffic after improvements [vehicles]	Id	Link type	Original link status	Improved link status	**Levels of improvment**	Initial traffic [vehicles]	Traffic after improvements [vehicles]
	1	0	2	2	**0**	177500	177500	1	0	2	2	**0**	177500	177500
	2	3	1	1	**0**	0	0	2	3	1	3	**2**	0	83622
	3	0	2	2	**0**	92500	92500	3	0	2	2	**0**	92500	92500
	4	0	2	2	**0**	70000	70000	4	0	2	2	**0**	70000	70000
	5	0	2	2	**0**	45000	45000	5	0	2	2	**0**	45000	45000
	6	3	1	1	**0**	0	0	6	3	1	1	**0**	0	0
	7	0	2	2	**0**	105000	105000	7	0	2	2	**0**	105000	105000
	8	3	1	1	**0**	6587	7220	8	3	1	1	**0**	6587	16622
	9	0	2	2	**0**	50000	50000	9	0	2	2	**0**	50000	50000
	10	3	1	1	**0**	0	0	10	3	1	3	**2**	0	55000
	11	1	1	1	**0**	8875	8875	11	1	1	1	**0**	8875	4694
	12	1	1	1	**0**	3917	3102	12	1	1	1	**0**	3917	3852
	13	1	1	1	**0**	12792	11977	13	1	1	1	**0**	12792	8546
	14	1	1	1	**0**	3500	3500	14	1	1	1	**0**	3500	3500
	15	1	1	1	**0**	1921	1889	15	1	1	1	**0**	1921	1419
	16	1	1	1	**0**	7372	6588	16	1	1	1	**0**	7372	3627
	17	1	1	1	**0**	5250	4503	17	1	1	1	**0**	5250	5250
	18	1	1	1	**0**	2122	2084	18	1	1	1	**0**	2122	1127
	19	2	0	0	**0**	0	0	19	2	0	5	**5**	0	52.3
	20	2	4	5	**1**	14.2	19	20	2	4	4	**0**	14.2	15.5
	21	2	2	3	**1**	14.2	19	21	2	2	5	**3**	14.2	61.9
	22	2	0	3	**3**	0	9.3	22	2	0	0	**0**	0	0
	23	2	2	2	**0**	6.6	7.2	23	2	2	2	**0**	6.6	16.6
	24	2	2	2	**0**	6.6	7.2	24	2	2	2	**0**	6.6	16.6
	25	2	0	0	**0**	0	0	25	2	0	0	**0**	0	0
	26	3	1	1	**0**	14152	30470	26	3	1	1	**0**	14152	15466
	27	3	1	1	**0**	0	14934	27	3	1	1	**0**	0	0
	28	2	2	3	**1**	7.6	5.2	28	2	2	2	**0**	7.6	27.5
	29	3	1	1	**0**	7565	8316	29	3	1	1	**0**	7565	27465

The improvement solutions can be consulted in Tables 18 and 19. The detailed results from the application of the network optimization model to both networks under the two different scenarios can be consulted in annexes A9 to A12.The solutions obtained for the two scenarios are considerably different, both in networks A and B, which reflects the adaptability of the model to different situations, as a bigger budget allows for more ambitious network interventions.

The solution adopted for scenario 1 in network A improves link 38, which represents the virtual link of the rail node, which is justified by the fact that in the original network configuration this node is congested, thereby it benefits from a capacity improvement. Also, rail links 19 and 20 are improved to their best possible level, in order to reduce the rail transport costs by operating trains with a higher cargo capacity. By contrast, on the solution obtained for scenario 2 the rail node link is not improved. This is justified by the fact that the higher available budget allowed for the construction of link 25, which is a new rail link that diverts rail traffic from the congested node, meaning that it no longer needs a capacity improvement. The construction of this new rail link, combined with the improvement of rail links 19 and 24, makes rail transport much more competitive for certain routes. This causes a sharp rise in rail traffic and in the amount of cargo that uses intermodal terminals 2 and 27, which therefore need to be improved in order to accommodate for this traffic growth.

As for the results obtained for network B, the solution adopted for scenario 1 includes the improvement of rail links 20, 21 and 28, and the construction of link 22. These improvements make the rail mode more competitive, by lowering its costs. Even so, the fact that link 21 did not need a capacity improvement indicates that the amount of rail freight traffic is still not very big. The higher available budget available in scenario 2 allows for the implementation of a bolder solution, which includes the construction of link 19 and the improvement of both the capacity and the transport costs of link 21. By connecting centroid 1 with the rail network, this solution generates a significant increase in rail freight traffic, which justifies the improvement of link 21, as well as the improvement of intermodal terminals 2 and 10.

The application of the network optimization model on these two networks under different investment scenarios delivered positive results which emphasize the model's adaptability to different conditions. This is patent in the very different outcomes that were obtained for the two scenarios in both networks, which are justified by the fact that the bigger budget of scenario 2 enabled the adoption of more ambitious solutions. The model used the higher available investment in scenario 2 to test and implement solutions that are radically different from those adopted in scenario 1. This was clear in the case of network A, where the improvement of the rail node was only justified in scenario 1, as the larger investment available for scenario 2 allowed for the construction of a new rail link that bypassed some of the traffic from the rail node, which eliminated the need to improve the rail node, as there were no longer congestion problems. This is an interesting result which highlights the fact that some ambitious transport infrastructure investments may make other smaller investments unnecessary, stressing the need to coordinate the short term and long term goals of transport infrastructure planning. Otherwise, there is a considerable risk of investing too much money in transport infrastructures that may soon become obsolete, or at least not as relevant as they used to be.

3.5. Conclusions

This chapter presents a rail network optimization model designed to assess the type of infrastructure investments needed to improve the transport of freight. The model is conceived for a strategic level of planning, modeling the major road and rail links, as well as specific congested rail nodes and intermodal terminals. It is an innovative model in the fact that it is not limited, allowing for both upgrades in the quality of existing rail and intermodal terminal links as well as the construction of new ones. This is attained by defining a set of possible link levels for each link type, including the mere possibility of building a link. The model can be applied to very different networks, not having a limit on the number or variety of network improvement possibilities, as it allows for the planners to freely define the characteristics of each link type and the potential improvement operations. The quality of the improvement solutions is assessed based on the reduction of the total generalized costs and CO2 emissions, in order to account for both the transport costs as well as the environmental impacts caused by freight transport. Regarding the optimization process, it is addressed using a heuristic based on a local search algorithm which delivers good results and can be run in a reasonable amount of time.

The practical application of the model on two different networks produced satisfactory results which display its ability to propose improvements in the quality of existing links, as well as the construction of new links. The contemplation of two different investment scenarios highlighted the model's adaptability to different situations, as a bigger budget allowed for more ambitious improvement solutions. The application also drawn attention to the fact that some ambitious transport infrastructure investments may make other smaller investments unnecessary, which is something that has to be taken account when making strategic transport planning decisions.

The developed strategic rail network optimization model is a valuable tool for planning investments in transport networks, which can be used both for medium and long term planning, by considering smaller or larger investment budgets. Even so, there is still room for future improvement, particularly in two main aspects. The first main improvement possibility would be the inclusion of a time analysis, which could ultimately create a network optimization calendar for a given time span that would be dependent on the available investment for each period. In order to do so, it would be important to account for possible evolutions in the demand for freight over the years, and possibly also on other parameters. This should also include a method to evaluate the benefit of each improvement operation over a certain period of time, in order to assess if it would be reasonable to make certain improvement operations that would eventually become obsolete. The final result of the full implementation of a time analysis would be a model which would deliver gradual improvement solutions, being dependent on the available investment for each period. The second improvement possibility would be the inclusion of passenger transport, which would be dependent on the improvement of the traffic assignment model to also included passenger traffic. This would create a strategic rail network optimization model for both passengers and freight transport, which would be an extremely valuable planning tool.

Chapter 4

VALIDATING THE NETWORK OPTIMIZATION MODEL: THE CASE OF THE IBERIAN PENINSULA

4. VALIDATING THE NETWORK OPTIMIZATION MODEL: THE CASE OF THE IBERIAN PENINSULA

4.1. Introduction

This chapter is dedicated to the application, on a real network, of the optimization model that was developed in the previous chapter. Although the developed model had already been applied to fictitious networks, it had not been applied to any real world network, thereby its results had not been studied in a real world environment, and its traffic assignment process was not empirically validated. The purpose of this chapter is to study the application of the network optimization on the transport network of the Iberian Peninsula in order to validate its traffic assignment process and to analyze the results from the network optimization.

The strategic rail network optimization model is applied to the transport networks of Portugal and Spain in order to obtain an optimal solution for investing a specific amount of money in the region’s rail network. The data inputs, namely those relative to the freight demand, are derived from 2008, which is the last year for which the required regional data was available. Two different scenarios are considered: a scenario in which the only goal is the minimization of the freight transport costs and another in which the goal of reducing the CO2 emissions is also taken into account. The application serves to calibrate and empirically validate the traffic assignment process, by comparing the estimated rail traffic with the actual traffic from 2008. Furthermore, the results of the optimization process reveal some critical findings regarding the planning of investments in rail infrastructure.

This chapter is structured in six sections. The following section is dedicated to the description of the transport network on which the model is run. The third section describes the method that was used to estimate the demand matrices and the fourth section is dedicated to the calibration and validation process. The fifth section presents the results from the application and the last section is dedicated to the final conclusions.

4.2. Description of the network

4.2.1. General characteristics

The model was applied to the transport network of the Iberian Peninsula by considering the road and rail networks of Portugal and Spain, as well as the main intermodal terminals. Due to the vastness of the considered area and to the strategic nature of the model, the territory was divided into large European NUTS 2 regions, each of which is represented by a centroid. They comprise five Portuguese regions: North, Center, Lisbon, Alentejo and Algarve; and fifteen Spanish regions, which correspond to the autonomous communities: Andalusia, Aragon, Principality of Asturias, Basque Country, Cantabria, Castile la Mancha, Castile and León, Catalonia, Community of Madrid, Extremadura, Galicia, La Rioja, Region of Murcia, Navarre and Valencian Community. The Iberian Peninsula is almost entirely surrounded by water, except for a relatively narrow land connection with the rest of Europe, thus the seaports play an important role in the transport of freight. Therefore, it is reasonable to consider the main seaports as independent centroids, as they are the source of a significant part of the long distance land traffic of freight. The adopted criterion for the selection of the main ports was the handled annual cargo, with the minimum threshold being set at 10 million tonnes of cargo per year. This definition comprises three Portuguese ports: Leixões, Lisbon and Sines; and eleven Spanish ports: A Coruña, Bahia de Algeciras, Barcelona, Bilbao, Cartagena, Castellón, Ferrol, Gigón, Huelva, Tarragona and Valencia.

The model was run using data from the year 2008, which is the last year for which the required data was available, thus it simulates the transport conditions for that year. The transport network represents the road and rail transport network of the Iberian Network, as well as the relevant intermodal terminals. It also considers the two most important land connections between the peninsula and the rest of Europe, which are the Spanish-French borders located near the Atlantic Ocean and the Mediterranean Sea, which are used by the vast majority of the land traffic between the Iberian Peninsula and Europe. These land connections are represented by two centroids which simulate the traffic flows between the peninsula and the rest of Europe. Due to the macro nature of the model, only the most important road axes were included with the same logic applied to the rail network. Even so, the majority of the existing rail lines were included, which is explained by the fact that the rail network in Portugal and Spain is not very dense and the majority of the lines correspond to the main transport axis. All relevant intermodal terminals were included, and the majority of the terminals are associated with regional centroids, as they are located near major cities. In addition to their connection with the road network, some major ports have a direct link to the rail network, which is represented in the model by the inclusion of a direct connector between the port and the adjacent rail node.

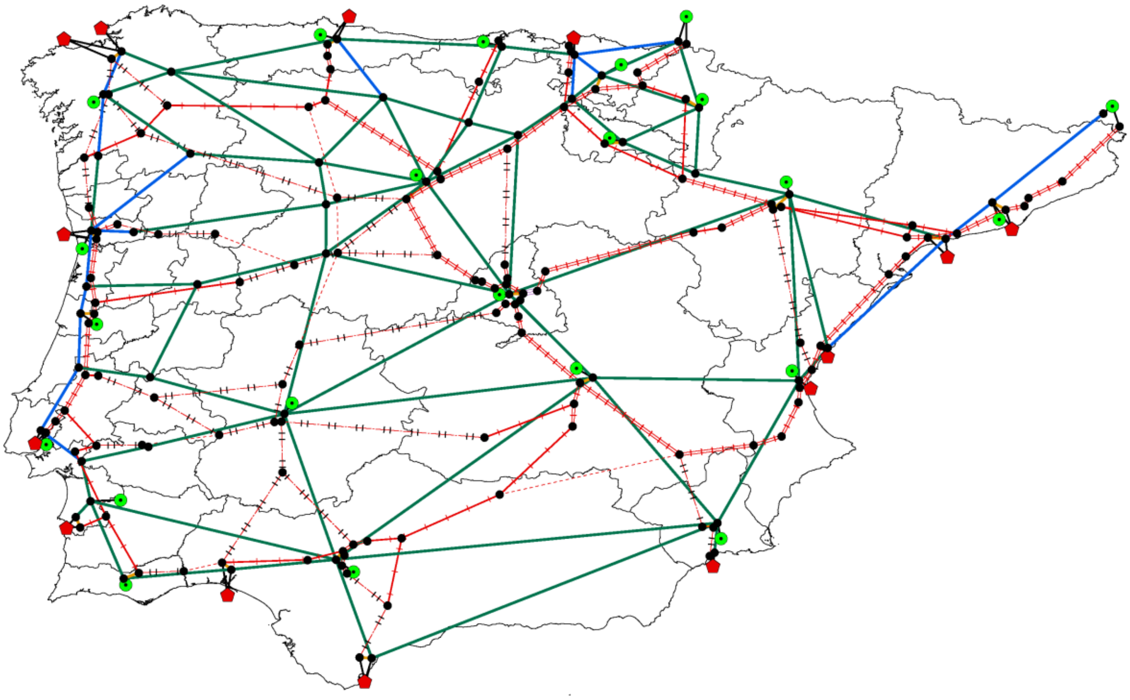


Figure 16 – Representation of the Iberian transport network

Figure 16 shows the adopted representation for the transport network: regional centroids are represented as large green dots and port centroids are represented as red pentagons. The road links are symbolized by plain blue and green lines, the rail links are depicted as crossed/dashed red lines and the connectors and intermodal terminals are represented by thin black and thick orange lines, respectively. The blue roads represent toll roads, whereas the green roads represent roads without tolls. The different representations of the rail lines distinguish between single-track and double track lines, as well as between electrified and non-electrified lines. In addition to the existing rail network, the model also considers potential future rail links, which are represented by dashed red lines.

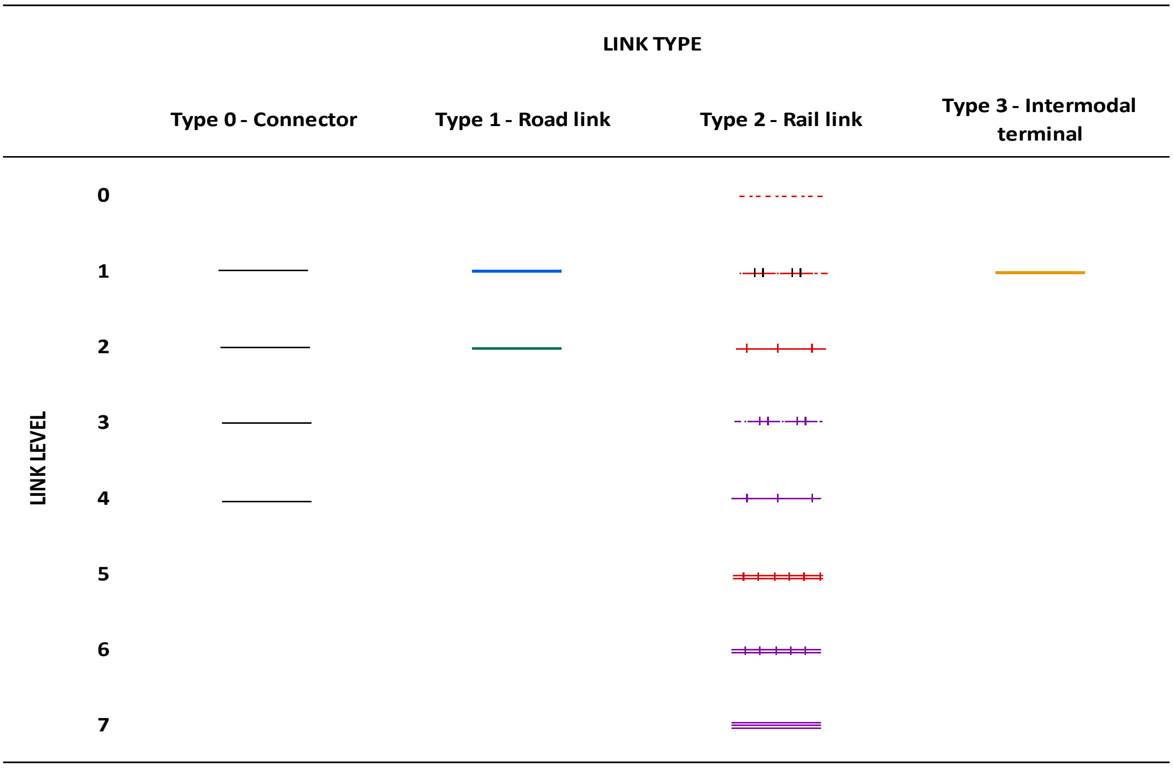
4.2.2.Adopted link characteristics

The model considers two types of road links, eight types of rail links, one type of intermodal terminal link and four types of connector links. This variety of links, which includes different types of existing and potential future links, does not represent a detailed list of all possible variations of road and rail links; instead, it represents a summary of the most important groups of links that is consistent with the macro nature of the model. The network does not contemplate any congested rail nodes, and as there is only one type of intermodal terminal link, there are no improvement possibilities for intermodal terminals.

Table 20 – Adopted network structure

		LINK TYPE			
		Type 0 - Connector	**Type 1 - Road link**	**Type 2 - Rail link**	**Type 3 - Intermodal terminal**
LINK LEVEL	**0**	-	-	Possible rail line	-
	1	Centroid to rail	Tolled main road	Non-electrified single line: *450m long trains*	Intermodal terminal
	2	Centroid to road	Non-tolled main road	Electrified single line: *450m long trains*	-
	3	Port to rail	-	Non-electrified single line: *750m long trains*	-
	4	Port to road	-	Electrified single line: *750m long trains*	-
	5	-	-	Electrified double line: *450m long trains*	-
	6	-	-	Electrified double line: *750m long trains*	-
	7	-	-	Electrified quadruple line: *750m long trains*	-

Table 21 – Link representation



LINK LEVEL	LINK TYPE Type 0 - Connector	 Type 1 - Road link	 Type 2 - Rail link	 Type 3 - Intermodal terminal
0			- - - - - - - - -	
1	———	———	·–++–·–++–·–	———
2	———	———	–+––+––+–	
3	———		–·–++–·–++––	
4	———		–+––+––+–	
5			═╪═╪═╪═╪═╪═╪	
6			═╪═╪═╪═╪═╪═	
7			≡≡≡≡	

As shown in Table 20, a distinction was made between major toll roads and major non-toll roads. This distinction is important because there are some road axes in the Iberian network that are served by toll motorways and do not have any suitable alternatives. The classification of rail links considers three main factors: the number of tracks, the existence or lack of electrification and the maximum allowable train length. Although other factors, such as the signaling system, could have been considered, a long list of link possibilities was not feasible, as that would add too much complexity to the optimization problem. The impact of different signaling systems, which may have a considerable impact on the capacity and speed of each line, was considered by assuming that electrified lines contain advanced signaling systems and non-electrified lines contain more rudimentary signaling systems. This assumption is reasonable for the Portuguese and Spanish networks because the busiest lines are electrified and contain modern signaling systems and there are virtually no non-electrified lines equipped with a modern signaling system.

The characteristics adopted for each type of link were based on information that was collected using various sources and on an expert opinion. With respect to the physical characteristics and performance of the links, the main sources of information included network statements from REFER (Rede Ferroviária Nacional - Portuguese rail network administrator) and ADIF (Administrador de Infraestructuras Ferroviarias - Spanish rail network administrator) (ADIF, 2009; REFER, 2009) and data relative to rail traffic in the Portuguese rail network in 2008, which were provided by REFER and other sources (Alvarez et al., 2010; Forkenbrock, 2001; Hanssen et al., 2012; Janic, 2007, 2008, 2009; Janic et al., 1999; Jeong et al., 2007; Tsamboulas and Moraitis, 2007; Vierth et al., 2009). Regarding the transport costs for using the various links, the main sources of information included the Spanish rail observatory (Alvarez et al., 2010) and several other important sources (Affuso et al., 2000; Bolis and Maggi, 2003; Forkenbrock, 2001; Hanssen et al., 2012; Janic, 2007, 2008, 2009; Kang et al., 2010; Racunica and Wynter, 2005; Tsamboulas and Moraitis, 2007; Vierth et al., 2009). The data relative to the CO2 emissions were collected using multiple sources (Alvarez et al., 2010; Cefic and ECTA, 2011; Forkenbrock, 2001; McKinnon, 2007); the same approach was used for the data relative to the costs of improving and building new rail links (Affuso et al., 2000; Baumgartner, 2001; BOE, 2003). The values of the link parameters and improvement costs from the various sources can be consulted in annex A13. The estimation of the initial link characteristics was based on the combination of the information collected from the various data sources with our sensibility and knowledge of the Iberian reality. Some of these initial characteristics were adjusted during the calibration process, resulting in the adopted link characteristics that are summarized in Table 22. The adopted rail link improvement costs are shown in Table 23.

Table 22 –Adopted link characteristics

Link type	**Link level**	**Description**	**Length [km]**	**Link capacity [veic/direction]**	**Speed (General cargo) [km/h]**	**Speed (Intermodal cargo) [km/h]**	**Vehicle capacity (General cargo) [ton]**	**Vehicle capacity (Intermodal cargo) [TEU]**	**Vehicle cost (General cargo) [€/km]**	**Vehicle cost (Intermodal cargo) [€/km]**	**Value of time (General cargo) [€/ton/h]**	**Value of time (Intermodal cargo) [€/TEU/h]**	**CO2 emissions (General cargo) [g/veic/km]**	**CO2 emissions (Intermodal cargo) [g/veic/km]**
Type 0 - Connector links	1	**Centroid to rail**	1	-	0.15	0.5	1	1	15.3	25.5	0.028	0.3304	150	1560
	2	**Centroid to road**	20	-	20	20	10	1	0.4	0.4	0.028	0.3304	1500	1560
	3	**Port to rail**	1	-	0.15	0.5	1	1	6.5	17.5	0.028	0.3304	150	1560
	4	**Port to road**	20	-	5	20	10	1	0.4	0.4	0.028	0.3304	1500	1560
Type 1 - Road links	1	**Tolled main road**	-	-	65	65	15	1.25	0.7	0.7	0.028	0.3304	1200	1219
	2	**Non-tolled main road**	-	-	65	65	15	1.25	0.5	0.5	0.028	0.3304	1200	1219
Type 2 - Rail links	0	**Possible rail line**	-	-	-	-	-	-	-	-	-	-	-	-
	1	**Non-electrified 450m single line**	-	25	45	45	600	47	20.8	16.2	0.028	0.3304	19200	17108
	2	**Electrified 450m single line**	-	35	55	55	600	47	17.8	14.1	0.028	0.3304	7200	6721
	3	**Non-Electrified 750m single line**	-	25	45	45	1000	79	25.9	20.3	0.028	0.3304	32000	28756
	4	**Electrified 750m single line**	-	35	55	55	1000	79	22.2	17.7	0.028	0.3304	12000	11297
	5	**Electrified 450m double line**	-	135	55	55	600	47	17.8	14.1	0.028	0.3304	7200	6721
	6	**Electrified 750m double line**	-	135	55	55	1000	79	22.2	17.7	0.028	0.3304	12000	11297
	7	**Electrified 750m quadruple line**	-	220	55	55	1000	79	22.2	17.7	0.028	0.3304	12000	11297
Type 3 - Intermodal terminal links	1	**Intermodal terminal**	1	-	0.15	0.5	1	1	6.5	17.5	0.028	0.3304	150	1560

Table 23 – Rail link improvement costs

Link improvement costs [million euro / km] From / To	**1 - Non-electrified 450m single line**	**2 - Electrified 450m single line**	**3 - Non-Electrified 750m single line**	**4 - Electrified 750m single line**	**5 - Electrified 450m double line**	**6 - Electrified 750m double line**	**7 - Electrified 750m quadruple line**
0 - Possible rail line	5.1	5.6	5.7	6.2	7.2	8	10.3
1 - Non-electrified 450m single line	-	2	2.3	3.5	5.1	6.4	9.9
2 - Electrified 450m single line	-	-	-	2.5	4.1	5.4	9.4
3 - Non-Electrified 750m single line	-	-	-	2	-	5.3	9.3
4 - Electrified 750m single line	-	-	-	-	-	4.3	8.9
5 - Electrified 450m double line	-	-	-	-	-	3.6	8.5
6 - Electrified 750m double line	-	-	-	-	-	-	6.7

4.2.3. Considered scenarios

In order to run the optimization model in the described network, it is necessary to define certain conditions, such as the universe of links that may be constructed or improved. This universe should be vast enough to consider different possible alternatives but should also be as limited as possible to avoid wasting computational time testing unfeasible solutions. To reach a desirable compromise, we have opted to consider only the major freight corridors, as defined by the European Commission (Comission, 2011) and other relevant links. The set of improvable links is depicted by the set of shaded rail links in Figure 17.

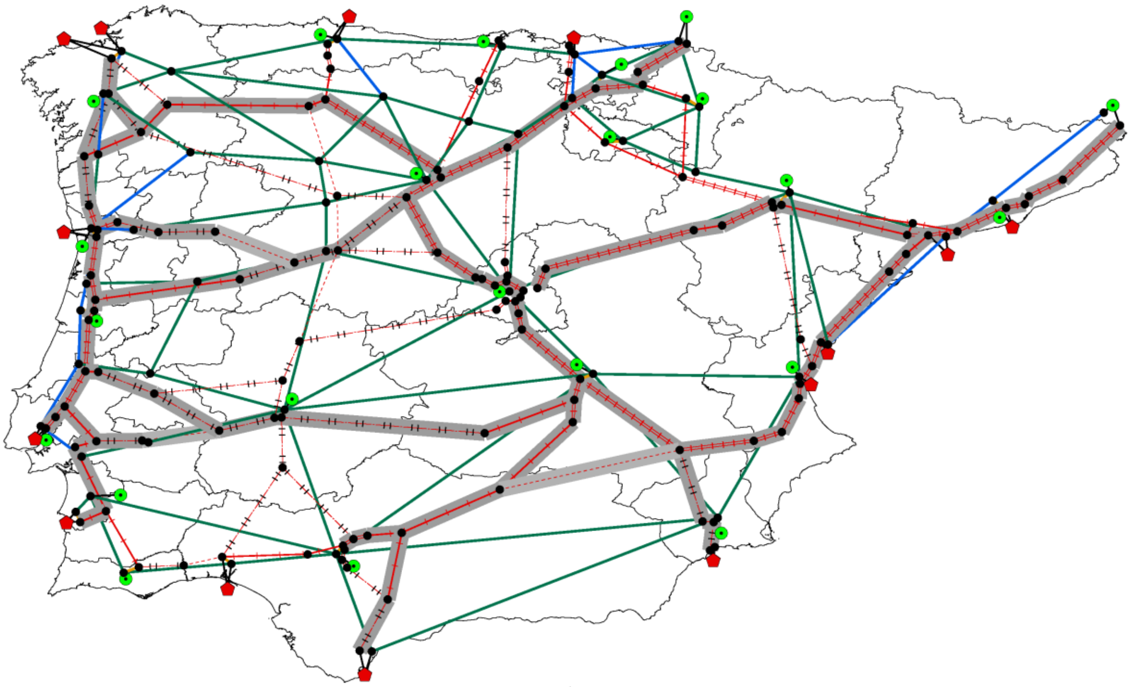


Figure 17 –Improvable links

The network investment budget that was adopted for this application was of 10 000 million euros, which is a reasonable value for a network of this size, as it enables the enactment of important network improvement operations while simultaneously limiting the amount of possible operations.

Two different scenarios were considered for the weights assigned to the minimization of costs and CO2 emissions: one scenario that only considers the minimization of total transport costs, by assigning a weight of zero to the emissions parameter; and another scenario in which the impact of the reduction of CO2 emissions was emphasized by attributing it a weight equal to that of the transport costs. These two different scenarios reflect very different policy priorities: a policy with complete disregard for environmental impacts, and a policy with significant consideration for environmental impacts. They are intended to provide an analysis on the impact and relevance of imparting more or less importance to the environmental impacts.

4.3. Estimation of the demand matrices

4.3.1. General description of the adopted method

The construction of the O/D matrices for the freight demand was based on various data sources, with the basic assumption that all intermodal cargo was generated at the major ports, representing the containerized cargo that is handled by those ports. Although some intermodal cargo is not generated at the major ports, such as containers transported by small ports or swap bodies and containers not used for sea-shipping operations, this is a reasonable assumption given the macro nature of the model. We opted to measure containerized cargo in TEUs because this is the most commonly used measurement unit for this type of cargo, and the model enables the use of different measurement units for different types of cargo. All remaining cargo was classified as general cargo and was quantified in tonnes.

The construction of the freight demand matrices involved building an O/D matrix for all cargo and then subtracting the intermodal cargo generated by the ports, obtaining the general cargo matrix. The data relative to the amount of cargo transported by land within and between Portugal and Spain, as well as the cargo transported from those countries to Europe was obtained from Eurostat statistics (Eurostat, 2009). The data regarding the generalized cargo and containerized cargo transported by the main ports was obtained by consulting the individual port statistics, which provided the tonnage of total cargo and the tonnage and TEUs of containerized cargo. The cargo transported by the ports was then subtracted to the total cargo transported in the countries within which they are inserted, as they function as independent centroids.

The freight movements between the different regions and main ports were estimated using the data from Eurostat statistics (Eurostat, 2009) and the total amount of cargo handled by the ports. The distribution of freight between the different O/D pairs was estimated using a gravitational model. The attraction of each region was based on its total GDP (Eurostat, 2011b) and the attraction of the main ports was measured by the total amount of cargo handled by them. As for the impedance between each O/D pair, it was assessed based on the physical distance between them. The final result of this process is the complete set of O/D matrices for both general cargo and intermodal cargo for the entire network.

The last step was the construction of the O/D matrix of the movement of passenger trains within the network, which is required by the model to assign the passenger trains to the network and calculate the remaining capacity that is left for freight trains. The movement of the passenger trains was estimated using data on rail traffic for the Portuguese rail network in 2008, which was provided by REFER, and the timetables for passenger train services in Spain in 2008, which was provided by the Spanish rail operator Renfe Operadora.

4.3.1. Construction of the freight O/D matrices

The construction of the freight demand matrices involved various steps, starting with the data relative to the movements of freight within and between the European countries (Eurostat, 2009) and with the cargo handled by the ports, and delivering the O/D matrices for both general cargo and intermodal cargo.

The first step consisted in the estimation of the aggregate freight movements between the three geographical areas under study: Spain, Portugal and Rest of Europe. This was transformed into eight relevant aggregate movements: Spain-Spain, Portugal-Portugal; Spain-Portugal; Portugal-Spain; Spain-Europe; Europe-Spain; Portugal-Europe; Europe-Portugal. The annual freight values for these movements were taken from Eurostat statistics (Eurostat, 2009), which had to be adjusted from the year 2006 to 2008, because the available data on country to country movements of freight was from 2006. This was made by assuming that the variation in freight movements from country to country between 2006 and 2008 was proportional to the variation in total freight movements in the origin country (Eurostat, 2011a) within that period.

The second step was the inclusion of two other main groups: Spanish main ports and Portuguese main ports. The values for the loaded and unloaded general cargo and intermodal cargo of each port were obtained by consulting the cargo statistics of each individual port. The total freight traffic from the ports of each country was subtracted to the total cargo transported in the respective countries, as they function as independent centroids. It was assumed that there are no land movements of freight between the seaports, given that the most logical way to move freight between two ports is by sea.

The third step consisted in the estimation of the movement of freight between the five main aggregate traffic generating poles: Spain, Portugal, Rest of Europe, Spanish main ports and Portuguese main ports. This was achieved by sub-dividing the eight aggregate movements estimated in the first step in order to include the movements of freight to and from the ports. This division was made according to the importance of each pole, which was measured by the amount of cargo moved by it. The outcome of this step was the construction of a matrix containing the aggregate freight movements between the five main traffic generating poles.

Table 24 – Aggregate freight movements

		DESTINATION				
		Spain	**Portugal**	**Rest of Europe**	**Spanish ports**	**Portuguese ports**
ORIGIN	**Spain**	Spain - Spain	Spain - Portugal	Spain - Rest of Europe	Spain - Spanish ports	Spain - Portuguese ports
	Portugal	Portugal - Spain	Portugal - Portugal	Portugal - Rest of Europe	Portugal - Spanish ports	Portugal - Portuguese ports
	Rest of Europe	Rest of Europe - Spain	Rest of Europe - Portugal	Rest of Europe - Rest of Europe	Rest of Europe - Spanish ports	Rest of Europe - Portuguese ports
	Spanish ports	Spanish ports - Spain	Spanish ports - Portugal	Spanish ports - Rest of Europe	Spanish ports - Spanish ports	Spanish ports - Portuguese ports
	Portuguese ports	Portuguese ports - Spain	Portuguese ports - Portugal	Portuguese ports - Rest of Europe	Portuguese ports - Spanish ports	Portuguese ports - Portuguese ports

Each of the twenty five aggregate freight movements that are displayed on Table 24 may now be dealt with as an individual O/D table. Thereby, all the traffic distribution techniques that are detailed in the subsequent steps were independently applied to each of these individual O/D tables, in order to estimate the detailed traffic distribution within each aggregate movement and build a global O/D table.

The fourth step involved the estimation of the amount of freight produced (output) and received (input) by each centroid in every individual O/D table. That estimation was based on a gravity weight factor that accounts for the importance of each centroid and the average distance between each centroid and the opposite traffic generating pole. The latter parameter was only applied in case there were significant differences between those distances, namely on the cross border aggregate freight movements: Spain – Portugal, Spain - Rest of Europe, Spain - Portuguese ports, Portugal – Spain, Portugal - Rest of Europe, Portugal - Spanish ports, Rest of Europe – Spain, Rest of Europe – Portugal, Rest of Europe - Spanish ports, Rest of Europe - Portuguese ports, Spanish ports – Portugal, Spanish ports - Rest of Europe, Portuguese ports – Spain and Portuguese ports - Rest of Europe.

Table 25 – Estimation of the freight input and output for each centroid

			DESTINATION				
			Input from ***centroid Y1***	**Input from** ***centroid Y2***	**…**	**Input from** ***centroid Yn-1***	**Input from** ***centroid Yn***
		Weight factor	W_{Y1}	W_{Y2}	…	W_{Yn-1}	W_{Yn}
ORIGIN	**Output from** ***centroid X1***	W_{X1}	TOTAL AGGREGATE FREIGHT FLOW (TAFF)				
	Output from ***centroid X2***	W_{X2}					
	…	…					
	Output from ***centroid Xm-1***	W_{Xm-1}					
	Output from ***centroid Xm***	W_{Xm}					

The calculation of the weight factor for each centroid is given by the following equation:

$$W_i = I_i * \left(\frac{1}{d_i}\right)^2 \qquad \forall i \in I \qquad (20)$$

with:

$$I = set \quad of \quad centroids \quad belonging \quad to \quad the \quad same \quad traffic \quad generating \quad pole$$

where:

W_i – weight factor associated with centroid i

I_i – importance attributed to centroid i

d_i – average distance between centroid i and the opposite traffic generating pole

The value of the importance I_i is equal to the total GDP of the region represented by the centroid (Eurostat, 2011b), in case of regional centroids, or to the total amount of cargo handled by the port, in the case of port centroids. In the cases where there were no significant differences in the average distance between each centroid and the opposite traffic generating pole, the value of the parameter d_i was assumed to be equal to 1. After having calculated the weight factors for all the centroids in each individual O/D table, it is possible to estimate the freight input or output of each centroid using the following equation:

$$I/O_i = TAFF * \frac{W_i}{\sum_{i \in I} W_i} \qquad \forall i \in I \qquad (21)$$

with:

$$I = set \quad of \quad centroids \quad belonging \quad to \quad the \quad same \quad traffic \quad generating \quad pole$$

where:

I/O_i – Input/output of centroid i

TAFF – Total aggregate freight flow

W_i – weight factor associated with centroid i

After having calculated the freight input and output for all the centroids in all the individual O/D tables, the fifth step in the traffic distribution process consisted in the estimation of all the individual movements inside each O/D table. This was achieved by using an iterative gravity traffic distribution process that uses the square of the distance between the centroids as the impedance factor.

Table 26 –Structure of the iterative traffic distribution process

			INPUT				
			Y_1	Y_2	…	Y_{n-1}	Y_n
		Balance factor	B_{Y1}	B_{Y2}	…	B_{Yn-1}	B_{Yn}
OUTPUT	X_1	B_{X1}	$F_{X1\ Y1}$	$F_{X1\ Y2}$	…	$F_{X1\ Yn-1}$	$F_{X1\ Yn}$
	X_2	B_{X2}	$F_{X2\ Y1}$	$F_{X2\ Y2}$	…	$F_{X2\ Yn-1}$	$F_{X2\ Yn}$
	…	…	…	…	…	…	…
	X_{m-1}	B_{Xn-1}	$F_{Xm-1\ Y1}$	$F_{Xm-1\ Y2}$	…	$F_{Xm-1\ Yn-1}$	$F_{Xm-1\ 2\ Yn}$
	X_m	B_{Xn}	$F_{Xm\ Y1}$	$F_{Xm\ Y2}$	…	$F_{Xm\ Yn-1}$	$F_{Xm\ 2\ Yn}$

The initial value adopted for the traffic flow between each pair of centroids was the inverse of the impedance, laying the foundations for the gravity process. The calculation of the balance factors for each row and column was calculated using the following equations:

$$Initial\ \ F_{kl} = \left(\frac{1}{d_{kl}}\right)^2 \qquad \forall k \in K, \quad l \in L \qquad (22)$$

$$Bx_k = \frac{X_k}{\sum_{l \in L} F_{kl}} \qquad \forall k \in K, \quad l \in L \qquad (23)$$

$$By_l = \frac{Y_l}{\sum_{k \in K} F_{kl}} \qquad \forall k \in K, \quad l \in L \qquad (24)$$

with:

$K = set \ \ of \ \ rows$

$L = set \ \ of \ \ columns$

where:

F_{kl} – flow of cargo between the centroids of row *k* and column *l*

d_{kl} – average distance between the centroids of row *k* and column *l*

Bx_k – balance factor of row *k*

By_l – balance factor of column *l*

X_k – freight output of the centroid of row *k*

Y_l – freight input of the centroid of column *l*

The iterative process alternatively balances the rows or the columns, by multiplying the flows of each row or column by the corresponding balance factor obtained in the previous iteration. This process gradually leads to a balanced outcome, where the sum of the flows of each line or column is equal to the output or input of the corresponding centroid. In the present case, the final traffic balance was reached after 200 iterations, with a maximum difference of around *0.10%* between the sum of the flow of the most imbalanced line and the output of the corresponding centroid. The final result of this process was the complete O/D matrix for all freight cargo.

The final step consisted in the separation between general cargo and intermodal cargo. This was done by separating the cargo that originated or terminated at ports into intermodal cargo and general cargo, according to each port's share of intermodal cargo as a percentage of the total handled cargo. As the ports are the only sources of intermodal cargo, the outcome of this process was the O/D matrix for intermodal cargo. The O/D matrix for general cargo was then obtained by subtracting the intermodal cargo to the total transported cargo. Finally, the O/D matrix for intermodal cargo was converted from tonnes to TEU's, as the statistics for intermodal cargo were displayed in both tonnes and TEU's. The obtained O/D matrices for general cargo and for intermodal cargo, as well as the matrix of the movement of passenger trains within the network can be consulted in annexes A14 to A16.

4.4. Calibration and validation

The assignment of traffic to the network was calibrated in order to define the value of the μ parameter in Equation (1) and to determine whether the model delivered realistic results. This calibration was based on the rail traffic data for the Portuguese network that was provided by REFER, specifically for the lines that cross the borders between the regions of each centroid, and on rail traffic data for the easternmost Spanish-French border at Portbou (Arroyo, 2009). The use of border crossings as calibration points was justified to prevent the mixing of long-distance traffic (centroid to centroid) with local traffic (within each centroid) because the latter is not considered in the model. The eight points that were utilized to calibrate freight rail traffic are illustrated in Figure 18. Note that only general cargo traffic is assigned using a logit distribution; therefore, the calibration of the μ

parameter only affects the distribution of this type of traffic. Intermodal cargo traffic could only be calibrated by modifying the adopted transport costs for that type of cargo.

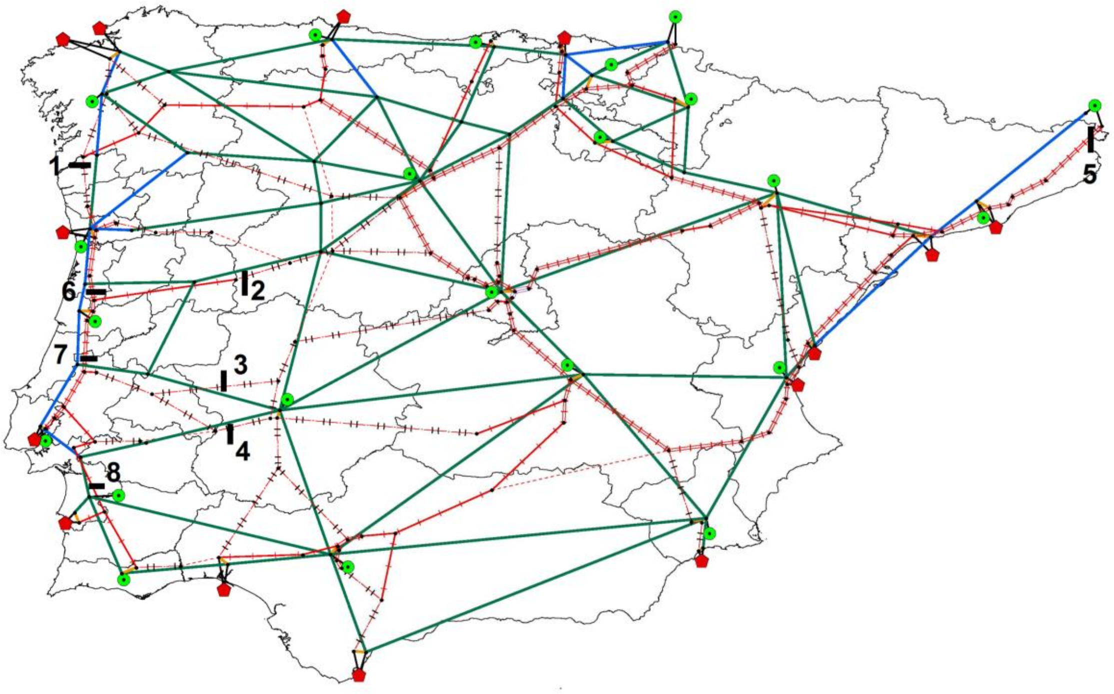


Figure 18 –Calibration points

After running the traffic assignment using the initial link characteristics and a value of *0.10* for the *μ* parameter, the results indicated that the model's intermodal rail traffic was significantly lower than actual traffic. This finding may have more than one justification, such as the fact that the model only considers intermodal traffic that originates in large ports, disregarding traffic from smaller ports and land containers, and the fact that there is no traffic distribution between road and rail transport, therefore rail transport is only used if it is cheaper in absolute terms than the road alternative. However, because the traffic was considered too low, we tested many possible cost variations and decided to apply a *30%* cost reduction in to the initial vehicle costs for intermodal cargo in connectors, rail links and intermodal terminals, which caused the model to deliver more realistic results, while remaining considerably lower than the actual traffic. With respect to the calibration of general cargo traffic, we adopted a value of *0.14* for the μ parameter after testing and assessing the results for many possible values. The rail on the Spanish-French border at Portbou was calibrated by adjusting the cost of the connectors that link the centroids to rail links, which are only used for the two rail connections between Spain and France, to include the cost of the break of gauge that exists at those borders (1668mm Iberian gauge on the Spanish side and 1435mm standard gauge on the French side).

Table 27 –Results after the calibration process

			AVERAGE RAIL FREIGHT TRAFFIC IN EACH DIRECTION								
			REALITY			MODEL					
	Calibration point		**Total traffic**	General cargo traffic	Intermodal cargo traffic	**Total traffic**	***Diference (%)***	General cargo traffic	*Diference (%)*	Intermodal cargo traffic	*Diference (%)*
INTERNATIONAL BORDERS	**1**	**Border Valença (PT-SP)**	**0.8**	0.6	0.2	**1.8**	***125%***	1.5	*162%*	0.2	*20%*
	2	**Border Vilar Formoso (PT-SP)**	**2.0**	1.9	0.1	**3.6**	***81%***	3.0	*61%*	0.6	*345%*
	3	**Border Ramal Cáceres (PT-SP)**	**0.0**	0.0	0.0	**1.0**	**-**	0.9	-	0.1	-
	4	**Border Elvas (PT-SP)**	**1.6**	1.3	0.3	**2.1**	***31%***	1.8	*45%*	0.3	*-20%*
	5	**Border Portbou (SP-FR)**	**10.0**	-	-	**9.1**	***-9%***	7.8	-	1.4	-
PORTUGUESE INTERNAL BORDERS	**6**	**Border Estarreja (PT NORTE - PT CENTRO)**	**11.7**	6.1	5.6	**8.1**	***-31%***	5.9	*-3%*	2.2	*-61%*
	7	**Border Caxarias (PT CENTRO - PT SUL)**	**11.0**	4.5	6.5	**11.2**	***2%***	7.3	*63%*	3.9	*-40%*
	8	**Border Ermidas Sado (PT LISBOA - PT ALENTEJO)**	**12.9**	9.9	3.0	**8.9**	***-30%***	6.3	*-36%*	2.7	*-12%*

The comparison between the results obtained by the model after the calibration process and the actual traffic is presented in Table 27. There are two relevant remarks about the real traffic values: the traffic in border point 3 is nearly zero, being rounded to zero, although a limited number of freight trains border regularly, indicating that there might have been works on the line or other problems in 2008; the real traffic considered for internal border point 8 was altered by us in order to remove the movements caused by a specific type of traffic, from the port of Sines to a coal power plant, which was significant but too specific to be contemplated by this model, due to its macro nature, and could negatively impact the results.

The main conclusion that can be derived from the results is that the traffic along the Portuguese internal borders is lower than real traffic, whereas the traffic along the Portuguese-Spanish border is higher than real traffic. In our opinion, the circumstance that the internal rail traffic in Portugal is lower than in reality is mainly due to the fact that the model does not contemplate empty trips, which are more prevalent in short trips such as the internal Portuguese movements. Although the traffic values obtained for the Portuguese-Spanish borders are high, their values are not excessive in absolute terms, due to the low levels of traffic. This difference is probably justified by the fact that the model does not consider that the physical borders have any impact on the movement of freight trains. This condition is not realistic, as there are technical and commercial issues that interfere with the free flow of freight trains between the two sides of the border, such as the existence of different railway companies and signaling systems in the two countries. The authors have nonetheless opted to maintain the model as it is, because the recent liberalization of the rail freight market and the efforts that are being made at the European level to facilitate the rail freight traffic between different states will tend to make this barrier effect disappear over time. Therefore, it would not make sense to include it in this long-term planning model. This logic is not applicable to the Spanish-French border, because in that case there is a break of gauge that represents a real physical barrier, which is not going to disappear in the foreseeable future. Thereby, it was necessary to consider this specific effect in the calibration process.

The favorable results delivered by the traffic assignment model constitute an empirical validation of its suitability for this type of strategic level transport networks.

4.5. Results

4.5.1. Network optimization results

The network optimization model was applied under the two previously mentioned scenarios: one that only considers the minimization of total generalized transport costs and assigns a weight of zero to the CO2 emissions parameter, and another scenario in which the impact of a reduction in emissions was assigned a weight equal to that of the minimization of costs. The model was run under the following conditions: the traffic was introduced into the network in 10 interactions, each local search process consisted of 50 cycles and the shaking process considered 50 shaking cycles. This optimization process took approximately 8 days to run in a core i7 2.0 GHz processor, with the two scenarios running simultaneously. The duration of the optimization process is mainly justified by the time spent running the traffic assignment process, as the optimization process is quite fast and straightforward. This is a reasonable computing time for a strategic planning tool of this nature, given the size of the network and the wide range of improvement possibilities. The detailed results from the application of the network optimization model to both scenarios can be consulted in annexes A17 and A18, and the summarized results are listed in Table 28.

Table 28 –Results for each scenario

	Scenario 1	**Scenario 2**
Weight given to the reduction of total generalized costs	1	1
Weight given to the reduction of total CO2 emissions	0	1
Total network investment (million euros)	9920	9934
Percentage of change in total generalized cost after the application of the improvments	0.04%	0.39%
Percentage of change in total CO2 emissions after the application of the improvments	-1.70%	-4.53%
Combined weighted percentage of change after the improvments	**0.04%**	**-2.07%**

The most significant conclusion that can be drawn from the results is that the model was not able to find a solution that reduced the total generalized costs of freight transport, indicating that the solution obtained for scenario 1 has no significance because it slightly increases, rather than decreases, the total generalized costs. This increase in total generalized costs is explained by the traffic assignment technique that is used for general cargo, which distributes the traffic by the least expensive road and rail alternatives. Thus, a reduction in rail transport costs will attract additional traffic to this mode, even if it is still more expensive than the road alternative. This means that the total generalized costs may increase when rail costs decrease, if the average cost of using rail transport remains higher than the average cost of road transport, as long as the increase in generalized costs due to the higher modal share of rail is greater than the reduction attributed to lower rail costs.

The results for scenario 2 indicate the reasonableness of considering generalized costs and CO2 emissions, because while it is not possible to reduce total generalized costs, there is a significant reduction in total CO2 production due to the lower emissions per tonne produced by rail transport. The variation in rail traffic after the network improvements, considering the network points that were employed in the calibration process, is illustrated in Table 29.

Table 29 – Rail freight traffic after network improvements for scenario 2

			TOTAL RAIL FREIGHT TRAFFIC IN EACH DIRECTION				
	Calibration point		**Initial**	**After the improvements**	***Diference in traffic (%)***	**Was the link improved?**	***Diference in transported cargo (%)***
INTERNATIONAL BORDERS	**1**	**Border Valença (PT-SP)**	**1.8**	**1.2**	***-32%***	**Y**	***13%***
	2	**Border Vilar Formoso (PT-SP)**	**3.6**	**4.9**	***35%***	**Y**	***125%***
	3	**Border Ramal Cáceres (PT-SP)**	**1.0**	**0.1**	***-92%***	**N**	***-92%***
	4	**Border Elvas (PT-SP)**	**2.1**	**2.4**	***16%***	**N**	***16%***
	5	**Border Portbou (SP-FR)**	**9.1**	**8.8**	***-3%***	**Y**	***61%***
PORTUGUESE INTERNAL BORDERS	**6**	**Border Estarreja (PT NORTE - PT CENTRO)**	**8.1**	**9.7**	***20%***	**N**	***20%***
	7	**Border Caxarias (PT CENTRO - PT SUL)**	**11.2**	**9.0**	***-20%***	**Y**	***34%***
	8	**Border Ermidas Sado (PT LISBOA - PT ALENTEJO)**	**8.9**	**10.2**	***14%***	**N**	***14%***

There is an overall increase in rail freight traffic, which is justified by the increased number of trains and/or by their increased capacity, with the improvements in the maximum allowable length of trains resulting in higher train capacities and reductions in the costs per tonne, making rail transport more attractive.

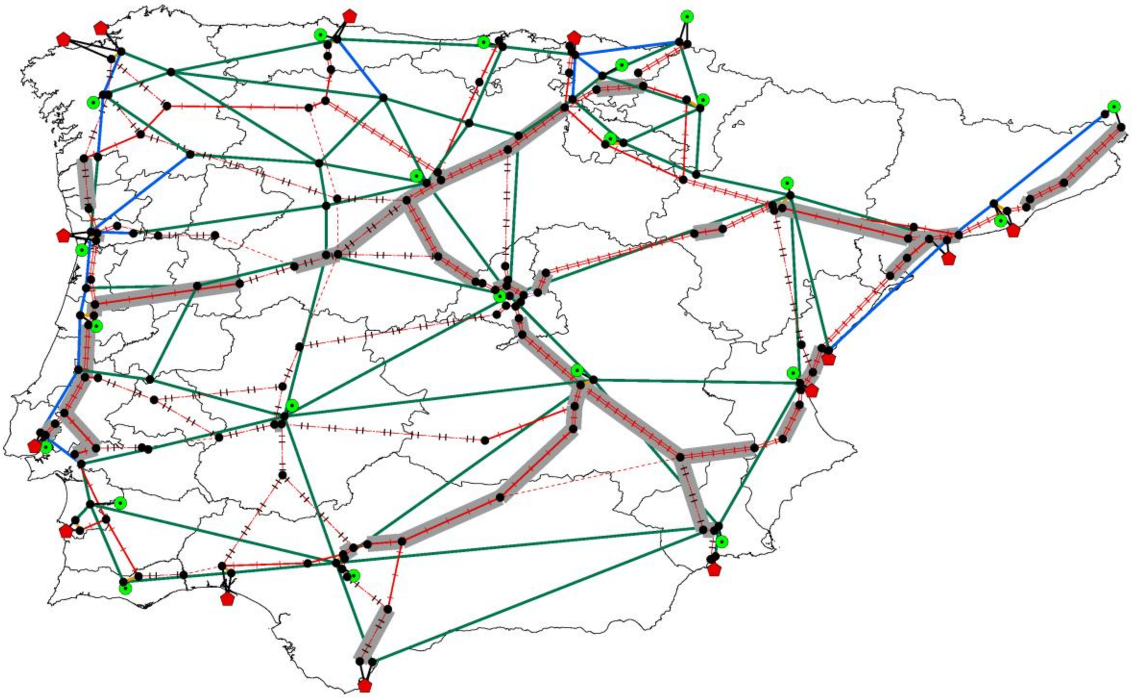


Figure 19 –Improved links in scenario 2

As illustrated in Figure 19, the links that were improved in scenario 2, such as the connection from Portugal and from Madrid to Europe and the Spanish Mediterranean line (along the east Mediterranean coast), are primarily located on the main transport axes. Nearly all improvements produce an increase in the maximum allowable train length, but not in the line capacity. This is consistent with the fact that the majority of Iberian long-distance lines are not congested, and with the efforts that are being made at the European level to increase the maximum length of freight trains, in order to make rail freight more competitive.

4.5.2.Discussion

The results revealed some relevant findings for planning rail infrastructure investments in the Iberian Peninsula, namely that the use of the minimization of the total generalized cost as a single parameter may not be feasible. This finding can be attributed to the fact that, although the improvement in rail links can increase the modal share of rail freight transport, that may not reduce the total generalized cost of transport. This occurs because rail transport is, on average, more expensive than road transport, even after significant rail improvements. Although the improvements may reduce the cost of long-distance movements, in which rail tends to be more competitive, this is not enough, by itself, to reduce the total generalized costs, at least in the Iberian network under study. This is due to the fact that the traffic assignment technique distributes the traffic of general cargo between the road and rail modes, and would not occur if all traffic was assigned to the least expensive route, in which case the improvements in the rail link would never generate an increase in the total cost. With the adopted assignment technique, the total generalized costs may increase when rail costs decrease if the average cost of rail transport remains higher than the average cost of road transport, as long as the increase in generalized costs due to the higher modal share of rail is greater than the reduction attributed to lower rail costs.

The distribution of traffic and its impact on the total generalized cost is dependent on the calibration parameters, particular on the μ parameter that is used in the logit distribution. The solution that was adopted for this application was to assume constant parameters and no significant changes in the traffic distribution, regardless of the improvements in rail transport. This was justified by the fact that the model was calibrated based on data from a single year, which makes it difficult to predict how these parameters would evolve under different conditions. However, it is possible that a significant improvement in the rail network would alter these traffic distribution parameters, which could lead to different results from those obtained considering constant parameters.

The limitations of a network optimization process that is based solely on the minimization of generalized costs justify the use of a complementary parameter, such as the environmental gain, which was the solution adopted for scenario 2. The network optimization results for this scenario, which consider the reduction of generalized costs and CO2 emissions, indicate that the improvements in the rail links are beneficial to the environment because they reduce the total CO2 emissions caused by freight transport. They also support the need to increase the maximum allowable length of freight trains to a minimum of 750 m to make rail freight more competitive and efficient, both economically and environmentally. Nearly all of the improvements involved increasing the maximum length of trains from 450m to 750m, alongside a few electrifications and only one capacity improvement, on a short single track section of the Mediterranean line near Tarragona. These results indicate that, apart from this specific link, the Iberian rail network does not require capacity improvements for long-distance links, although they may be required at specific points of conflict, such as junctions and rail yards, which are beyond the scope of this macro model. This is not surprising, as the majority of long-distance Iberian rail lines do not experience heavy passenger traffic, and some of the most important Spanish lines experience minimal traffic due to the recent construction of parallel high-speed lines. However, the optimization model considers fixed-demand conditions for freight train and passenger train trips, indicating that capacity issues may arise if there is a significant increase in freight demand or passenger train trips.

4.6. Conclusions

In this chapter, the previously developed rail network optimization model for freight transport is applied to the transport network of the Iberian Peninsula. Although the optimization model had already been tested on fictitious networks, it had not been applied to any real world network. Thereby, the purpose of this application was to validate the model's traffic assignment process and to study the results obtained from the application of the network optimization model to a real network.

The traffic assignment process was empirically validated based on the comparison between the actual traffic and the traffic that was estimated in the model. This involved a calibration process that served to adjust the necessary parameters of the model, which delivered favorable results that validate the traffic assignment process. As for the optimization process, it considered two different scenarios: one that only contemplates the minimization of the freight transport costs and another in which the goal of reducing the CO2 emissions is also taken into account. The results from the application of the network optimization model under two scenarios revealed some remarkable findings, particularly regarding the infeasibility of a reduction in the total generalized cost of freight transport via improvements to the rail network. This is due to the fact that on average, and even after significant network improvements, rail transport in the Iberian Peninsula is more expensive than road transport. This fact, in combination with the employed traffic assignment technique, which distributes general cargo between the least expensive road and rail alternatives, causes total generalized costs to increase when the rail network is improved.

The limitations of assessing the quality of a network optimization solution based solely on the minimization of the generalized transport costs justify the use of a complementary parameter, such as reduction of CO2 emissions. The results from the optimization of the second scenario indicate that the improvements to the rail network had a positive impact on the environment by reducing total CO2 emissions, even though generalized transport costs were not reduced. They also support the need to increase the maximum allowable length of freight trains to improve the competitiveness of freight rail transport, whereas the capacity of most long-distance links of the Iberian rail network seems to be sufficient.

Chapter 5

ASSESSING THE IMPACT OF RAIL NETWORK IMPROVEMENTS ON FREIGHT TRANSPORT UNDER DIFFERENT SCENARIOS

5. ASSESSING THE IMPACT OF RAIL NETWORK IMPROVEMENTS ON FREIGHT TRANSPORT UNDER DIFFERENT SCENARIOS

5.1. Introduction

This chapter is dedicated to the application of the rail network optimization model to a real network under different scenarios, in order to study the differences between the resulting outcomes. The goal of this application is to simulate possible evolutions in key variables, in order to assess how these different circumstances will affect the distribution of traffic and the impact of the possible network investments.

The network optimization model is applied under different conditions to the transport network of the Iberian Peninsula, which was modeled in the previous chapter. This is accomplished by considering twelve different scenarios that simulate possible future evolutions for the demand for freight and for the price of oil, in order to study the influence of these changes on the distribution of traffic and on the impact of the network improvements. A robustness analysis is then performed to identify solutions that can adequately address all the considered scenarios. Some final conclusions are drawn regarding the influence that the hypothetical future scenarios will have on freight traffic and on the impact of planned network investments.

This chapter is structured in six sections. The second section is dedicated to the description of the network and base conditions under which the model was run and the third section describes the various considered scenarios. The fourth section presents the results from the application and includes a robustness analysis of the obtained solutions. The last section presents the final conclusions.

5.2. Description of the network and baseline conditions

5.2.1. General description of the network

The application of the network optimization model under different scenarios was performed on the transport network of the Iberian Peninsula. This network was detailed in chapter 4, containing the road and rail networks of Portugal and Spain, as well as the main intermodal terminals. Due to the large area and the macro nature of the model, the territory was divided into large European NUTS 2 regions, each of which represented by a centroid. The main ports of the peninsula, which handle a minimum of 10 million tonnes of cargo per year, were also included as independent centroids. The model uses data from the year 2008; therefore, the transport network was modeled for the Iberian road and rail networks in 2008 and the relevant intermodal terminals. Due to the macro nature of the model, only the most important transport axes were included, as well as all relevant intermodal terminals.

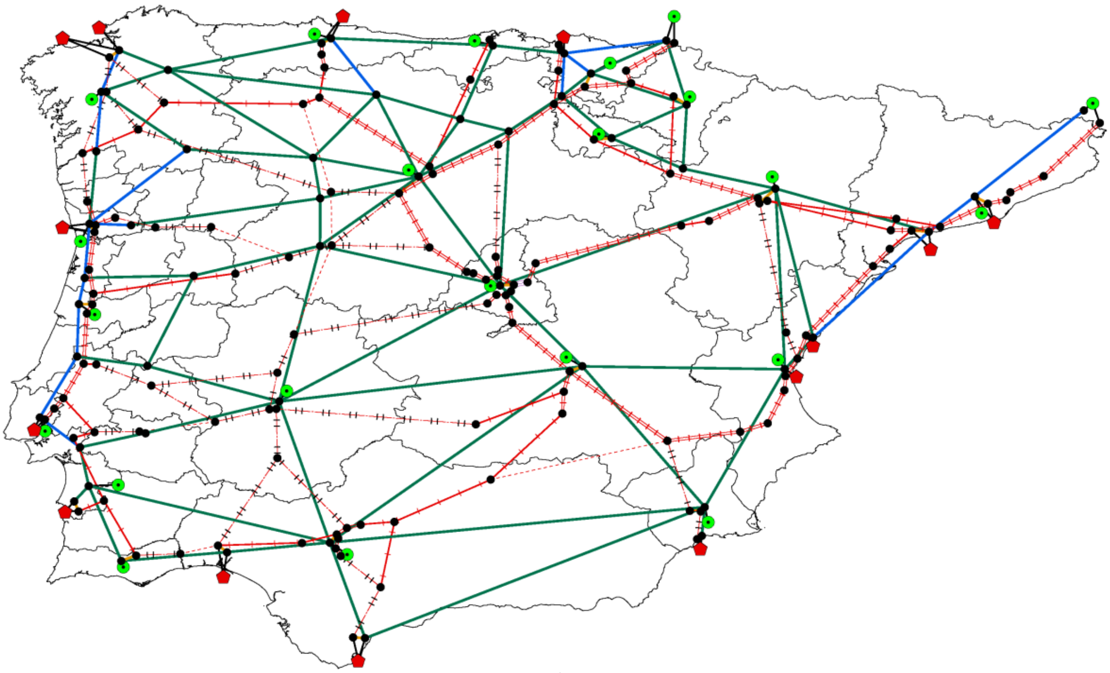


Figure 20 –Representation of the Iberian transport network.

Figure 20 shows the adopted representation for the Iberian transport network: regional centroids are represented as large green dots and port centroids are represented as red pentagons. The road links are symbolized by plain blue and green lines, the rail links are depicted as crossed/dashed red lines and the connectors and intermodal terminals are represented by thin black and thick orange lines, respectively. The blue roads represent toll roads, whereas the green roads represent roads without tolls. The different representations of the rail lines distinguish between single-track and double track lines, as well as between electrified and non-electrified lines. In addition to the existing rail network, the model also considers potential future rail links, which are represented by dashed red lines.

5.2.2. Adopted link characteristics

The model considers two types of road links, eight types of rail links, one type of intermodal terminal link and four types of connector links, containing the most important types of links. The classification of rail links considers three main factors: the number of tracks, the existence or lack of electrification and the maximum allowable train length. Although other factors, such as the signaling system, could have been considered, a long list of link possibilities was not feasible, as that would add too much complexity to the optimization problem. The impact of different signaling systems, which have a considerable impact on the capacity and speed of each line, was considered by assuming that electrified lines contain advanced signaling systems and non-electrified lines contain more rudimentary signaling systems

Table 30 – Network structure

		LINK TYPE			
		Type 0 - Connector	**Type 1 - Road link**	**Type 2 - Rail link**	**Type 3 - Intermodal terminal**
LINK LEVEL	**0**	-	-	Possible rail line	-
	1	Centroid to rail	Tolled main road	Non-electrified single line: *450m long trains*	Intermodal terminal
	2	Centroid to road	Non-tolled main road	Electrified single line: *450m long trains*	-
	3	Port to rail	-	Non-electrified single line: *750m long trains*	-
	4	Port to road	-	Electrified single line: *750m long trains*	-
	5	-	-	Electrified double line: *450m long trains*	-
	6	-	-	Electrified double line: *750m long trains*	-
	7	-	-	Electrified quadruple line: *750m long trains*	-

The characteristics that were adopted for each type of link were based on information that was collected using various sources, as detailed in chapter 4. The baseline link characteristics, obtained after the calibration process, are summarized in Table 31, and the rail link improvement costs are shown in Table 32.

Table 31 – Baseline link characteristics

Link type	Link level	Description	Length [km]	Link capacity [veic/direction]	Speed (General cargo) [km/h]	Speed (Intermodal cargo) [km/h]	Vehicle capacity (General cargo) [ton]	Vehicle capacity (Intermodal cargo) [TEU]	Vehicle cost (General cargo) [€/km]	Vehicle cost (Intermodal cargo) [€/km]	Value of time (General cargo) [€/ton/h]	Value of time (Intermodal cargo) [€/TEU/h]	CO2 emissions (General cargo) [g/veic/km]	CO2 emissions (Intermodal cargo) [g/veic/km]
Type 0 - Connector links	1	**Centroid to rail**	1	-	0.15	0.5	1	1	15.3	25.5	0.028	0.3304	150	1560
	2	**Centroid to road**	20	-	20	20	10	1	0.4	0.4	0.028	0.3304	1500	1560
	3	**Port to rail**	1	-	0.15	0.5	1	1	6.5	17.5	0.028	0.3304	150	1560
	4	**Port to road**	20	-	5	20	10	1	0.4	0.4	0.028	0.3304	1500	1560
Type 1 - Road links	1	**Tolled main road**	-	-	65	65	15	1.25	0.7	0.7	0.028	0.3304	1200	1219
	2	**Non-tolled main road**	-	-	65	65	15	1.25	0.5	0.5	0.028	0.3304	1200	1219
Type 2 - Rail links	0	**Possible rail line**	-	-	-	-	-	-	-	-	-	-	-	-
	1	**Non-electrified 450m single line**	-	25	45	45	600	47	20.8	16.2	0.028	0.3304	19200	17108
	2	**Electrified 450m single line**	-	35	55	55	600	47	17.8	14.1	0.028	0.3304	7200	6721
	3	**Non-Electrified 750m single line**	-	25	45	45	1000	79	25.9	20.3	0.028	0.3304	32000	28756
	4	**Electrified 750m single line**	-	35	55	55	1000	79	22.2	17.7	0.028	0.3304	12000	11297
	5	**Electrified 450m double line**	-	135	55	55	600	47	17.8	14.1	0.028	0.3304	7200	6721
	6	**Electrified 750m double line**	-	135	55	55	1000	79	22.2	17.7	0.028	0.3304	12000	11297
	7	**Electrified 750m quadruple line**	-	220	55	55	1000	79	22.2	17.7	0.028	0.3304	12000	11297
Type 3 - Intermodal terminal links	1	**Intermodal terminal**	1	-	0.15	0.5	1	1	6.5	17.5	0.028	0.3304	150	1560

Table 32 – Rail link improvement costs

Link improvement costs [million euro / km] From / To	**1 - Non-electrified 450m single line**	**2 - Electrified 450m single line**	**3 - Non-Electrified 750m single line**	**4 - Electrified 750m single line**	**5 - Electrified 450m double line**	**6 - Electrified 750m double line**	**7 - Electrified 750m quadruple line**
0 - Possible rail line	5.1	5.6	5.7	6.2	7.2	8	10.3
1 - Non-electrified 450m single line	-	2	2.3	3.5	5.1	6.4	9.9
2 - Electrified 450m single line	-	-	-	2.5	4.1	5.4	9.4
3 - Non-Electrified 750m single line	-	-	-	2	-	5.3	9.3
4 - Electrified 750m single line	-	-	-	-	-	4.3	8.9
5 - Electrified 450m double line	-	-	-	-	-	3.6	8.5
6 - Electrified 750m double line	-	-	-	-	-	-	6.7

5.2.3. O/D matrices

The construction of the O/D matrices for the freight demand was based on various data sources, with the basic assumption that all intermodal cargo was generated at the major ports and represented the containerized cargo that is transported by those ports. Although some intermodal cargo is not generated by the major ports, such as containers that are handled by small ports or swap bodies and containers that are not used for sea-shipping operations, this is a reasonable assumption, given the macro nature of the model. We chose to measure containerized cargo in TEUs, as this is the most commonly used unit of measurement for this type of cargo. All remaining cargo, which was considered general cargo, was measured in tonnes.

The construction of the freight demand matrices, which was detailed in chapter 4, employs various data sources, with the construction of the O/D matrices being based on a gravity model. The construction of the O/D matrix for the movement of passenger trains within the network used data on rail traffic for the Portuguese rail network in 2008, which was provided by REFER, and the timetables for passenger train services in Spain in 2008, which was provided by the Spanish rail operator Renfe Operadora. The O/D matrices for the baseline scenario of general cargo and for intermodal cargo, as well as the matrix of the movement of passenger trains within the network can be consulted in annexes A14 to A16.

5.2.4. Calibration

The initial assignment of traffic was calibrated based on real rail traffic data, in order to estimate the value of the μ parameter in Equation (1) and to determine whether the adopted costs delivered realistic results. The calibration was performed using the rail traffic data for the Portuguese network that was provided by REFER, specifically for the lines that cross the borders between the regions of each centroid, including the borders between Spain and Portugal, and using rail traffic data for the easternmost Spanish-French border (Arroyo, 2009). This calibration process was detailed in chapter 4, with the final calibrated values for the baseline scenario being presented in Table 31. The adopted value for the μ parameter was *0.14*.

5.3. Scenarios

To run the optimization model in the described network, it is necessary to define the universe of links that can be built or improved. This universe should be vast enough to consider different possible alternatives, but should also be as limited as possible to avoid wasting computational time testing unfeasible solutions. To reach a desirable compromise, we chose to consider only the major freight corridors, as defined by the European Commission (Comission, 2011), and other relevant links. The set of improvable links is represented by the shaded rail links in Figure 21.

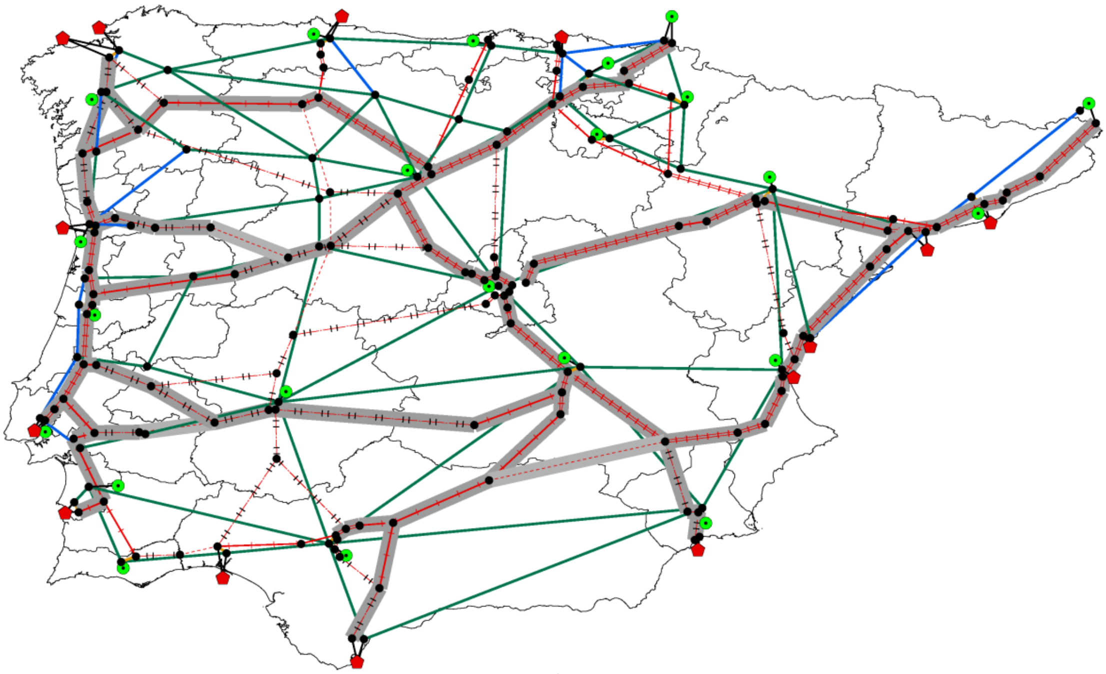


Figure 21 –Improvable links

The weights that were assigned to the minimization of generalized transport costs and CO2 emissions were equal, reflecting a policy that gives an equal weight to both of these goals. This approach is supported by the results obtained from the application of the model on chapter 4.

The different scenarios that were tested in this application were conceived to simulate the influence that important changes to some key variables have on freight traffic and on the impact of the rail network improvements. The key variables that were taken into account were the demand for freight transport and the price of oil, which is reflected in the vehicle costs. The choice of these variables was justified by their significant impact on the transport of freight and their potential volatility. This volatility is confirmed by the significant changes in the demand for inland freight transport in Spain, which grew by *68,4%* between 2000 and 2007, and shrank by *18,7%* between 2007 and 2009 , with Portugal following a similar trend (Eurostat, 2011a). As for the average price of a barrel of Brent crude oil, it went from 31€ in 2000 to 66€ in 2008, down to 45€ in 2009, and increasing again to around 87€ in 2012 (ECB, 2013).

As shown in Table 33, three different settings were considered for the freight demand and two for the vehicle costs. This, combined with the contemplation of two different alternatives for available budget for rail network improvements resulted in a total of twelve scenarios.

Table 33 –Scenarios

Scenario	Freight demand	Vehicle costs	Availible budget
1	Baseline conditions	Baseline conditions	10000 M€
2	50% increase in all freight demand	Baseline conditions	10000 M€
3	100% increase only in intermodal freight demand	Baseline conditions	10000 M€
4	Baseline conditions	25% increase for diesel powered transport	10000 M€
5	50% increase in all freight demand	25% increase for diesel powered transport	10000 M€
6	100% increase only in intermodal freight demand	25% increase for diesel powered transport	10000 M€
7	Baseline conditions	Baseline conditions	20000 M€
8	50% increase in all freight demand	Baseline conditions	20000 M€
9	100% increase only in intermodal freight demand	Baseline conditions	20000 M€
10	Baseline conditions	25% increase for diesel powered transport	20000 M€
11	50% increase in all freight demand	25% increase for diesel powered transport	20000 M€
12	100% increase only in intermodal freight demand	25% increase for diesel powered transport	20000 M€

Three different settings were considered for freight demand: baseline conditions, representing the current demand for freight (data from 2008); a *50%* increase in all freight demand, simulating the possibility of a significant increase in the demand for freight transport in the future; and a *100%* increase only in intermodal freight demand, simulating the possibility of a large increase in the amount of containerized cargo handled by the ports, which is consistent with the significant growth that Iberian ports have witnessed in the last years. The two possibilities for vehicle costs were baseline conditions, representing the current vehicle costs (data from 2008), and a *25%* increase in the vehicle cost per km (excluding the costs associated with road tolls and rail infrastructure charges) of all the diesel powered transport modes, namely road transport and un-electrified rail transport. This is intended to simulate a significant increase in the price of oil, assuming that its impact on the vehicle costs of electric powered locomotives would be negligible. The order of magnitude of this increase should roughly correspond to a *140%* increase over the prices in 2008, assuming that a *1%* increase in the price of oil causes a *0,5%* increase in the price of diesel (a 31% increase in the price of oil between 2008 and 2012 (ECB, 2013) was reflected in a 15% increase in the price of diesel in Portugal (PORDATA, 2013)) and that the cost of fuel accounts for approximately *35%* of the total vehicle costs (FTA, 2013), excluding the costs associated with road tolls and rail infrastructure charges. Two different budgets were tested: a network investment budget of 10 000 million euros, which is a reasonable value for a network of this size; and a network investment budget of 20 000 million euros, which should be large enough to cause a significant improvement to the quality of the rail network.

It is important to stress that the considered scenarios are simplified versions of some of the changes in key factors that could occur in the future, and are not meant to be a detailed study about all the possible changes that could affect freight transport in the future. The number of scenarios needs to be limited, in order for them to be individually run by the network optimization program in a reasonable amount of time. Each different setting is intended to simulate a broad change that could affect a key variable in the future, not a very specific or localized variation, which is consistent with the macro nature of the network optimization model.

5.4. Results and robustness analysis

5.4.1.Network optimization results

The network optimization model was applied to the twelve different scenarios under the following conditions: the traffic was introduced into the network in 10 interactions, each local search process consisted of 50 cycles and the shaking process comprised 25 cycles. This process was run in a core i7 2.0 GHz processor, with four scenarios being run at the same time, which took approximately five days for each run and a total of approximately fifteen days for all the scenarios, which is a reasonable computing time for a strategic planning model of this nature. The duration of the optimization process is mainly justified by the time spent running the traffic assignment process, as the optimization process is quite fast and straightforward. The summary of the results for all the scenarios is listed in Table 34, and the detailed results can be consulted in annexes A19 to A30.

Table 34 –Rail freight traffic and network improvement results

Scenario	**Rail freight traffic before network improvements**	**Network improvement results** (relative to initial conditions for each scenario)			**Rail freight traffic after network improvements**	
	(relative to baseline conditions)	**Change in total generalized transport costs**	**Change in total CO2 emissions**	**Weighted change**	*(relative to initial conditions for each scenario)*	*(relative to baseline conditions)*
1	***baseline conditions***	+0.39%	-4.53%	**-2.07%**	*+58%*	*+58%*
2	**+39%**	+0.50%	-4.89%	**-2.19%**	+68%	+134%
3	**+14%**	+0.27%	-4.94%	**-2.34%**	+59%	+81%
4	**+69%**	-0.74%	-8.45%	**-4.59%**	+68%	+184%
5	**+123%**	-0.60%	-8.37%	**-4.49%**	+71%	+283%
6	**+91%**	-0.96%	-8.65%	**-4.80%**	+80%	+244%
7	***baseline conditions***	+0.49%	-7.45%	**-3.48%**	*+98%*	*+98%*
8	**+39%**	+0.66%	-7.58%	**-3.46%**	+111%	+193%
9	**+14%**	+0.25%	-8.04%	**-3.90%**	+99%	+127%
10	**+69%**	-1.68%	-13.37%	**-7.52%**	+104%	+246%
11	**+123%**	-1.34%	-13.80%	**-7.57%**	+128%	+408%
12	**+91%**	-2.03%	-14.29%	**-8.16%**	+110%	+303%

The rail freight traffic was estimated based on the total amount of freight that passes through fourteen critical rail lines in the Iberian network, which are indicated in Figure 22. These fourteen points provide a representative sample of the Iberian rail traffic, due to their importance and strategic location.

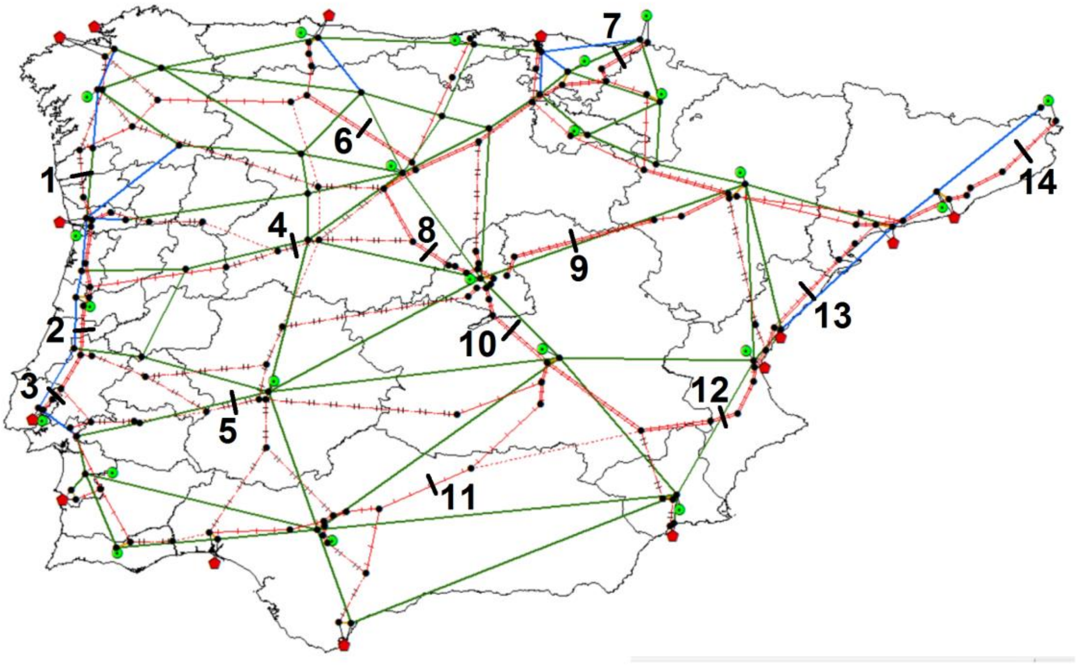


Figure 22 –Traffic measurement points

The first aspect that deserves scrutiny is the impact that the different scenarios have on the distribution of freight traffic. By analyzing the initial rail freight traffic for each scenario, it is possible to see that the different conditions produce important variations in rail freight traffic. This indicates that even if no rail network improvements are performed, the amount of rail traffic that will use the Iberian network in the future will be highly influenced by changes that are beyond the control of planners. The results indicate that a *50%* increase in total freight demand will cause an increase of only *39%* in rail freight traffic, indicating that rail transport will tend to lose market share to road transport as the freight demand increases, which is likely due to capacity limitations at specific points. The doubling of container port traffic creates a *14%* increase in rail traffic, while the *25%* increase in vehicle costs for diesel powered transport causes a *69%* increase in rail traffic. This shows that a sharp increase in the price of oil may have a greater impact in the rail freight traffic than a significant variation of freight demand. The combination of the *50%* increase in freight demand with the increase in oil prices causes a *123%* increase in rail freight traffic, which is more than the sum of the traffic increases that are caused by these changes individually. This occurs because the increase in the price of oil renders rail freight more competitive, absorbing a greater share of the increase in freight demand, even though it is constraint by capacity limitations.

The rail network improvements caused a combined weighted change between *-2.07%* and *-8.16%*. It is important to note that in half of the scenarios, there was an increase in total generalized transport cost, which is an occurrence that was also observed in the application performed in chapter 4 and is due to the type of assignment technique that is used for general cargo, combined with the fact that road transport is, on average, more competitive than rail transport. This indicates that the improvements to the rail network may increase the total generalized costs if the average cost of using rail remains higher than the average cost of road transport, and as long as the increase in generalized costs due to the higher modal share of rail is greater than the reduction that is attributed to lower rail costs. However, this does not occur in scenarios that consider a significant increase in the price of oil, where there are reductions in the total generalized cost. This occurs because the increase in the vehicle costs for diesel powered transport modes causes a substantial surge in the costs of road transport, rendering rail transport more competitive, causing the rail network improvements to reduce the total generalized costs. The network improvements also have a significant impact on the rail freight traffic, causing increases ranging from *58%* to *128%*, with changes compared to the initial baseline conditions ranging from *58%* to *408%*. This five-fold increase of rail freight traffic indicates that this type of traffic may experience considerable changes in the future, depending on the investments that are made as well as on external conditions, such as the evolution of the demand for freight and of the price of oil.

Most of the performed network improvements are increases in the maximum allowable train length. This is consistent with the general European consensus that the maximum length of freight trains should be increased to at least 750 m. There are also a few capacity improvements, namely the quadruplication of some lines around Madrid and of a short stretch near Valencia, and the duplication of the single track line between Zaragoza and Reus and on of a short single track stretch in the Mediterranean line, near Tarragona. These capacity improvements affect only a few specific points, which is due to the fact that the majority of Iberian long-distance lines are not congested.

5.4.2.Robustness analysis

Given the difficulty of making any accurate prediction about the future evolution of the price of oil or of the demand for freight, it is advisable to favor improvement solutions that are as robust as possible, capable of coping with different possible future scenarios. Consequently, we studied two robust network improvement solutions: one for investments up to 10 000M€ (scenarios 1 to 6) and another for investments up to 20 000 M€ (scenarios 7 to 12). The method used for the definition of the robust network improvement solutions was to only implement the improvement operations that were employed in at least *80%* of the scenarios, which delivers quite robust solutions. This *80%* threshold was freely defined by us and should not be interpreted as a fixed value. The goal of this robustness analysis is not to define a specific robust solution, based on the *80%* threshold, but to define a method for the creation of robust solutions, which can be based on any threshold value, according to the characteristics and needs of each situation.

The adopted method produced satisfactory results, with robust solutions totaling an investment of 8 821 M€ for investments up to 10 000M€ and of 16 937 M€ for investments up to 20 000M€. This indicates that it is possible to spend approximately *85%* of the total planned investment just by improving links that are part of these robust solutions. This is a sensible way to invest the majority of the total planned investment, limiting the possibility of making prediction errors in the allocation of investment funds.

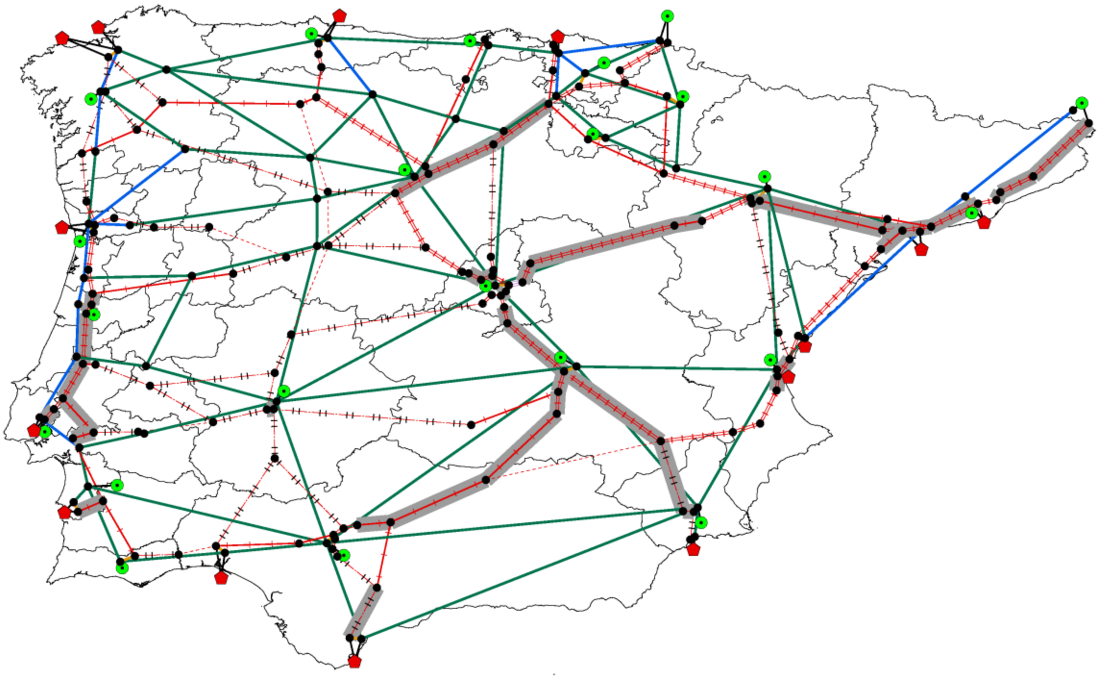


Figure 23 –Robust improvement solution for investments up to 10 000M€

Figure 23 shows that the robust solution for investments up to 10 000M€ improves some of the main rail axes, namely the major line running north-south along Portugal, and the main Spanish lines connecting Madrid with the rest of the country and to Europe. Apart from the quadruplication of a short stretch of double line around Madrid, the only capacity improvements were the duplication of the single track line from Zaragoza to Reus and of the short single track section on the Mediterranean line, near Tarragona.

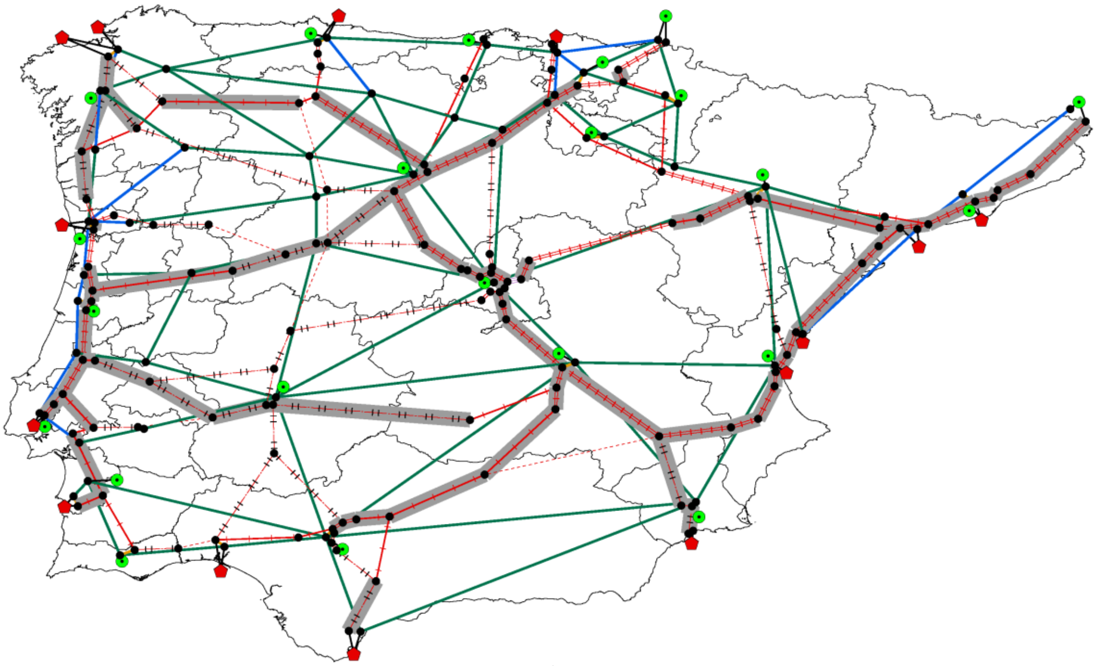


Figure 24 –Robust improvement solution for investments up to 20 000M€

The robust improvement solution for investments up to 20 000M€ is considerably more extensive than the 10 000 M€ solution, including the improvement of two corridors between Portugal and Spain, which also provide a link between Portugal and Europe. Other than those corridors, it is also worth mentioning the significant rail improvements in the region of Galicia, in the northwest of Spain, including the connection between that region and the center of Spain. Additionally, this solution considers two new capacity improvements: the quadruplication of another short stretch of double line around Madrid and of a small stretch close to Valencia.

5.4.3.Discussion

The main conclusions regarding the impact that the different scenarios have on the freight traffic are that, all other factors being equal, road transport will benefit the most from an increase in freight demand, while rail will benefit greatly from a significant increase in the price of oil. The former is due to the existence of capacity limitations at specific points of the rail network, limiting its ability to cope with significant traffic increases without some network improvements, while the latter is a consequence of road transport's high dependence on oil, as a sharp increase in the price of oil would render rail freight considerably more competitive relative to road freight.

The network improvements create large increases in rail traffic, causing a reduction in total CO2 emissions and having a mixed impact on the total generalized costs. The improvements caused a reduction in generalized costs in the scenarios that consider a higher price of oil, and an increase in all the other scenarios. This is justified by the fact that a significant increase in the price of oil makes rail transport much more competitive with road transport. The comparison between scenarios 1 to 6 and 7 to 12 reveals that doubling the investment resulted in *58%* to *70%* improvements in the weighted reduction in costs and CO2 emissions. These numbers reveal a gradual reduction in the benefit that is gained from each invested monetary unit, as each further improvement is less critical than the previous one. The majority of the improvements are increases in the maximum allowable train length from 450 m to 750 m, with only a few capacity improvements. These results are consistent with the need to run longer freight trains to make them more competitive, and with the characteristics of the Iberian network, which has few congestion problems.

The results from the robustness analysis indicate that it is possible to spend approximately *85%* of the planned budget investing only on improvement operations that are part of the defined robust solution, which is a sensible way to allocate funds for long-term infrastructure investments, as it will limit the risk from future uncertainty. This risk limitation is extremely important, due to fact that it is very difficult to predict the long term evolution of many critical factors that have an important influence on the demand and modal distribution of freight transport. Although these results cannot be extrapolated to other networks and different scenarios, the method used for the definition of robust network improvement solutions can be used in any network under any set of scenarios, in order to create robust solutions in which to invest an important part of the total network investment budget.

Finally, it is important to note that all the scenarios were run using the parameters that were obtained from the calibration of the baseline scenario. However, there is the possibility that a significant improvement of the rail network or a surge in the price of oil or in the demand for freight would modify behaviors, or lead to the introduction of disruptive new technologies. Therefore, the interpretation of results should consider the degree of uncertainty that is associated with all long-term analyses.

5.5. Conclusions

In this chapter, the developed rail network optimization model for freight transport is applied to the transport network of the Iberian Peninsula, considering twelve different scenarios. The scenarios simulate conceivable future variations on certain key variables to study the influence of these changes on the distribution of traffic and on the impact of the rail network investments.

The twelve scenarios contemplated three different settings for freight demand, simulating future increases in freight demand, two different settings for the vehicle costs, simulating a significant increase in the price of oil, and two different network investment budgets. The application of the optimization model considering the various scenarios revealed that the impact of the network improvements varies significantly from one scenario to another, although there is an important increase in rail freight traffic in all of the scenarios. The factor that appears to have the greatest influence on freight transport and on the impact of future network improvements is a possible sharp increase in the price of oil, which would render rail freight much more competitive, increasing the positive effects from investing in rail network improvements. The broad nature of the network improvements is essentially the same for all scenarios: an increase in the maximum allowable train length to 750 m in the main transport axes, and some localized capacity improvements.

A robustness analysis was performed on the obtained results for the twelve different scenarios. The obtained results indicate that it is possible to allocate an important part of the network improvement budget just by investing on very robust network improvement operations. This is a wise method for allocating a significant portion of the total planned investment for long-term infrastructure investments, as it limits the risks derived from the uncertainty around the long term evolution of critical factors that have an important influence on the demand and modal distribution of freight transport.

Chapter 6

NETWORK OPTIMIZATION PROGRAM

6. NETWORK OPTIMIZATION PROGRAM

6.1. Introduction

This chapter presents a brief description of the network optimization program that was developed in the scope of this thesis. The program assesses the optimal way in which to perform strategic investments in rail transport networks, in order to improve the conditions for the transport of freight. As such, it is a useful strategic planning tool which can be used in the future to address different planning problems. The purpose of this chapter is to provide a basic explanation on how to operate the program, namely on how to insert data and read the obtained results, in order to facilitate its potential future use by different users. This is justified by the fact that the model does not have a user-friendly interface, which means that the users have to be relatively familiarized with it in order to use it.

This chapter is structured in five sections. The following section is devoted to a brief description on the structure of the program. The third section is dedicated to the insertion of data in the model, while the fourth section explains how to read the obtained results. The final section is dedicated to the concluding remarks, including a suggestion for future improvements.

6.2. Structure of the program

The network optimization program was written in C++ programing language and is divided in two interconnected source files, which follow the structure of the program. The main source file contains the network optimization process and makes the connection with the secondary source file, which runs the traffic assignment process. Every time that the network optimization process needs to assign the traffic to the network, it calls the secondary source file, which delivers the results of the assignment process. The programing code for the rail network optimization process and the associated traffic assignment process, can be consulted in annexes A31 and A32, being configured for scenario 1 from chapter 5.

The data inputs necessary to run the model, namely the characteristics of the transport network and the demand matrices, are inserted in the code of the main source file using a straightforward process. The conditions under which the optimization program is performed are manually inserted each time the program is run. As for the output of the program, it is presented in a text document which contains the various steps performed during the program, as well as the final results.

While it is possible to change the optimization structure of the program in order to adapt it to different network conditions with diverse improvement possibilities, that process involves making slight changes to the improvement possibilities in the programming code. Although this is not a complex process, it requires some basic programming skills, as the user has to be familiarized with C++ programming language.

6.3. Inserting data

The data inputs relative to the characteristics of the transport network and the demand matrices are inserted in the code of the main source file in the appropriate section, named *Class InsertData*. The necessary inputs are: total number of nodes; total number of links; O/D matrix of general cargo; O/D matrix of intermodal cargo; O/D matrix of passenger trains; matrix containing the characteristics of all the links in the network; matrix containing the rail link improvement costs; matrix containing the intermodal terminal link improvement costs (if applicable); matrix containing the virtual link improvement costs (if applicable); matrix containing the general characteristics of the rail links; matrix containing the general characteristics of the intermodal links (if applicable); matrix containing the general characteristics of the virtual links (if applicable).

The matrix containing the characteristics of all the links in the network follows the structure presented in Table 35. The exact meaning of each variable can be consulted in Table 4, from chapter 2.

Table 35 –Structure of the matrix containing the characteristics of all the links in the network

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
		ID	**Type**	**Improvable**	**LinkStatus**	**Startnode**	**Endnode**	**Length**	**Speed_gen**	**Veicap_gen**	**Veicost_gen**	**VOT_gen**	**Veicost_int**	**Capacity_each_way**	**CO2_km_gen**	**Veicap_int**	**Bidirectio**	**Speed_int**	**CO2_km_int**	**VOT_int**
1	**Link 1**																			
...	**...**																			
n	**Link n**																			

As for the matrices containing the link improvement costs, they follow the structure presented in Table 36.

Table 36 –Structure of the matrices containing the link improvement costs

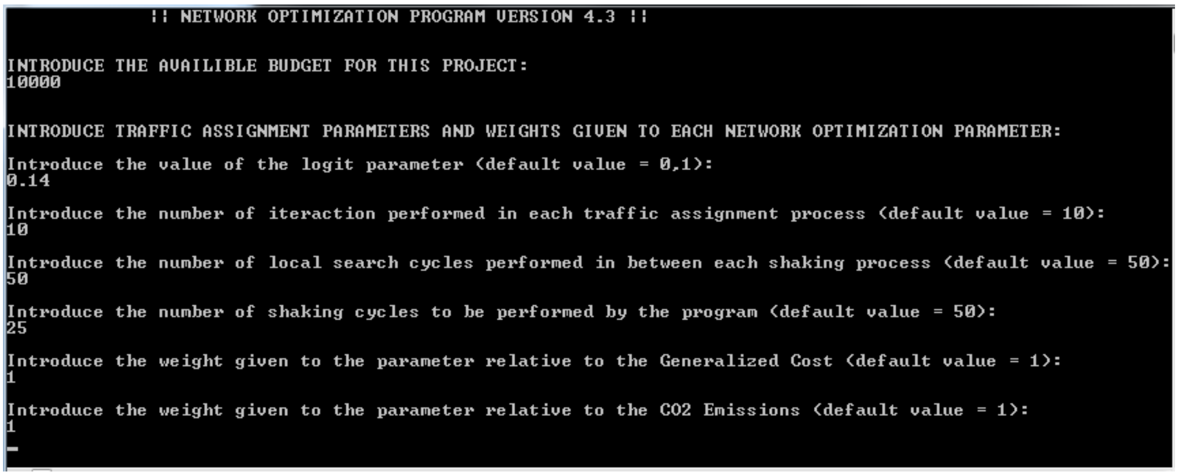
		1	...	q
	From / To	**Link level 1**	**...**	**Link level q**
1	**Link level 0**			
...	**...**			
q	**Link level q-1**			

The matrices containing the general characteristics of the different link types follow the structure presented in Table 37.

Table 37 –Structure of the matrices containing the general characteristics of the different link types

		1	2	3	4	5	6	7	8	9	10	11	12	13
		Length	**Speed_gen**	**Veicap_gen**	**Veicost_gen**	**VOT_gen**	**Veicost_int**	**Capacity_each_way**	**CO2_km_gen**	**Veicap_int**	**Bidirectio**	**Speed_int**	**CO2_km_int**	**VOT_int**
1	**Link level 1**													
...	**...**													
q	**Link level q**													

After having inserted all the data inputs relative to the characteristics of the transport network and the demand matrices, the program is ready to be run. Every time that it is run, the program asks for the definition of certain conditions: available investment budget; value of the μ logit parameter; number of iterations performer in each traffic assignment process; number of local search cycles; number of shaking cycles; weight given to the minimization of the total generalized transport costs; weight given to the minimization of the total CO2 emissions. This data is inserting using a pop-up window, as depicted in Figure 25, which displays the conditions for scenario 1 from chapter 5.



```
            !! NETWORK OPTIMIZATION PROGRAM VERSION 4.3 !!

INTRODUCE THE AVAILIBLE BUDGET FOR THIS PROJECT:
10000

INTRODUCE TRAFFIC ASSIGNMENT PARAMETERS AND WEIGHTS GIVEN TO EACH NETWORK OPTIMIZATION PARAMETER:
Introduce the value of the logit parameter (default value = 0,1):
0.14

Introduce the number of iteraction performed in each traffic assignment process (default value = 10):
10

Introduce the number of local search cycles performed in between each shaking process (default value = 50):
50

Introduce the number of shaking cycles to be performed by the program (default value = 50):
25

Introduce the weight given to the parameter relative to the Generalized Cost (default value = 1):
1

Introduce the weight given to the parameter relative to the CO2 Emissions (default value = 1):
1
_
```

Figure 25 –Inserting initial data

The program will then run in the background until it reaches the final network optimization solution.

6.4. Results display

The final results are presented in a text file which contains all the initial input data, the initial traffic conditions before the improvements, and the various steps of the improvement process. An example of an output text file can be consulted in Annex A19, which presents the results for scenario 1 from chapter 5.

The output file is divided in four sections: a first section containing data from the initial network conditions, before any improvement operation; a second section with the results and network conditions after the application from the initial greedy algorithm; a third section displaying the partial results from the iterative optimization process; and a fourth section with the final results and network conditions after the optimization process.

The description of the network conditions in each stage includes the detailed characteristics of all the links in the network and the total traffic flow in each link, as well as the total generalized transport costs and CO2 emissions. The results for the initial greedy algorithm and for the global optimization process display the change in the two optimization variables as well as the network improvement solution, detailed by the type of improvements performed in each link. As for the partial results from the iterative optimization process, they reveal the combined weighted change of the optimization parameters after each local search iteration as well, as the performed link improvements and link reversals.

6.5. Conclusions

This chapter presents the network optimization program that was developed in the scope of this thesis, with the goal of providing a brief explanation on how to operate the program. This explanation is intended to facilitate its future use by different users, describing the general structure of the program, the process of inserting input data, and how to read the obtained results.

The operation of the program under the current network structure is a straightforward process. Even so, it is not very easy to change the network structure in order to adapt it to different network conditions with diverse improvement possibilities, as that implies changes to the programming code. Although these changes are not complex, they require the user to be familiarized with C++ programming language. This limitation of the program may be addressed in the future with the development of a user friendly interface. Such an interface would enable users to define the desired network structure and to insert the needed input data without having to access the programming code, which would facilitate its use by people not familiar with the program.

Chapter 7

CONCLUSIONS AND FUTURE DEVELOPMENTS

7. CONCLUSIONS AND FUTURE DEVELOPMENTS

7.1. Conclusions

This thesis addresses the strategic planning of investments in rail transport infrastructures, with the goal of improving the conditions for the transport of freight. The focus on freight transport is the main differentiating factor of this work, and it required the development of planning tools designed specifically for this type of transport. The major contribution of this thesis was the development of a strategic rail network optimization model for freight, aimed at helping transport planners to make strategic network investment decisions.

The four main objectives that were set at the beginning of this thesis were fully accomplished, delivering a positive final outcome. The first objective was the development a traffic assignment model for freight, designed to model transport networks at a strategic planning level, which was described in chapter 2. The second objective was the creation of a strategic rail network optimization model for freight, capable of determining how to invest a specific volume of capital on a rail network, in order to improve the conditions for freight transport, as detailed in chapter 3. The third objective was the application of the network optimization model on a real network, in order to properly validate and calibrate the traffic assignment model and to analyze the results from the network optimization, which was done in chapter 4. Lastly, the forth objective was to apply the network optimization model on a real network considering possible modifications on some key variables that affect freight, in order to study the influence of these changes on the distribution of traffic and on the impact of the network improvements, which was described in chapter 5.

The divulgation of the work developed in this thesis also followed the initial plans, with the production of a scientific paper for each of the chapters from 2 to 5. The first paper describes the traffic assignment model developed in chapter 2, and was published in the *Computers in Industry* journal. The second paper is dedicated to the network optimization model developed in chapter 3, and was accepted for publication by the *Transport Research Record* journal. The third paper and fourth papers are dedicated to the applications of the network optimization model that are described in chapters 4 and 5, respectively. They are both research papers which we plan to submit for future publication.

The most relevant findings that were drawn from the work developed in this PhD thesis are described in the concluding section of each chapter. Nevertheless, it is important to summarize the main conclusions that emanate from this PhD thesis.

The first main conclusion is that the freight and passenger transport have very distinct characteristics, thus the impact that different network investments have on each of those types of transport may be very different. Therefore, when planning for future investments in transport networks, it is important to use models adapted to each type of transport.

Second, when planning for long term investments on transport infrastructure, it is advisable to contemplate other evaluation parameters besides the simple minimization of total transport costs, as it may be unfeasible to evaluate solutions based on just one parameter. The inclusion of a complementary parameter, such as reduction of CO2 emissions, results in a more comprehensive analysis.

The third and final conclusion is that the future is uncertain and it is thus impossible to accurately predict the future evolution of key factors that affect freight transport, such as the demand for this kind of transport and the price of oil. This latter factor has an important influence on the distribution of freight traffic, as a sharp increase in the price of oil would render rail freight much more competitive, increasing the positive effects from investing in rail network improvements. Thereby, when planning for long term investments in transport infrastructures, it is important to consider different possible evolutions for those key factors, as they may have a decisive influence on the impact caused by the network investments. This uncertainty problem can be addressed by performing a robustness analysis, in order to minimize the risk of making network investments that may become obsolete if there are changes in some critical variables.

7.2. Future developments

Although the work developed in this PhD thesis delivered a very positive outcome, resulting in the creation of a traffic assignment model and a strategic rail network optimization model for freight, there is still room for future improvement. The main possibilities for future developments can be separated on three main topics, namely the traffic assignment model, the rail network optimization model and the resulting strategic network optimization program.

The developed traffic assignment model for freight has room for future improvement on four significant aspects. The first one is the assignment technique used for intermodal cargo, which may be improved by considering some sort of traffic distribution between various possible routes, which would involve the development of a system for the creation of one or more feasible alternative routes. The second aspect with room for improvement is the limited number of cargo types that are contemplated in the model. This may be improved by separating general cargo into various sub-categories and possibly considering different assignment techniques for the various sub-categories of cargo, according to their characteristics. The third improvable aspect is the estimation of empty trips, which is currently not taken into consideration. The explicit estimation of empty freight trips would make the model more realistic, by improving the accuracy of its traffic estimations. To finalize, the fourth improvement possibility would be the integration of the model with a passenger traffic assignment model. This would lead to the creation of a comprehensive strategic surface transport model for both passengers and freight transport, which would be an extremely valuable planning tool.

As for the developed strategic rail network optimization model for freight, there are two main aspects that may be improved the future. The first one would be the inclusion of a time analysis, which could ultimately lead to the creation of a network optimization calendar for a given time span that would be dependent on the available investment for each period. This should include an estimation of future variations on the demand for freight and possibly other parameters, and a method to evaluate the benefit of each improvement operation over time. The full implementation of a time analysis would enable the model to deliver gradual improvement solutions that would be dependent on the available investment for each period of time. The second improvement possibility would be the inclusion of passenger transport, which is dependent on the integration of passenger transport in the assignment model, creating a strategic rail network optimization model for both passengers and freight transport.

To finalize, the strategic network optimization program that was developed in the scope of this thesis may be improved with the development of a user friendly interface. This would facilitate its use by different users, as they would no longer need to be familiarized with the programming code in order to make changes to the network conditions. The development of a user friendly interface would enable users to define the desired network structure and to insert the needed input data without having to access the programming code.

Future developments in these relevant areas of research will further improve the planning tools that were developed in this PhD thesis. Nevertheless, we believe that the developed work already offers a valuable contribution to the strategic planning of investments in transport infrastructure.

REFERENCES

Abdelwahab, W., Sargious, M. (1992) 'Modelling the Demand for Freight Transport', Journal of Transporte Economics and Policy, January, pp. 49-70

Abdelwahab, W.M. (1998) 'Elasticities of Mode Choice Probabilities and Market Elasticities of Demand: Evidence From a Simultaneous Mode Choice/Shipment-Size Freight Transport Model', Transportation Research Part E, 34(4), pp. 257-266

ADIF (2009) 'Declaración sobre la Red 2009'.

Affuso, L., Masson, J., Newbery, D. (2000) 'Comparing Investments on New Transport Infrastructure: Roads vs. Railways?', Department of Applied Economics, University of Cambridge.

Alvarez, A.G., Aguilar, J.M.J., Vijande, I.P., Santos, L.E.M., Melero, R.R., Pardo, A.O.S., Cañizares, P.M., Vega, M.J. (2010) 'Informe 2010 - Observatorio del Ferrocarril en España', Fundación de los Ferrocarriles Españoles.

Andersen, J., Christiansen, M. (2009) 'Designing new European rail freight services', Journal of the Operational Research Society, 60, pp. 348-360

Arnold, P., Peeters, D., Thomas, I. (2004) 'Modelling a rail/road intermodal transportation system', Transportation Research Part E, 40, pp. 255-270

Arroyo, F. (2009) 'El tren de mercancías no arranca', El País (06/07/2009).

Ashtakala, B., Murthy, A.S.N. (1988) 'Optimized gravity models for commodity transportation', Journal of Transportation Engineering, 114(4), pp. 393-408

Baumgartner, J.P. (2001) 'Prices and Costs in the Railway Sector', École Polytecnique Fédérale de Lausanne.

Bereskin, C.G. (2001) 'Sequential Estimation of Railroad Costs for Special Traffic', Transportation Journal, 40(3), pp. 33-45

Beuthe, M., Jourquin, B., Geerts, J.-F., Há, C.K.à.N. (2001) 'Freight Transportation Demand Elasticities: A Geographic Multimodal Transportation Network Analysis', Transportation Research Part E, 37, pp. 253-266

Bilegan, I.C., Crainic, T.G., Gendreau, M. (2007) 'Forecasting freight demand at intermodal terminals using neural networks – an integrated framework'.

BOE (2003) 'Boletín Oficial del Estado - Viernes 30 mayo 2003 - Pags. 21100-21102', Agencia Estatal Boletín Oficial del Estado.

Bolis, S., Maggi, R. (2003) 'Logistics Strategy and Transport Service Choices: An Adaptive Stated Preference Experiment', Growth and Change, 34, pp. 490-504

Bontekoning, Y.M., Macharis, C., Trip, J.J. (2004) 'Is a new applied transportation research field emerging?—A review of intermodal rail–truck freight transport literature', Transportation Research Part A, 38, pp. 1-34

Cefic, ECTA (2011) 'Guidelines for Measuring and Managing CO2 Emission from Freight Transport Operations'.

Chen, M., Alfa, A.S. (1991) 'A Network Design Algorithm Using a Stochastic Incremental Traffic Assignment Approach', Transportation Science, 25(3), pp. 215-224

Comission, E. (2011) 'Handbook on the Regulation concerning a European rail network for competitive freight (Regulation EC 913/2010)'.

Cook, P.D., Das, S., Aeppli, A., Martland, C. (1999) 'Key Factors In Road-Rail Mode Choice In India: Applying The Logistics Cost Approach', 1999 Winter Simulation Conference, Squaw Peak, Phoenix, AZ, USA.

Crainic, T.G. (2000) 'Service Network Design in Freight Transportation', European Journal of Operational Research, 122, pp. 272-288

Crainic, T.G., Florian, M., Léal, J.-E. (1990) 'A Model for the Strategic Planning of National Freight Transportation by Rail', Transportation Science, 24(1), pp. 1-24

Crainic, T.G., Laporte, G. (1997) 'Planning Models for Freight Transportation', European Journal of Operational Research, 97, pp. 409-438

ECB (2013) 'Statistical Data Warehouse - Oil price', European Central Bank.

Eurostat (2009) 'Panorama of Transport', Eurostat - Statistical books.

Eurostat (2011a) 'Energy, transport and environment indicators', Eurostat - Pocketbooks.

Eurostat (2011b) 'Eurostat regional yearbook 2011', Eurostat - Statistical books.

Fernandez, J.E., De Cea, J., Giesen, R. (2004) 'A strategic model of freight operations for rail transportation systems', Transportation Planning And Technology, 27(4), pp. 231-260

Floyd, R.W. (1962) 'Algorithm 97: Shortest path', Communications of the ACM, 5(6), pp. 345

Forkenbrock, D.J. (2001) 'Comparison of external costs of rail and truck freight transportation', Transportation Research Part A, 35, pp. 321-337

FTA (2013) 'Fuel as a percentage of hgv operating costs', Freight Transport Association.

Gallo, M., D'Acierno, L., Montella, B. (2012) 'A meta-heuristic algorithm for solving the road network design problem in regional contexts', 15th meeting of the EURO Working Group on Transportation, Paris, France.

Guélat, J., Florian, M., Crainic, T.G. (1990) 'A Multimode Multiproduct Network Assignment Model for Strategic Planning of Freight Flows', Transportation Science, 24(1), pp. 25-39

Hanssen, T.-E.S., Mathisen, T.A., Jørgensen, F. (2012) 'Generalized transport costs in intermodal freight transport', 15th meeting of the Euro Working Group on Transportation, Paris, France.

Hautzinger, H. (1984) 'The prediction of interregional goods vehicle flows: some new modeling concepts', Ninth International Symposium on Transportation and Traffic Theory, VNU Science Press, Delft, The Netherlands.

Holguín-Veras, J., Thorson, E. (2003) 'Modeling commercial vehicle empty trips with a first order trip chain model', Transportation Research Part B, 37, pp. 129-148

Iding, M.H.E., Meester, W.J., Tavasszy, L.A. (2002) 'Freight trip generation by firms', 42nd European Congress of the Regional Science Association, Dortmund.

Janic, M. (2007) 'Modelling the full costs of an intermodal and road freight transport network', Transportation Research Part D, 12, pp. 33-44

Janic, M. (2008) 'An Assessment of the Performance of the European Long Intermodal Freight Trains (LIFTS)', Transportation Research Part A, 42, pp. 1326-1339

Janic, M. (2009) 'Modelling the Cost Performance of a Given Logistics Network Operating Under Regular and Irregular Conditions', European Journal of Transport and Infrastructure Research, 9(2), pp. 100-120

Janic, M., Reggiani, A., Nijkamp, P. (1999) 'Sustainability of the European Freight Transport System: Evaluation of Innovative Bundling Networks', Transportation Planning And Technology, 23, pp. 129-156

Jeong, S.-J., Lee, C.-G., Bookbinder, J.H. (2007) 'The European freight railway system as a hub-and-spoke network', Transportation Research Part A, 41, pp. 523-536

Jong, G.d., Gunn, H., Walker, W. (2004) 'National and International Freight Transport Models: An Overview and Ideas for Future Development', Transport Reviews, 24(1), pp. 103-124

Jourquin, B. (2005) 'A Multi-flow Multi-modal Assignment Procedure Applied to the European Freight Transportation Networks', Studies in Regional Science, 35, pp. 929-945

Jourquin, B., Beuthe, M. (1996) 'Transportation Policy Analysis with a Geographic Information System: The Virtual Network of Freight Transportation in Europe', Transportation Research Part C, 4(6), pp. 359-371

Jourquin, B., Limbourg, S. (2006) 'Equilibrium Traffic Assignment on Large Virtual Networks: Implementation Issues and Limits for Multi-modal Freight Transport', European Journal of Transport and Infrastructure Research, 6(3), pp. 205-228

Kang, K., Strauss-Wieder, A., Eom, J.K. (2010) 'New Approach to Appraisal of Rail Freight Projects in South Korea - Using the Value of Freight Transit Time Savings', Transportation Research Record, 2159, pp. 52-58

Levine, B., Nozick, L., Jones, D. (2009) 'Estimating an origin–destination table for US imports of waterborne containerized freight', Transportation Research Part E, 45, pp. 611-626

Ma, L., Liang, J., Gao, D., Sun, J., Li, Z. (2012) 'The future demand of transportation in China: 2030 scenario based on a hybrid model', 15th meeting of the EURO Working Group on Transportation, Paris, France.

Maia, L.C., Couto, A.F.d. (2013a) 'Assessing The Impact Of Rail Network Improvements On Freight Transport Under Different Scenarios', Research paper.

Maia, L.C., Couto, A.F.d. (2013b) 'An Innovative Freight Traffic Assignment Model For Multimodal Networks', Computers in Industry, 64(2), pp. 121-127

Maia, L.C., Couto, A.F.d. (2013c) 'A Strategic Rail Network Optimization Model For Freight Transportation', Transportation Research Record (in press).

Maia, L.C., Couto, A.F.d. (2013d) 'Validating A Network Optimization Model For Freight: The Case Of The Iberian Peninsula', Research paper.

McKinnon, A. (2007) 'CO2 Emissions from Freight Transport: An Analysis of UK Data', Commission for Integrated Transport.

Müller, S., Wolfermann, A., Huber, S. (2012) 'A nation-wide macroscopic freight traffic model', 15th meeting of the EURO Working Group on Transportation, Paris, France.

Oum, T.H. (1979) 'Derived Demand for Freight Transport and Inter-modal Competition in Canada', Journal of Transporte Economics and Policy, May, pp. 149-168

Peterson, B.E., Southworth, F. (2010) 'Generation of a U.S. Commodity Flows Matrix Using Log-Linear Modeling and Iterative Proportional Fitting (Freight Analysis Framework 3)'.

PORDATA (2013) 'Preços médios de venda ao público dos combustíveis líquidos e gasosos – Continente em Portugal'.

Racunica, I., Wynter, L. (2005) 'Optimal location of intermodal freight hubs', Transportation Research Part B, 39, pp. 453-477

Rao, P.S. (1978) 'Forcasting the demand for railway freight services', Journal of Transporte Economics and Policy, January, pp. 7-26

REFER (2009) 'Directório da Rede 2010'.

Santos, B., Antunes, A., Miller, E. (2009) 'Multiobjective Approach to Long-Term Interurban Multilevel Road Network Planning', Journal of Transportation Engineering, (September), pp. 640-649

Santos, B., Antunes, A., Miller, E.J. (2008) 'Integrating Equity Objectives in a Road Network Design Model', Transportation Research Record, (2089), pp. 35-42

Santos, B.F., Antunes, A.P., Miller, E.J. (2010) 'Interurban Road Network Planning Model with Accessibility and Robustness Objectives', Transportation Planning and Technology, 33(3), pp. 297-313

Shinghal, N., Fowkes, T. (2002) 'Freight mode choice and adaptive stated preferences', Transportation Research Part E, 38, pp. 367-378

Southworth, F., Peterson, B.E. (2000) 'Intermodal and international freight network modeling', Transportation Research Part C, 8, pp. 147-166

Tavasszy, L.A. (2006) 'FreightModelling – An overview of international experiences', Transportation Research Board Conference on Freight Demand Modelling: Tools for Public Sector Decision Making, Washington DC, USA.

Tsamboulas, D., Moraitis, P. (2007) 'Methodology for Estimating Freight Volume Shift in an International Intermodal Corridor', Transportation Research Record, 2008, pp. 10-18

Vanek, F.M., Smith, R.I. (2004) 'Prospects for Rail Freight from Peripheral Regions: The Case of North-East Scotland and Grampian Country Foods', International Journal of Logistics, 7(1), pp. 59-70

Vierth, I., Lord, N., Daniel, J.M. (2009) 'Representation of the Swedish transport and logistics system', (Logistics Model Version 2.00), VTI - Swedish National Road and Transport Research Institute

Vilain, P.B., Muhammad, T., Li, C., Hallissey, M. (2010) 'Freight and Passenger Modeling Using Hybrid Econometric-Network Models', Transportation Research Record, 2187, pp. 97-105

Wardrop, J.G. (1952) 'Some Theoretical Aspects of Road Traffic Research', Proceedings of the Institute of Civil Engineers, 1(Part II), pp. 325-378

Weidner, T., Hunt, J., Knudson, B., Picado, R. (2009) 'Sensitivity Testing with the Oregon Statewide Integrated Model (SWIM2)', Transportation Research Board 88th Annual Meeting, Washington DC, USA.

Wigan, M., Southworth, F. (2006) 'Whats Wrong With Freight Models and What Should We Do About It?', Transportation Research Board 85th Annual Meeting, Washington DC, USA.

Wong, W.G., Niu, H., Ferreira, L. (2002) 'A Fuzzy Method for Predicting the Demand for Rail Freight Transportation', Journal of Advanced Transportation, 37(2), pp. 159-171

Xu, W., He, S., Song, R., Li, J. (2009) 'Reliability based assignment in stochastic-flow freight network', Applied Mathematics and Computation, 211, pp. 85-94

Yaghini, M., Foroughi, A., Nadjari, B. (2011) 'Solving railroad blocking problem using ant colony optimization algorithm', Applied Mathematical Modelling, 35(12), pp. 5579-5591

Yamada, T., Russ, B.F., Castro, J., Taniguchi, E. (2009) 'Designing Multimodal Freight Transport Networks: A Heuristic Approach and Applications', Transportation Science, 43(2), pp. 129-143

Yang, L., Gao, Z., Li, K. (2011) 'Railway freight transportation planning with mixed uncertainty of randomness and fuzziness', Applied Soft Computing, 11(1), pp. 778-792

Zhang, G., Lu, J. (2008) 'Genetic Algorithm for Continuous Network Design Problem', Journal of Transportation Systems Engineering and Information Technology, 7(1), pp. 101-105

Zlatoper, T.J., Austrian, Z. (1990) 'Freight Transportation Demand - A Survey of Recent Econometric Studies', Transportation, 16(1), pp. 27-46

ANNEXES

A23 – Results for chapter 5 - Scenario 5

A24 – Results for chapter 5 - Scenario 6

A25 – Results for chapter 5 - Scenario 7

A26 – Results for chapter 5 - Scenario 8

A27 – Results for chapter 5 - Scenario 9

A28 – Results for chapter 5 - Scenario 10

A29 – Results for chapter 5 - Scenario 11

A30 – Results for chapter 5 - Scenario 12

A31 – Programming code for the network optimization process

A32 - Programming code for the traffic assignment process

The annexes can be consulted in the accompanying CD.