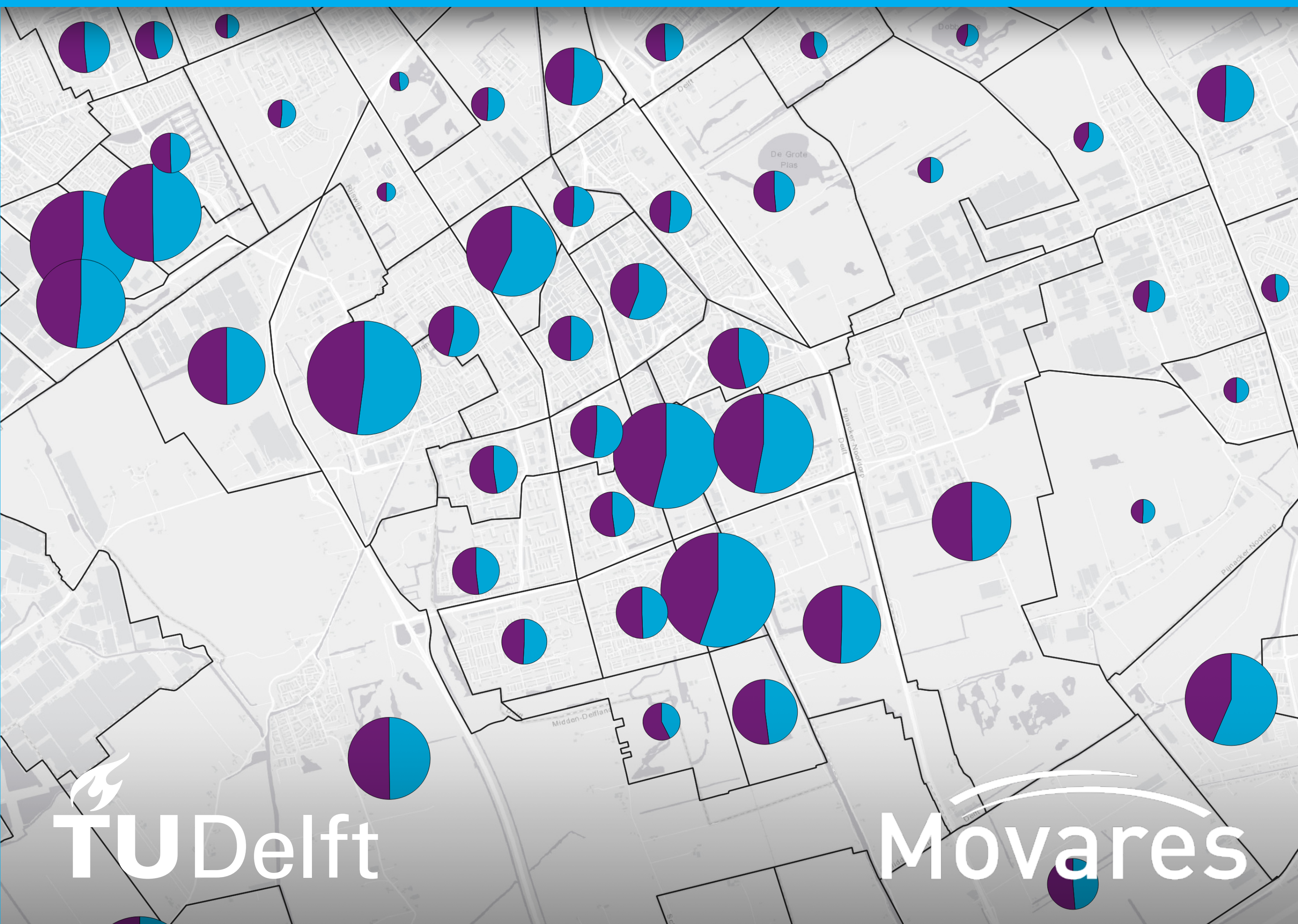
Quantifying freight trip and freight generation from spatial developments in the Netherlands

W.G.J. Gommans



Quantifying freight trip and freight generation from spatial developments in the Netherlands

by

W.G.J. Gommans

to obtain the degree of
Master of Science in Transport, Infrastructure and Logistics
at the Delft University of Technology, to be defended publicly on
Thursday 14:00, 30 May 2024.

Thesis committee:	Prof.dr.ir. L.A. (Lóri) Tavasszy	TU Delft
	Dr.ir. M.A. (Michiel) de Bok	TU Delft
	Dr.ir. M. (Maarten) Kroesen	TU Delft
	Dr. J.A. (Jan Anne) Annema	TU Delft
	Ir. T.S. (Thymo) Vlot	Movares
	J.A.M. (Jessica) van Rijn MSc	Movares

Preface

This master’s thesis, titled ”Quantifying freight trip and freight generation from spatial developments in the Netherlands,” represents the completion of my graduate studies in Transport, Infrastructure and Logistics (TIL) at Delft University of Technology. The research presented addresses the critical need for accurate and reliable quantification of logistics in the Netherlands by identifying key factors influencing freight activities in urban areas, using a freight model and other data sources to inform policymakers for better integration into spatial development. The research was performed under the supervision of my graduation committee and it was made possible by the engineering and consultancy firm Movares. Movares is already well-known for its expertise with index numbers in various fields, such as the parking norms provided by CROW. However, the company is particularly interested in how policymakers could more easily implement logistics in cities and what metrics could be used to quantify these logistics. One of the main concerns today is the presence of postal vans in city centers. This pressing issue prompted a comprehensive and detailed regression analysis, the results of which are presented in this research, marking the beginning of addressing this knowledge gap. I hope this thesis will serve as a useful resource for researchers and policymakers, and that it will inspire further studies to advance our understanding of how freight is generated.

I want to thank my supervisors for their support, clear feedback, patience and guiding me through the process of doing research and writing a thesis. It was a long journey, with a meandering takeoff, but you all kept supervising me with the same enthusiasm which led to me getting also very excited about this study, results and future opportunities. First of all, I want to express my gratitude towards Movares for giving me the opportunity to complete this thesis on an internship basis and simultaneously show me how consulting work is done and what it is like to work at a company like Movares, both formally and informally. Jessica, who guided me to the topic of logistics, always had a sympathetic ear and encouraged me to attend the Vervoerslogistieke Werkdagen, where I got a deep dive into the world of (city) logistics. There, I got a deep dive into the world of city logistics and informally met two of my TU Delft committee members, Lóri and Michiel, initiating the conversations about quantifying freight trips. And most notably Thymo, who, for 14 months now, listened each week to my personal and thesis struggles, my questions about working life or the uncertainties when writing a thesis. He has always been keen on open conversations about all other subjects, work or no work, and reassured me that each topic I chose for my thesis was the right one. Furthermore, I would like to thank the TU Delft’s committee for guiding me towards the end of this final mark of my student life and showing me the urge and beauty of this topic. Lóri, the chair of the committee, consistently provided me with a clear perspective on what is feasible, realistic, and valuable to include in a TIL thesis. His straightforward and honest guidance was exactly what I needed to complete this thesis, and I greatly appreciated his support. Maarten, whose enduring optimism gave me the confidence to finish this project and showed my work is valuable and has great results, but also was clear in his feedback and showed what could be improved. Jan Anne, who jumped in at the last minute to join this committee but also guided me beforehand in writing a research proposal and highlighting areas of improvement in my writing style. And in particular, Michiel, who directed me to this specific and interesting topic, provided me with this very comprehensive dataset called the BasGoed freight tour data, and introduced me to the world of regression models. He also listened to my struggles on a biweekly basis, was always open to answering my questions, provided me with constructive feedback and showed me ways to get this thesis done. Finally, I want to thank all my colleagues at Movares for their warmth in welcoming me, showing me their interesting projects or thesis struggles, and introducing me to the informal side effects of work or doing an internship at Movares.

On a personal note, I express my deepest gratitude to my family and friends for their consistent support and patience throughout this journey. Most certainly for all of my questions, doubts, and again (the same) questions on how to approach certain steps, e-mails or paragraphs in my process. Moreover, the coffee breaks, the 24/7 Python helpdesk, the motivational texts and the interesting discussions have helped me a lot in this journey. Their belief in my abilities, and their being or having been in the same boat, has been a constant source of motivation. Like I said before, this thesis marks the end of my time as a student in Delft. I am very grateful to have been able to study in Delft, all the friends that I made and all the new insights that came with it. It has given me a wonderful time.

W.G.J. Gommans
Delft, May 2024

Summary

This research aims to improve the understanding and quantification of logistics in the Netherlands by identifying explaining factors influencing freight activities in urban areas, using a freight model, location and zonal data to inform policymakers for better integration into spatial development.

Inefficient spatial planning leads to increased congestion and delays in logistics, causing higher costs and environmental impacts. Poor alignment of distribution centers with transportation networks disrupts goods distribution and neighbourhood environments. The rapid expansion of logistics, driven by e-commerce and urbanization, presents challenges for cities, including finding space for logistic functions and managing freight movements. Efforts to promote sustainability, such as zero-emission zones and consolidation centers, require a detailed study on logistics quantification. Without accurate data on logistics movements, municipalities struggle to plan effectively, leading to congestion, safety issues, and inefficiencies. Improved spatial planning tools and index numbers for logistics are needed to optimize freight operations, reduce environmental impacts, and ensure sustainable urban development.

The methodology involves developing a regression model using synthetic data from the BasGoed freight tour module, enriched with CBS location data and Dutch zonal plan data at a disaggregated level. This model incorporated 26 variables such as employment, surface area, and urban density level to predict freight (trip) generation in the Netherlands. The model was validated using statistical tests and comparisons with existing literature, ensuring its reliability and accuracy. Furthermore, a use case on a spatial planning scenario in the A12 zone in Utrecht was conducted to identify the practical purposes of the model. This turned out to be very relevant and led to valuable insights for policymakers. Policymakers can use these insights to design logistics-friendly urban environments and optimize infrastructure. However, the study notes limitations, including certain simplifications of statistical procedures and the reduced model accuracy in seaport and transhipment areas, as well as the possible endogenous relationship between distribution center size and trip generation. Despite these constraints, the model offers valuable predictions for predicting logistics under various spatial scenarios.

By analyzing multiple datasets using linear regression, the study found significant factors for Freight Trip Generation and Freight Generation with over 97.5% confidence. These factors include the surface area of distribution centers, employment in various sectors, the presence of rail or inland waterway terminals, urban density levels, and the surface area of business, industrial, and office spaces. Employment and surface area were particularly impactful, aligning with existing literature. According to the model results, employment has a greater influence on freight production and surface area on freight attraction. The study also shows that age distribution does not affect freight generation, while distribution centers primarily attract freight weight. The developed prediction model offers substantial improvements for spatial planning scenarios, such as urban growth and economic development, by providing accurate logistics movement predictions. The results of this model can inform policymakers and be integrated into other models, like traffic flow models, enhancing urban logistics planning and sustainability.

The study reveals several topics for further investigation. Exploring the model performance using the original CBS XML microdata, which includes location-specific trips in the Netherlands, could validate current findings. Non-linear relationships, such as those identified by Sánchez-Díaz et al., should be considered to refine regression models. Additionally, examining explanatory variables for shipment size and vehicle type could improve the model's practical application and its impact on urban environments. Future studies should also focus on the implications of liveability aspects, such as emissions and congestion, and extend this methodology to datasets of vans or other vehicles. Implementing these insights into policy and practice could significantly benefit urban logistics management. Finally, future work should explore causality between variables to address endogeneity, potentially using instrumental variables to correct bias caused by positive trade-offs between variables.

This study makes significant contributions on both scientific and societal fronts. Scientifically, it introduces a statistically significant model for understanding Freight Trip Generation and Freight Generation, incorporating recent and very rich data and identifying new explanatory factors. Societally, it offers a guide for municipalities to quantify logistics requirements, assisting in effective spatial planning and decision-making.

Contents

List of Figures

List of Tables

List of Abbreviations

1
Introduction

This research aims to enhance the understanding of the quantification of logistics in the Netherlands by identifying key factors influencing logistics generation. Through the utilization of one freight model combined with other data sources, the study seeks to quantify freight activities in urban areas. Additionally, it aims to provide valuable insights for policymakers to optimize logistics integration into spatial development projects.

This chapter addresses background information regarding the problem on which this thesis is focused. After the problems have been described, the scientific relevance that follows from the literature review in chapter 3 will be explained. Moreover, the societal relevance will also be outlined. Those gaps will be translated into a research goal and operationalised with a research design and questions in chapter 2. This chapter continues with an expected contribution of this research and ends with a reading guide.

1.1. Background

Inefficient spatial planning leads to increased congestion and delays in logistic movements, resulting in higher operational costs and environmental impact. For instance, poor alignment (and the construction) of distribution center locations with transportation networks can cause significant bottlenecks ("Overlast door bouw logistieke 'blokkendozen'", 2022), disrupting the overall efficiency of goods distribution and more importantly, the living environment of the neighbourhood. Moreover, solutions need to be found - mostly afterwards - for the many vehicles that regularly stop in the streets ("'Smart zone' wordt mogelijke oplossing voor laad- en losprobleem in Utrechtse Twijnstraat", 2022). Logistics does currently not yet have a fixed place in the process of spatial planning and is therefore often only taken into account implicitly or even afterwards (CROW & Topsector Logistiek, 2023). One of the reasons is the lack of knowledge about the amount of logistics generated by a spatial planning scenario. This problem emphasizes the need for a detailed study on the quantification of logistics to inform better spatial planning decisions, ensuring streamlined logistic operations and minimizing negative consequences on both the economy and the environment.

Over the past few years, logistics has experienced a significant expansion, the share of transport volume to, from and in the Netherlands is 1,925 megatons (Matthijs Otten & Peter Scholten, 2020). With the rise of e-commerce and urbanisation, cities are experiencing a significant increase in freight movements, leading to growing challenges in managing logistical activities. Nonetheless, we find ourselves in a world where almost anything can be conveniently ordered online and delivered right to our doorsteps within minutes, hours, or days, all from the comfort of our homes. However, it is important to acknowledge that alongside parcel and express logistics, which only addresses 3% of city logistics

(Otten et al., 2016), five other components embody modern city logistics. These include general cargo, temperature-controlled logistics, facility logistics, waste logistics, and construction logistics, making up a total of six distinct city logistics segments (den Boer et al., 2017). Rijkswaterstaat[1] even adds three more categories to the Logistic Segment (LS)s, making it a total of nine: food (non-conditioned) and dangerous goods are added, and parcel and express are divided into last-mile flows and consolidated flows between sorting centres ("BasGoed Functionele Documentatie", 2023).

Several vehicles from various logistic segments have become an integral part of our view of city life. However, their integration into city streets is not without challenges. Delivery vans often struggle to find appropriate parking spaces, resulting in frustrations within cities (Couzy, 2016). Furthermore, wrongfully parked vans in narrow streets in an industry where efficiency is the name of the game leads to unsafe situations (Messelink, 2020). In fact, of all accidents involving motorised vehicles, 60% of those injured fell within built-up areas. Heavy goods vehicles and vans are overrepresented in these accidents; two to three times more frequent than accidents involving passenger vehicles (Ploos van Amstel, 2017). Municipalities increasingly show a preference for encouraging low-traffic areas as part of their urban planning strategies (Gemeente Amsterdam, 2023; Gemeente Delft, n.d.; Merwede, 2022). After all, we would all like to stay in a livable environment, or even in a car-free environment or somewhere with a car-free city centre (Kock, 2023). These low-traffic areas typically restrict or discourage private vehicle usage, promoting sustainable modes of transportation such as walking, cycling, and public transit. For city logistics, this preference for low-traffic areas presents both challenges and opportunities. Logistics operators face constraints in terms of access and delivery times, requiring them to adapt their operations and implement innovative solutions such as micro-consolidation centres, electric cargo bikes, or alternative delivery methods. The earlier-mentioned developments, like consolidation centres outside city centres and collaboration among logistical stakeholders, also come to mind. Although these challenges seem hard, they also encourage operators to think out of the box and cooperate.

Furthermore, the Netherlands has set ambitious climate goals to address the challenges of climate change. One of the key objectives is to achieve carbon neutrality by 2050, aiming to reduce or even eliminate greenhouse gas emissions from the country's economy. To support this goal, the Netherlands will implement zero-emission zones in 30 to 40 city centres of the bigger municipalities, where only vehicles with zero emissions are allowed to utilise city logistics ("Klimaatakkoord", 2019). These zones promote the use of electric vehicles and other sustainable transportation options, reducing air pollution and creating a cleaner environment. As a result, logistics operators must transition their fleets to zero-emission vehicles to continue operating within these zones. While this transition presents initial challenges regarding vehicle availability and costs, it also offers opportunities for innovation and sustainable practices like collaboration and the consolidation of logistics operations among different companies. By adapting their operations, including reevaluating or collaborating in delivery routes, (jointly) optimizing vehicle utilization, and investing in charging infrastructure. This collaborative approach allows for optimising freight movements, reducing the number of vehicles on the road and minimizing congestion. Besides, this collaborative model encourages a more sustainable and integrated city logistics ecosystem, promoting the sharing of infrastructure, data, and expertise among different stakeholders.

It is estimated that by 2050, 68% of the world population is likely to be living in cities or other urban centres (United Nations Department of Economic and Social Affairs, 2018). These urban areas, where logistics takes place, are characterized by limited space, high population density, diverse economic activity, and competing demands for infrastructure and services. However, urban areas also strive for a healthy living environment that is influenced by various factors such as our living and working conditions, dietary habits, exposure to air pollution, social interactions, and lifestyle choices (Data- en Kennishub Gezond Stedelijk Leven (DKH GSL), n.d.). Efficient utilization of space on a street level is crucial, and determining the appropriate amount of space needed for each function is a complex

[1] Rijkswaterstaat is a Directorate-General of the Ministry of Infrastructure and Water Management of the Netherlands.

challenge to solve. Moreover, in the Netherlands, there currently is a housing task present to construct a total of 1,000,000 homes (Nederlandse Omroep Stichting, 2021), which will implicate the increase of urban developments with its corresponding challenges.

Due to the substantial growth of logistics in urban areas, driven by e-commerce, urbanization, and other factors, managing freight movements has become increasingly challenging. With a rise in urban developments, finding suitable space for each function presents a complex puzzle. Efforts to encourage sustainability, such as implementing zero-emission zones, consolidation centers, and transitioning to electric vehicles, are underway. However, there's an evident lack of understanding regarding the quantity of city logistics operations across these domains. While some traffic counts have been conducted, insufficient information is available regarding the amount of logistics movements occurring in existing and new areas. This makes it impossible for municipalities to answer questions such as: "Will this new urban development lead to traffic congestion due to logistics movements?" "How many parking spaces to reserve for (un)loading logistic vehicles?" "Can road safety be guaranteed in our city centre?" "Can business owners in our city centre receive their supplies in time?" Without this information, municipalities cannot target their spatial planning and supporting (traffic) policies to guarantee a clean, green and pleasant living environment.

Moreover, the spatial planning of logistical facilities, such as terminals or distribution centers, lacks a lot of knowledge and tools. Logistical facilities play a central role in optimizing freight trips and services within urban areas, but it comes with its fair share of challenges. Coordination among various stakeholders, such as transport companies, local authorities, policymakers, and businesses, often proves to be a complex task. Diverse interests and objectives can hinder efficient facility placement and design. Traffic congestion is another issue, as poorly planned logistical centres can contribute to increased congestion on already congested urban road networks. Furthermore, accessibility and safety for both freight traffic and their personnel is a critical concern. Ensuring convenient access to transportation while minimizing negative impacts on local communities remains a precise balancing act in spatial planning for logistical facilities. No tools are currently available to address these challenges and analyse the impact on traffic conditions, congestion, accessibility, emission and safety.

Index numbers for the generation of logistics are a key piece of the puzzle in determining the number of logistics movements generated by a predefined area. For example, the number of freight trips generated by a neighbourhood in a city suburb with an x amount of inhabitants. Such index numbers offer valuable information enabling efficient traffic management and planning. Especially concerning the limited amount of public space, efficient and conclusive decisions can be made concerning developing areas. Logistics facilities, which are a key piece in the city logistics puzzle, do not have any policy regarding spatial planning and growth (Boer, n.d.). By including logistics facilities in the design phase, the dwell time and disturbance of logistics vehicles can be significantly reduced. Both for now and in the future, since spatial plans will determine the space for logistics in cities for years to come ("Logistieke hubs voor emissievrije stedelijke distributie", 2023). Index numbers can be the game changer, revealing the potential and minimum necessity in a developing area regarding spatial use.

1.2. Scientific Relevance

An analysis of the previous work regarding quantifying freight (trip) generation, found in chapter 3, shows that the Netherlands is barely represented in the studies conducted. A fair share is conducted in the United States of America, some across Europe and some in parts of Asia. A large part of the studies have a regression analysis as a method for the model, have vehicle trips as the dependent variable and show those trips on a disaggregate level. Moreover, little models show an extensive analysis per commodity and the highest number of sample size is 135,564 trucks.

This study aims to fill several knowledge gaps using detailed data from multiple sources. The main data source, the BasGoed freight tour data, has a very high level of detail, fully covering the Netherlands and containing over 4 million trips, with information about the weight of the shipment, commodity type and location of origin and destination (chapter 4). Additionally, the BasGoed freight tour data is derived from XML microdata containing over 5 million observations. Finally, the data will be enriched with two other data sources, ultimately testing 26 explaining factors (Figure 3.2).

1.3. Societal Relevance

The societal relevance of this research lies in its potential to clarify and elaborate on the assumptions made for the quantification of city logistics which is found in current area development reports. Examples are found in the area development of Beurskwartier and Merwede of the municipality of Utrecht (Gemeente Utrecht, n.d.-a, Gemeente Utrecht, n.d.-b). Suitable statistics and tools are absent in the field of logistics.

1.4. Research Goal

The goal of this research is to gain more insight into the quantification of logistics by determining the explaining factors for logistics generation in the Netherlands considering urban environments (i.e. the freight trips or freight volumes generated by an urban area, considering its parameters) and provide knowledge for spatial planners to more efficiently implement logistics into urban developments and know its impact.

The first part of the research goal, which is gaining more insight into the quantification of logistics by determining the explaining factors for logistics generation in the Netherlands, is defined as quantifying logistics in terms of index numbers (explaining factors) coming from a Freight Trip Generation (FTG)[2] or Freight Generation (FG)[3] model. These index numbers can indicate the number of freight generated by each Logistic Segment for a given geographical area. For example, the number of trips generated by an area which has an x amount of inhabitants and a y amount of people working in the retail industry in the city centre of Utrecht and which specifically belongs to the retail Logistic Segment (LS).

The second part, which is defined as providing knowledge for spatial planners to more efficiently implement logistics into urban developments and know its impact, has its goal to offer insights into the required space and efficient implementation for logistics and logistical facilities regarding urban development. Since knowledge of the number of logistics is lacking severely, efficiently implementing logistical facilities and their required area, as well as the impact of implementing logistics in urban environments (from implementing hubs or loading docks in residential areas to the impact on traffic), can be evaluated more severely.

1.5. Scientific and Societal Contribution

The scientific contribution of this study is addressed by having a statistically significant model which gives the explaining factors for FTG and FG models. Moreover, the model is based on the most recent available data and the analysis is based on all commodity types instead of one trip count of weight count. Furthermore, this model intends to analyse and identify new explaining factors which have not been found significant in the literature.

Regarding societal contribution, due to the absence of suitable statistics and tools in the field of logistics, there is an urge for a guide that can assist municipalities in identifying and understanding the specific logistics requirements of a given area and acquiring these figures easily. With this guide, presented as index numbers, municipalities can quantify what kind of logistics is generated in a particular area. By

[2]FTG, i.e., the generation of vehicle trips.
[3]FG, i.e., the generation of the cargo that is transported by the vehicle trips.

providing this valuable information, the guide can contribute to educated decision-making and effective planning regarding spatial use within city boundaries as well as outside.

1.6. Reading Guide

After this introduction, chapter 2 will discuss the design of this research, consisting of the research gaps, objectives, questions and methodology. Next, chapter 3 will highlight the current literature regarding Freight Trip Generation (FTG)/Freight Generation (FG) models. Chapter 4 will discuss and evaluate the sources for (city) logistics data in the Netherlands and the data used in this study. Then, chapter 5 will show how the model used in this study is developed and estimated. It will start with an overview of the development, show which variables are taken into account, and discuss the verification and validation of the model. The results of the model are then shown in the same chapter but applied in chapter 6 where the use case is shown and elaborated. This report finishes with a discussion, conclusion and recommendations in chapter 7, chapter 8 and chapter 9, respectively.

2

Research Design

This chapter describes the research design, including the research gaps, objectives, research (sub-)question(s) and research methodology.

2.1. Research Gaps

Although there has been published a lot regarding (city) logistics (Dolati Neghabadi et al., 2019), *much is still unknown about the actual generation (production and attraction) of logistics movements in urban areas*. Current transport models consist of aggregated production- and attraction figures or very local, hand-calculated, estimates for a smaller neighbourhood. This research will focus on freight generation in terms of Logistic Segment and what happens within urban areas, scaling it up towards a universal tool which can be used in the Netherlands to estimate the production and attraction for a specific area. Additionally, *there are also some uncertainties about the impact of planning policies on logistical facilities* (Boer, n.d.). That is why, this research can complement those new insights for policymakers regarding logistical facilities.

2.2. Research Objectives

To accomplish the goal of this research, which is to gain more insight into the quantification of logistics by determining the explaining factors for logistics generation in the Netherlands and provide knowledge for spatial planners to more efficiently implement logistics into urban developments and know its impact, six research objectives are defined which have to be accomplished throughout the research. The objectives are more thoroughly described in section 2.4, but in summary the objectives are defined as follows:

1. To describe current freight traffic generation models
2. To determine the feasibility of a regression model with Dutch (freight) data
3. To validate the regression model
4. To determine the explaining factors of the developed regression model
5. To determine how this model can be applied to an urban development use case
6. To determine which spatial planning scenarios can be considered using the outcome of the model and define the learning lessons from this model on policy implications

2.3. Research Question

This research is conducted in collaboration with Movares, a consultancy and engineering firm, which also is interested in the insights given by this research's goals. With that in mind and given the presented research goal, gaps and objectives, which suggest the importance of providing insight regarding the quantification of logistics, the main research question is defined by:

Main research question

How can the logistics movements of urban environments in the Netherlands be predicted?

This main research question is answered using a total of eight sub-questions which are divided into a scientific and practical part. The scientific part will answer the first goal of this research and the practical part the second. Together they will then answer the main research question. The sub-questions are in turn divided according to the three different phases in this research. An overview of the phases and how the main research question is answered is depicted in Figure 2.1.

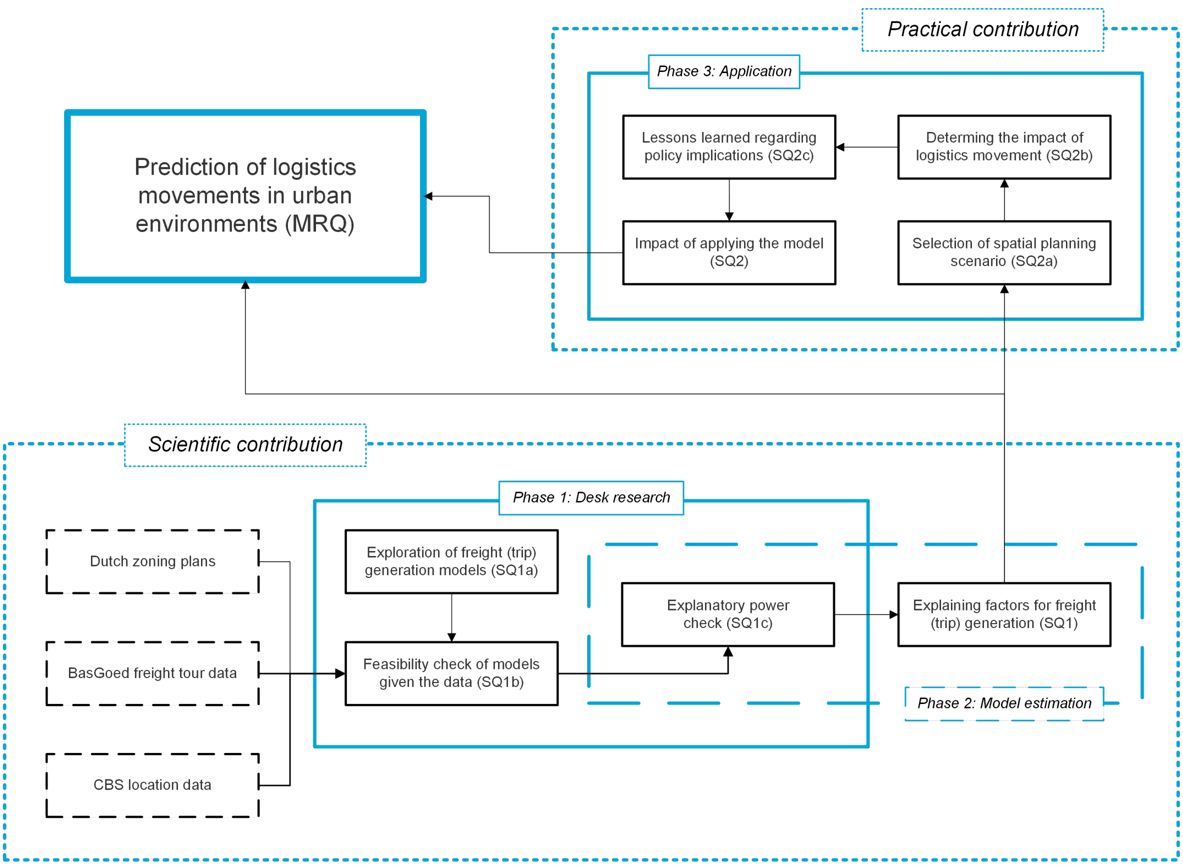


Figure 2.1: Overview of the research methodology

The first part is considered as the scientific part. The goal of this part is to determine the explaining factors from a regression analysis. The overarching question of the scientific part is defined in the following sub-question:

1. What are the explaining factors for freight (trip) generation of logistics in the Netherlands?

This sub-question is answered in phases one and two of the research. In phase one two sub-questions are answered which are created from the urge to get more insights into the current freight traffic models and what can be learned from them. Moreover, the second question defined what models are feasible with the data. This is conducted using desk research. The sub-questions are formulated as follows:

> 1. a What freight (trip) generation models exist, and what can be learned from them?

> 1. b Which models are feasible, given the limited available data?

In the second phase, where the model is developed and estimated, the question describes how strong the explanatory power of the regression model will be using the learned lessons from the existing FTG and FG models. This phase ends with the answering of the scientific contributed sub-question (1) and the following sub-question:

> 1. c How strong is the explanatory power of the model?

Finally, the third research phase is addressed with four more questions. These questions describe what the impact will be when the developed model is applied to an urban development use case. First, the spatial planning scenario is defined, the impact of the increase in trips is determined and finally, the lessons learned from the policy implications are considered. This phase is formulated using one overarching, practically oriented, sub-question which is divided into three sub-sub-questions:

> 2. What is the impact of applying the model to a use case?

> 2. a Which spatial planning scenarios can be considered using this model?

> 2. b What is the impact in terms of logistics movements?

> 2. c What policy implications can be learned from this use case?

2.4. Research Methodology

Now that the main research question and sub-questions are defined, this section explains the methodology of the three phases leading towards answering the main research question. Moreover, this section shows how the sub-questions are answered. An overview of the sub-questions and their intended method is depicted in Table 2.1.

No.	Sub-question	Method
1	What are the explaining factors for freight (trip) generation of logistics in the Netherlands?	Desk research & regression analysis
1a	What freight (trip) generation models exist, and what can be learned from them?	Desk research
1b	Which models are feasible, given the limited available data?	Desk research
1c	How strong is the explanatory power of the model?	Desk research & regression analysis
2	What is the impact of applying the model to a use case?	Use case
2a	Which spatial planning scenarios can be considered using this model?	Desk research
2b	What is the impact in terms of logistics movements?	Use case
2c	What policy implications can be learned from this use case?	Use case

Table 2.1: Overview of the different methods per sub-question.

2.4.1. Phase 1: Desk Research

Before the research phases, a desk research phase is conducted regarding current freight (trip) generation models and what explaining factors they entail. Furthermore, a brief overview of the data sources in the Netherlands is depicted. This can be found in chapter 3 and chapter 4, respectively. Moreover, the data used in this research is analysed. This is also shown in chapter 4. The steps defined in this phase are the following:

1. Define search query
2. Filter relevant papers
3. Define methods used in previous models
4. Define explaining factors from previous models
5. Take stock of available data sources
6. Analyse and report BasGoed freight tour data
7. Analyse and report CBS[1] geographical socio-demographic data
8. Analyse and report Dutch zoning plans

2.4.2. Phase 2: Freight (Trip) Generation Model

In phase 2, an empirical freight (trip) generation model will be developed that can specifically be used to predict logistics. To develop this model, the initial plan was to use CBS-microdata[2] containing freight traffic movements. However, due to circumstances regarding the availability of this data, obtaining it took more time than expected. Therefore, BasGoed[3] freight tour data is used instead to develop the model. The data of the BasGoed model is synthetic but based on the CBS-microdata, also see chapter 4 for a more detailed description of this data.

Next to the freight tour data, publically available CBS location data is used to make a differentiation between different zones of the freight tour data. Next to that, relevant Dutch zoning plans are incorporated into the model. In combination with the freight movements from the BasGoed freight tour data, the location and zoning data are used to investigate, on a more disaggregated scale, what different areas attract and produce.

[1] Dutch Central Agency for Statistics

[2] Microdata is linkable data at person, company and address levels that allow Dutch universities, scientific organisations, planning agencies and research bodies in a number of other EU countries to conduct their own statistical research under strict conditions (Centraal Bureau voor de Statistiek, n.d.).

[3] BasGoed is a strategic freight transport model used to produce forecasts for road, rail and inland waterways. It aims to identify the effects of economic developments and policy measures on freight transport ("BasGoed", n.d.).

With the estimated model, logistics in urban environments can be predicted based on a specific combination of parameters (i.e. amount of surface area in offices, residential or retail). Moreover, the model is put to the test to achieve its best predictive performance. To achieve this, the following sub-steps are defined:

1. Prepare BasGoed freight tour data
2. Prepare CBS location data
3. Prepare Dutch zoning data
4. Combine freight traffic-based, location and zoning data to construct a working dataset
5. Analyse which parameters achieve the model's best performance
6. Estimate model parameters
7. Validating and determining the explanatory power of the model according to the literature

2.4.3. Phase 3: Application

Phase 3 will be an application of the developed model. The goal of this research is the explanation of the logistic movements of an area, given the area's parameters. It will indicate the amount of trips an area generates. Therefore, the application is conducted to show the practical use of the model, to what extent the model is feasible to use in spatial planning scenarios and what results can be expected. Moreover, the application shows how this model can benefit in evaluating the impact of spatial scenarios. In phase 3, a use case will be determined, described and shown how the model is applied. The following sub-steps are defined in this phase:

1. Define what spacial scenarios are feasible with the model
2. Define the use case and its parameters
3. Simulate the impact on freight trip movements
4. Analysis of results on indicators and policy implications

3 Literature Review

This chapter describes the desk research conducted in phase one of this research. The desk research focuses on freight (trip) generation models of previous work, this part can be found in section 3.1. The method and explaining factors from these models are distilled from the studies thus the pros and cons, considering this research' dataset, can be summed. Finally, this chapter concludes by answering sub-question 1.

Starting from the study of Pani et al. presented by the committee, other literature of freight trip generation models used in this desk research is found using the following search query on Scopus:

TITLE-ABS-KEY (("freight trip generation" OR "freight generation") AND "regression analysis") AND PUBYEAR > 2017 AND PUBYEAR < 2025

The choice of this query is based on the study of Pani et al., given its overview table showed a method preference for regression analysis and an overview of the research conducted before the year 2018. The query resulted in a total of 13 papers. These papers are then scanned by title and abstract which resulted in a total of 10 papers used in this desk research, next to the studies used by Pani et al. Other additional papers are found using specific queries on Scopus and Google Scholar, are found using Connected Papers where a map is shown with links between a specific paper or are presented by the committee. A final source of literature is *Handbook on City Logistics and Urban Freight*, which is provided by the TU Delft Library.

3.1. Freight (Trip) Generation Models

Freight (trip) generation models are essential tools for predicting and managing the movement of goods within transportation networks. Unlike traditional models for passenger travel, these models focus on characteristics unique to freight, such as industry/goods type, land use, employee number and spatial distribution. By using data, these models help optimize logistics, guide infrastructure investments, and inform decision-making for businesses and policymakers. With the increasing demographic and economic growth, understanding and forecasting freight movements have become critical for efficient and sustainable transportation systems and spatial planning. This brief overview sets the stage for the following section. Here are the FG/FTG models from previous work described and their explaining factors presented. At first, this section describes the methods used in the literature. Next, the explaining factors found in the literature are described in depth. Table 3.1 gives an overview of the previous models based on the study area, the method used, results, whether it is on an aggregate or disaggregate level and finally the dependent[1] and independent[2] variables. Concerning the data collection and size of the models, a more recent study by Rodoshi et al. highlighted the data collection techniques used in FG/FTG models. The conclusion of this study was that truck trip-based studies have large sample sizes that range from 5,276 (Holguín-Veras & Patil, 2007) to 135,564 (Gonzalez-Calderon et al., 2021).

Study	Study area/country	Method; results; A/D[3]	Dependent variable	Independent variable
Iding et al., 2002	Germany	Regression; $R^2 \leq 0.88$; D	Vehicle trips	No of employees & area firm, Site area, Employment
Beagan et al., 2007	Wisconsin, Florida, Indiana, Tennessee	Regression; $R^2 \leq 0.98$; A	Weight, Vehicle trips	Employment type, No of households, Population
Institute of Transportation Engineers, 2008	USA	Trip rate; D	Vehicle trips	Land use
Bastida and Holguín-Veras, 2009	New York, Manhatten, Brooklyn	Regression, MCA[4]; D	Vehicle trips	Commodity type, Industry sector, Employment
Giuliano et al., 2010	Los Angeles	Input - output; A	Weight	Inter and intra-regional commodity flow, Employment
Holguín-Veras et al., 2011	New York, USA	Regression; $R^2 \leq 0.93$; D	Vehicle trips	Employment, Commodity type
Campbell et al., 2012	New York, USA	MCA; D	Vehicle trips	Employment, Commodity type
Lawson et al., 2012	New York, USA	Regression; $R^2 \leq 0.49$; D	Vehicle trips	Employment, Land Use, Commodity type
Alho and Silva, 2014	Lisbon, Portugal	Regression; $R^2 \leq 0.67$; D	Vehicle trips	Employment, Frontage width, Warehouse area, Sales area, Commodity type
Asuncion, 2014	Canterbury, New Zealand	Regression; $R^2 \leq 0.92$; D	Vehicle trips	Employment, Retail trading area, Storage and parking spaces, Product variation number
Holguín-Veras et al., 2014	New York, USA	Regression; $R^2 \leq 0.60$; D	Weight, Vehicle trips	Employment, Commodity type
Jaller et al., 2014	New York, New Jersey	Regression; $R^2 \leq 0.91$; D	Vehicle trips	Employment & area, Employment, Area, Commodity type
Ha and Combes, 2016	France	Regression; $R^2 \leq 0.68$; D	Weight	Employment, No of clients, No of carriers, Share of transport cost in product value
Mommens et al., 2017	Belgium	Regression; D	Vehicle trips	Employment, Area per activity, Population
Sánchez-Díaz, 2017	Gothenburg, Sweden	Regression; $R^2 \leq 0.75$; D	Weight, Vehicle trips	Employment, Area, Commercial sector
Pani et al., 2018	Kerala, India	Regression; $R^2 \leq 0.64$	Weight	No employees & gross floor area, No of years in business
Gonzalez-Feliu et al., 2020	Lyon, France	Regression; $R^2 \leq 0.88$; D	Vehicle trips	Population, Employment, Distance between zones
Sanchez-Diaz, 2020	Stockholm, Sweden	Regression; A	Vehicle trips	No of employees, Area
Cheah et al., 2021	Singapore	Regression; D	Vehicle trips	Employment, Establishment type
Oliveira et al., 2022	Brazil	Regression	Vehicle trips	No of employees, Area
Zhang and Yao, 2022	Shanghai	Regression; $R^2 \leq 0.55$; D	Vehicle trips	Population, Employment, Population density, Land use
Dhulipala and Patil, 2023	India	Regression; D	Vehicle trips	Employment

Table 3.1: Overview of past FG/FTG studies (adapted from Pani et al., 2018).

A general analysis of this overview shows that the Netherlands is barely represented in the studies conducted. A fair share is conducted in the United States of America, some across Europe and some in parts of Asia. A large part of the studies have a regression analysis as a method for the model, have vehicle trips as the dependent variable (which results in FTG models) and show those trips on a disaggregate level. Moreover, the amount of independent variables taken into consideration in these studies is quite considerate and ranges from share of transport costs to commodity type, and from employment to storage and parking spaces.

3.1.1. Methods

From previous work, four methods for FG/FTG models can be determined: linear regression my means of Ordinary Least Squares (OLS), Multiple Classification Analysis (MCA), trip rate and input-output. The regression method, being mostly based on a statistically robust technique, is seen most in previous work (Table 3.1). It is easy to use and can uncover patterns, gain insights into data and make predictions for

[1]The dependent variable is the outcome variable that is being predicted or explained by the independent variables in a regression model. It represents the variable that is being observed and is influenced by the independent variables in the model. Here the dependent variables can be the amount of freight trips or the amount of freight weight.
[2]The independent variable is the variable that is presumed to have an effect on the dependent variable.
[3]Aggregate/Disaggregate.
[4]Multiple Classification Analysis.

the future. Although by use of a MCA, parameters can freely change across intervals of the independent variable (Lawson et al., 2012), comparative analysis revealed that MCA performed slightly better than regression models for estimating Freight Trip Attraction (FTA), although there was only a very small difference to the overall error (Veras et al., 2012). Trip rate is a method and can be derived from trip- and tour data, trips can then be divided by a factor (i.e. employees) which shows the rates per employee. An example of such rates can be found in Figure 3.1. Finally, input-output is based on economic interdependencies between different zones. Trips or volumes are then predicted based on the economic value, often for aggregate areas.

Table 3.3 Daily Truck-Trip Rates Used in Factoring Truck Trips

SIC	Description	Trips/Employee
1-9	Agriculture, Forestry, and Fishing	0.500
10-14	Mining	0.500
15-19	Construction	0.500
20-39	Manufacturing, Total	0.322
40-49	Transportation, Communication, and Public Utilities	0.322
42	Trucking and Warehousing	0.700
50-51	Wholesale Trade	0.170
52-59	Retail Trade	0.087
60-67	Finance, Insurance, and Real Estate, Total	0.027
70-89	Services	0.027
80	Health Services (Including State and Local Government, Hospitals)	0.030
N/A	Government	0.027

Source: Minnesota DOT.

Figure 3.1: Trip rates per employee (Beagan et al., 2007).

3.1.2. Explaining factors

Explaining factors are the independent variables which affect the dependent variables (vehicle trips or volume/weight) in a model. From previous work, it can be derived that employment (17 times) and commodity type (7 times) are in most cases the explaining factors for FG/FTG models. Area, number of employees and population are also derived to be significant in some models. Pani et al. suggested in his study to use of multiple-variable models for quantifying freight activities from establishments subjected to data availability. Single-variable FG models may be used when the data availability is limited to one variable.

For employment, Jaller et al. results showed that the strongest correlation between employment and area is seen in commercial areas, office and retail in particular. The employment-based variables perform better than the area-based in these situations. Moreover, Pani et al. explained that employment is invariably considered the most preferred causal variable in FG/FTG models. It is often considered along with other variables such as business area, commodity type, industry segment, gross floor area, building area or parking area, or both while modelling the freight models. Also what is interesting regarding employment as an explaining factor, is that Sánchez-Díaz et al. described that FTA is better modelled using non-linear models for all industry sectors. Specifically, the FTA of business establishments is concave with employment, flattening as employment increases. That matched the result of Oliveira et al., where the results of the study showed that the number of employees has a more significant influence in small cities and a lower influence in medium-sized municipalities. Finally, Sahu and Pani stated that the predictive ability of business size variables suggests that employment is more suitable for representing FP, while area explains FA better.

3.1.3. Freight (Trip) Generation

The literature also points out a few arguments regarding whether a FTG or FG model should be used. Holguín-Veras et al. describes that the correlation of establishment size variables (e.g., employment) with FG is relatively stronger than FTG due to its direct physical correlation with the scale of commodity production and attraction. FTG, on the other hand, is influenced by shipment size and logistic decisions. Veras et al. depicts FG as an expression of economic activity performed at a business establishment. Here the input materials are processed and transformed generating an output that, in most cases, is transported elsewhere for further processing, storage, distribution, or consumption. FTG, on the other hand, is the result of the logistic decisions concerning how best to transport the FG in terms of shipment size, frequency of deliveries, and the vehicle or mode used. The ability of the distributor to change shipment size to minimize total logistic costs, as it allows carriers to increase the cargo transported (the FG) without proportionally increasing the corresponding FTG is of great importance. As a result, FTG cannot be universally assumed to be proportional to business size because large establishments could receive larger amounts of cargo without the accompanying increases in FTG. This has big implications for FTG modelling, as normally it is implicitly assumed that FTG and business size variables (e.g., square footage, employment) are in proportion. The case studies in the work of Veras et al. confirm that proportionality between FTG and business size only happens in a minority of industry segments.

3.2. Conclusion

In conclusion, the desk research shows various methods for modelling Freight Generation and Freight Trip Generation models, with a main focus on regression methods such as Ordinary Least Squares. While regression models are statistically robust and widely used for their convenience and predictive capabilities, Multiple Classification Analysis is found to perform slightly better in estimating Freight Trip Attraction. The analysis identifies employment and commodity type as key explaining factors in FG/FTG models, with employment consistently emerging as the most preferred causal variable. Remarkably, the correlation between employment and area varies across commercial sectors, influencing model performance. The literature suggests the use of multiple-variable models when data availability allows, with single-variable models employed under limited data conditions.

Furthermore, discussions are present regarding the choice between FG and FTG models. The correlation of establishment size variables, such as employment, is described as stronger with FG due to its direct physical connection with commodity production and attraction. In contrast, FTG is influenced by shipment size and logistics decisions. The assumption of the relation between FTG and business size variables is challenged, as large establishments may receive larger cargo without a proportional increase in freight trips. This challenges the conventional modelling assumption and emphasizes the importance of considering logistic decisions in FTG modelling. Overall, the study highlights the complexity of modelling both models, highlighting the need for nuanced approaches that provide diverse factors influencing freight activities.

Sub-question 1a: What freight (trip) generation models exist, and what can be learned from them?

In the last 20 years, a lot of FG/FTG have been developed. Many models were tailor-made to serve a specific purpose (i.e. production and attraction by restaurants (Zhou et al., 2018)), or developed for a specific city or urban area. Almost all of the models are based on a linear regression method, applied to dependent and independent variables of a specific dataset (many times retrieved by surveys). In logistics, regression would also be the best way to approach production and attraction. With the request for quantification of logistics in urban areas being the goal of this research, regression is the most robust and straightforward way to go over all the variables of the BasGoed freight tour data in combination with the area-specific parameters of the CBS location data and Dutch zonal plan data. Explaining factors like employment, commodity type (or logistic segment in this model), area (size and type) and population would be able to be derived in this to-be-developed model. Moreover, this model takes urban density also into account and places the model results on a wider geographical perspective. A conceptual model, showing the explaining factors considered in this study, is shown in Figure 3.2. The figure depicts the individual factors and their expected regression paths. An extensive description of the factors is found in chapter 4.

Given the previous information, the choice for this study's model method is made to be a regression analysis using a Ordinary Least Squares method. OLS has shown to be a robust method in this sector. Given the data specifics, which also will be elaborated in chapter 4, the models will consist of a Freight Trip Generation and Freight Generation prediction.

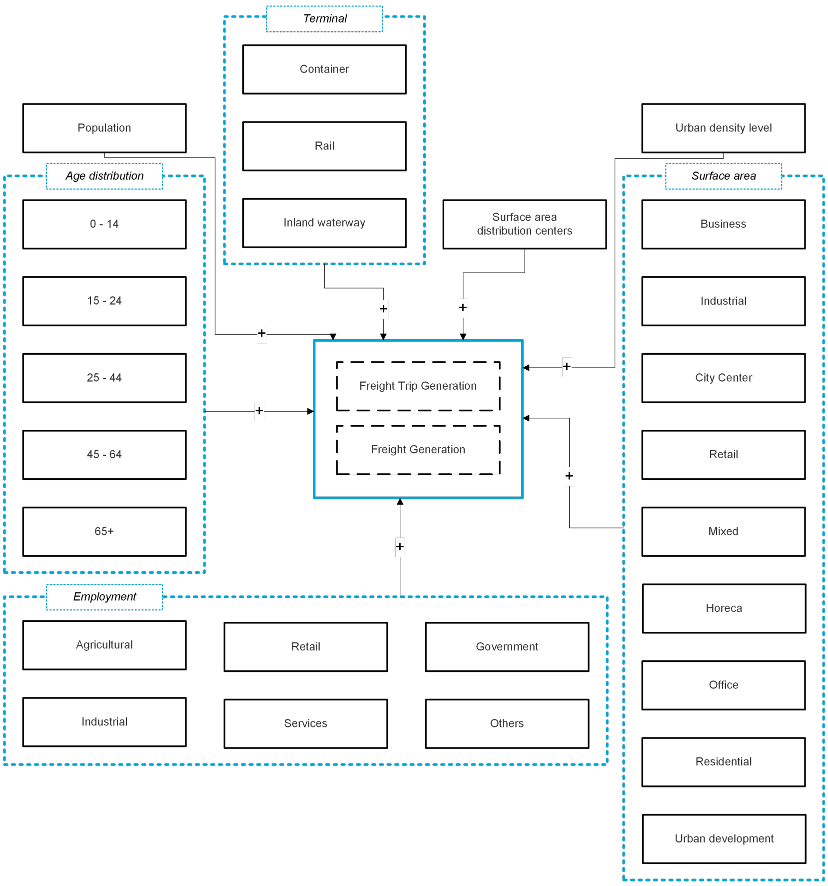


Figure 3.2: Conceptual model

4

Data Description

This chapter describes the data sources available in the Netherlands and puts them in perspective to the data used in this research. This is shown in section 4.1 and section 4.2, respectively. Next, the data sources used in this research and how they are collected are described in the same section: section 4.2. This includes the BasGoed freight tour data, the CBS location data and the zoning plan data. All three sections are divided into the available data characteristics, sample representativeness, data provision and data limitation and strengths. Next, section 4.3 shows what available variables are inside the provided data and an exploratory data analysis is conducted. Finally, a conclusion is given regarding the usability of the two data sources in this research. An overview of how the data is defined and used in this research is shown in Figure 4.1.

4.1. Logistics data in the Netherlands

This section gives a brief description of the key available data sources for logistics in the Netherlands. A short search on the internet gave the following options and their short descriptions:

1. CBS
   The Central Bureau of Statistics in the Netherlands provides a wide range of data related to demographics, employment, transportation, and economic activities, which can be useful for understanding logistics trends and patterns. In particular, CBS provides these sources:

   - CBS-microdata
     Microdata is linkable data at person, company and address levels that allow Dutch universities, scientific organisations, planning agencies and research bodies in several other EU countries to conduct their statistical research under strict conditions (Centraal Bureau voor de Statistiek, n.d.).

     + Coverage across the whole country
     + Trip coordinates are available
     – Data is very hard to get access to

   - CBS StatLine
     StatLine is the online database of the CBS. The information is conveniently classified by theme and freely available to everyone.

     + Coverage across the whole country
     + Free to use

    - Only available on aggregated (province/country) level

- Basisbestanden goederenwegvervoer
  This data source allows describing road traffic and transport within the Netherlands - as well as to and from the Netherlands - by both Dutch and foreign vehicles. The source consists of information on vehicle, trip, and shipment data and is compiled for Rijkswaterstaat's traffic and transport models.
    + Coverage across the whole country
    + Usable dataset
    – Trips available at so-called VAM-zones (550 across the Netherlands and Europe)

2. Rijkswaterstaat
   The Dutch Ministry of Infrastructure and Water Management's agency responsible for managing the country's roadways and waterways provides data on transportation infrastructure, traffic flows, and road conditions, which can be valuable for logistics planning and analysis. In particular, Rijkswaterstaat also provides the BasGoed freight model.
   + Coverage across the whole country
   + Usable dataset on a disaggregated level
   – Not publicly available
3. Transport Operators and Companies
   Logistics and transport companies operating in the Netherlands often collect data on their operations, including fleet management, route optimization, and delivery performance, which can be accessed through partnerships or data-sharing agreements.
   + Rich and real-time data insights
   – Limited accessibility and potential privacy concerns
   – Limited data disaggregation
   – Potential bias towards company operations
4. Research Institutes (TNO) and Universities
   Academic institutions and research organizations in the Netherlands conduct studies and surveys related to logistics and freight, generating valuable datasets for research purposes.
   + Objective and specialized analysis
   + Publicly available reports
   – Time-consuming data collection
5. Industry Associations
   Organizations such as Transport and Logistics Netherlands (TLN) and Topsector Logistiek (TL) may provide industry-specific data, reports, and insights on logistics trends, regulations, and best practices.
   + Industry-specific insights
   – Limited data disaggregation and potential bias towards member interests
   – Potentially restricted access

4.2. Used Data

This section describes the data used in this research. First, the selection of data is addressed, considering the knowledge of the previous section. Next, the three data sources that are used in this study are described. One of the datasets is known from the overview of the logistics data of the Netherlands, the other two are found in light of this research and its search for (new) explaining factors. Data enrichment, which is also conducted in this research, is very common in strategic freight models. Data availability is a known issue. Some data is omitted due to the commercial sensitivity, or including information on senders and receivers of the shipments is often very costly (Mohammed et al., 2023), or even some data lacks geographical areas and therefore a combination of multiple geographical datasets is necessary (Grebe et al., 2016). In this study, the three datasets are the BasGoed freight tour data (subsection 4.2.2), CBS location data (subsection 4.2.3) and Dutch zoning plan data (subsection 4.2.4). An overview of the data sources can be found in Figure 4.1 and in this section, key aspects of each dataset are described.

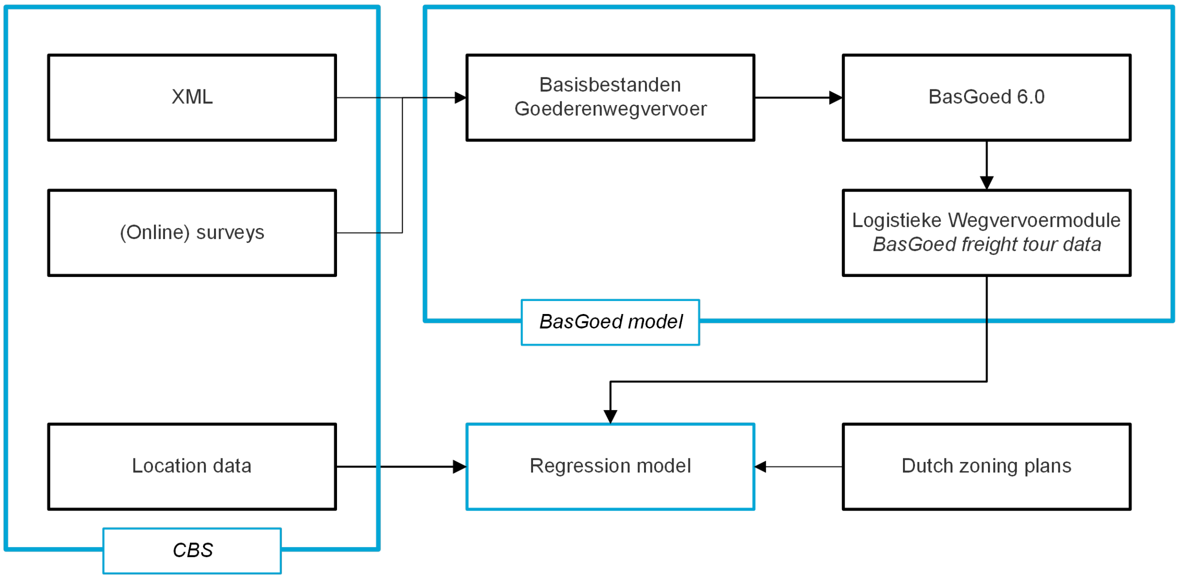


Figure 4.1: Overview of the data sources used in this research

4.2.1. Data Selection

To conduct a regression analysis for Dutch freight, data needs to be at hand or gathered. The question arises of which options there are to collect data for a model. The limited options roughly are:

1. Local/National (online) surveys - This has been seen many times in literature.
2. Statistics (CBS)
3. Simulated data using current models (BasGoed)

Next to the data availability, the data needs to meet a list of requirements:

- A nationwide coverage
- Even distribution across provinces/municipalities/rural areas
- Level of detail (disaggregate)
- Easy to maintain/update
- Fit within the scope of research

Within these requirements, especially within the time and funding of this relatively small research, simulated data - which has nationwide coverage, is evenly distributed, is disaggregate and easy to update, is an easy pick. BasGoed freight tour data fully meets these requirements. The next section will describe the dataset and what it entails.

4.2.2. BasGoed Freight Tour Data

The first dataset that was made available is the BasGoed freight tour data. The basic model of freight transport (Dutch: Basismodel Goederenvervoer, BasGoed) is a freight transport model developed by Rijkswaterstaat, the executive agency of the Ministry of Infrastructure and Water Management, which models future freight flows for target years of 2040 and 2050 with different (economic) scenarios ("Handleiding Webviewer BasGoed", 2022). The current BasGoed model, BasGoed 5.1, is in use as we speak. However, BasGoed 6.0 is in development and this dataset corresponds to the newest version of BasGoed. In that newest version, a freight logistics module is added in which logistics choice behaviour is taken into account more realistically and individual shipments are modelled which are consolidated into round trips. With these additions, the model is suitable for many more road freight transport policy questions.

Although BasGoed can make predictions for future scenarios, the available dataset is synthesized from the base model of 2018. This means that there has been data collection for 2018, which was raised to make it fit the whole area of the Netherlands and was finally made into the BasGoed model. From that model, synthetic data has been fabricated which is used in this research. How this is done is described in section 4.3. Furthermore, since the data is divided into different zones, the characteristics per zone (in 2018) have been provided as well.

Data Characteristics

The BasGoed freight tour data is gathered through two methods: surveys directly from freight carriers, shippers, or logistics companies and XML data (Extensible Markup Language) collected directly from freight transport systems placed in trucks.

Sample Representativeness

As the BasGoed freight tour data is a synthetic dataset which is extracted from a model, the sample representativeness therefore is assumed to be secured. That is: the simulated tours in this dataset are representative of all freight tours in the study area. A 100% sample of the synthetic tours is used. The synthetic tours are simulated by using CBS XML microdata from a road transport survey conducted in the Netherlands. The entire dataset contains over 5 million observations, with information about the vehicle type in which the shipment is transported and its weight, the commodity type (Logistic Segment and NST/R goods type) the location of origin and destination, and the trip characteristics.

Data Provision

The data is provided by Rijkswaterstaat.

Data Limitations and Strengths

BasGoed freight tour data offers detailed insights into freight movements, including origin, destination, routes, and commodities, providing an in-depth analysis of logistics patterns. Its disaggregate level allows for analysis at various spatial scales, from individual trips to broader regional or national patterns. However, the dataset only includes larger freight trucks. Moreover, the dataset is based on 2018 data and is today already 6 years old.

4.2.3. CBS Location Data

CBS publishes statistics by squares and postcodes. These include data on inhabitants, households, social security, income, housing, energy and proximity to facilities. The area classifications of the dataset are squares of 100 metres by 100 metres or bigger. In this study, the squares of 100 metres by 100 metres of 2018 are used. This is to match the dataset of BasGoed which also has its base model in 2018.

Data Characteristics

The data in this publication are derived from the Personal Records Database, the Basic Registration Addresses and Buildings, the WOZ register, the Basic Geographical Register and the Integral Income and Wealth Statistics of the Netherlands. The source for coordinates and postcodes of addresses is the Basic Registration of Addresses and Buildings. Missing postcodes in the Basic Registration Addresses and Buildings are supplemented as far as possible (Statistiek, 2021).

Sample Representativeness

As the CBS, which is short for the Central Bureau of Statistics in the Netherlands, location data comes from an organisation which has statistics as its top priority, the sample representativeness is assumed to be secured.

Data Provision

The data is collected using the website of the CBS (in Dutch).

Data Limitations and Strengths

The CBS location data comes with some limitations, these are described in the documentation provided with the data. Some relevant limitations are described here.

Due to the size of the file to be published for the Netherlands, only those squares in which at least 5 inhabitants or 5 houses are located will be published and these will also be published as an unconcealed, positive number. This means that 100 metres by 100 metres squares located above areas with no data or squares located above areas for which all data are concealed will not be published. Also, if a square is visible but for a specific characteristic the $N < 5$, the characteristic will also not be shown with a positive number but show -99997 instead.

Regarding population, the CBS population numbers include only persons registered in the population register of a Dutch municipality. In principle, everyone living indefinitely in the Netherlands is included in the population register of the municipality of residence. Persons for whom no fixed place of residence can be indicated are included in the population register of the municipality of The Hague. Not included are persons residing illegally in the Netherlands and persons to whom exceptional rules apply, for instance, diplomats and NATO military personnel.

4.2.4. Dutch Zoning Plan Data

Zoning plans are government regulations that designate specific land areas for different uses like residential, retail, or industrial purposes. These plans outline rules for building heights, densities, and other development aspects to guide urban growth and ensure harmony between different land uses. The Dutch zoning plans are a publicly available dataset and consist of nationwide designated land uses.

Data Characteristics

Zoning plans are legally binding documents and are provided to the data model IMRO 2008 and IMRO 2012 (IMRO is short for the Spatial Planning Information Model and is used to exchange data between national, provincial and municipal levels).

Sample Representativeness

As the zoning plans need to cover the whole country, it is assumed that the dataset secures the sample representativeness of the Netherlands. However, the zoning plans used for this research date the latest 31-12-2023, and while zoning plans reflect long-term planning, there might be some discrepancies.

Data Provision

The data is provided by Esri Nederland, Kadaster and ruimtelijkeplannen.nl and is collected using the GIS (Geographic Information System) program QGIS.

Data Limitations and Strengths

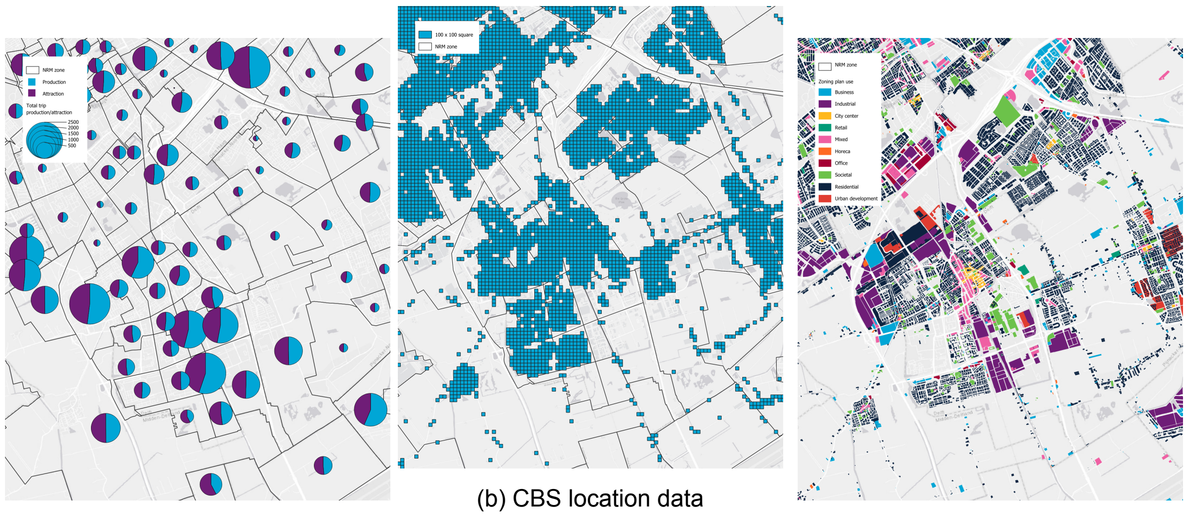
For this type of research, zoning plans normally are a good tool to use. They reflect long-term planning goals and intentions for land use, guiding development decisions over time. As the first dataset in this research dates from 2018, zonal plans would suit this research perfectly. However, although that coverage is supposed to be across the whole country, the amount of different land uses is limited. This would not be suitable for a more detailed analysis.

4.3. Data Collection

As briefly discussed in the previous section, the BasGoed freight tour data, CBS location data and the Dutch zoning plan data are enriched with a lot of information. This section will elaborate on the data collected following the categories of interest.

4.3.1. Available Data and Exploration

In this subsection, all three data sources will be thoroughly described and the rough available data will be presented. First, the BasGoed freight tour data will be described, consisting of the freight tour and zonal data. Next, the CBS location data is described and finally, the Dutch zoning plan data is elaborated. An impression of how the data is geographically visualised is shown in Figure 4.2. Here the NRM zones are shown and for each data source its distinguished data type is presented: trip production/attraction for the BasGoed freight tour data (Figure 4.2a, in trips/workweek), the 100x100 squares for the CBS location data (Figure 4.2b) and the zonal plans for the Dutch zoning plan data (Figure 4.2c).



(a) BasGoed freight tour data

(b) CBS location data

(c) Dutch zoning plan data

Figure 4.2: Impression of the three different data sources

BasGoed Freight Tour Data

First, the BasGoed freight tour data is described. As shown in Figure 4.1, the freight tour data consists of synthetic data required from the BasGoed freight tour module. The data consists of a table which is 4,018,280 rows long and 31 columns wide representing a total of 4,018,280 trips in 2018. Each trip has 31 columns describing the information about each trip. This can be the ID of the carrier, tour or trip, the origin of the trip and its destination, specified in different zoning notations. Furthermore, the trip information consists of the vehicle type used and its combustion type, the DC ID involved (if applicable), information about its cargo like the commodity type, logistic segment, number of shipments, weight, whether it is containerised and finally, the data specifies information about its trip like CO^2 emissions, distances travelled and departure or arrival time.

From this trip data found in the BasGoed freight tour data, the question arises as to how these trips can be explained. In other words, which explaining variables generate these trips? The dependent variables are therefore distilled from this list of trips. As the goal of the research is described as predicting (and explaining) the logistic movements, the trips per logistic segment are chosen. Moreover, as in the literature models are estimated for the generation of volumes and this dataset gives the opportunity to also analyse this, both trips as volumes of logistic segments are estimated. Finally, to add more reference material to this study and compare it to freight (trip) generation models internationally, the European encoding of commodities, Nomenclature uniforme des marchandises pour les Statistiques de Transport, Revisée (NST/R), is added as a dependent variable in trips and volume to estimate. An overview of the distribution of both the amount of trips and the weight of LS and NST/R is shown in Table 4.1 and Table 4.2, respectively.

Label	Description	% trips	% weight
LS 0	Food (general cargo)	12.94	14.56
LS 1	Miscellaneous (general cargo)	45.29	45.99
LS 2	Conditioned transport	11.31	11.88
LS 3	Facility logistics	5.02	3.70
LS 4	Construction logistics	9.51	10.86
LS 5	Waste	8.36	5.88
LS 6	Parcel (consolidated flows between sorting centres)	1.02	0.68
LS 7	Hazardous materials	6.55	6.45
-	**Total**	100.00	100.00

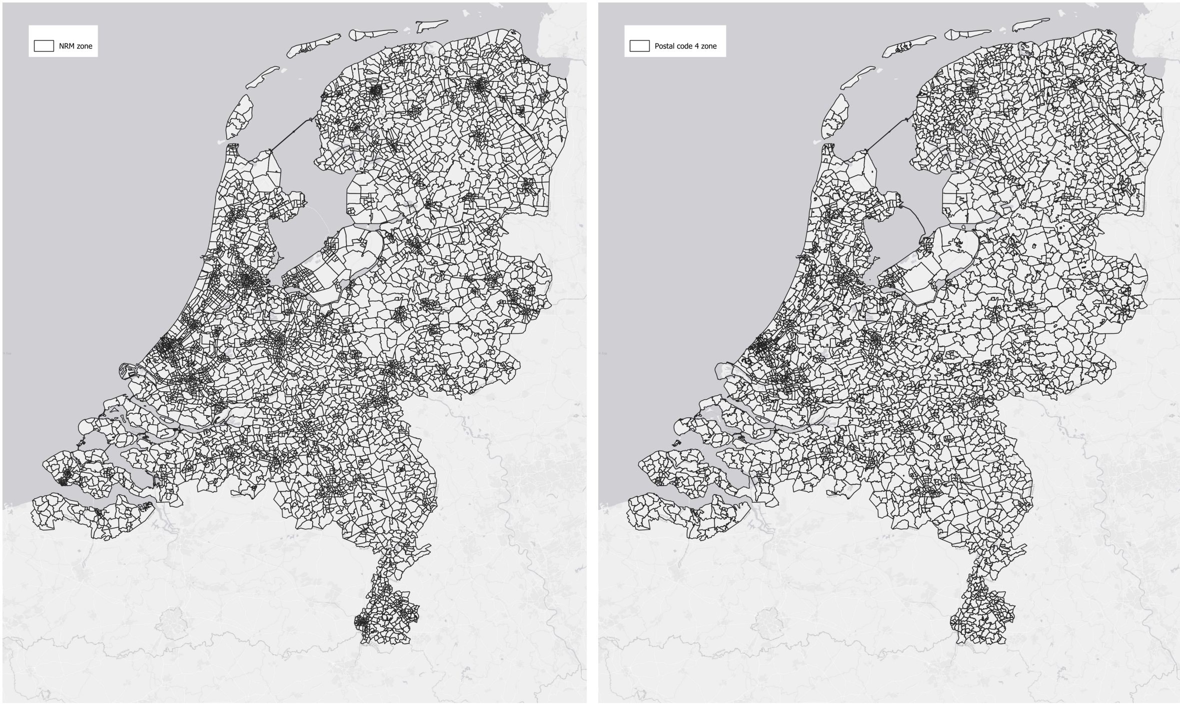
Table 4.1: Distribution of LS in the raw BasGoed freight tour data

Label	Description	% trips	% weight
0	Agricultural products and live animals	6.60	6.34
1	Foodstuffs and animal fodder	10.60	11.97
2	Solid mineral fuels	1.09	1.72
3	Petroleum products	0.38	0.50
4	Ores and metal waste	0.89	0.55
5	Metal products	2.23	2.02
6	Crude and manufactured minerals, building materials	3.52	5.12
7	Fertilizers	0.20	0.27
8	Chemicals	7.74	6.72
9	Machinery, transport equipment, manufactured articles and miscellaneous articles	65.16	64.15
xx	Arms and ammunition, military	0.00	0.00
-1	Empty trips	1.59	0.65
-	**Total**	100.00	100.00

Table 4.2: Distribution of NST/R in the raw BasGoed freight tour data

The trips of the freight tour data are, amongst others, based on the origin and destination labelled by NRM (Nieuw Regionaal Model) zoning. This is the most disaggregated level in the dataset and is described by 6930 zones. These zones are merged from the four NRM models of the Netherlands: NRM-North, NRM-East, NRM-South, NRM-West (zones $1 - 6441$), and a few zones located abroad (zones > 6441). The NRM zones each have their specifications (23 in total), like different ways of zoning labels, coordinates of the zone, area size and the province or municipality it is located in. To visualize it, the NRM zones roughly correspond to the postal code 4 zones (4071 zones). An overview of the NRM zones in the Netherlands and a comparison between the postal code 4 zones is shown in

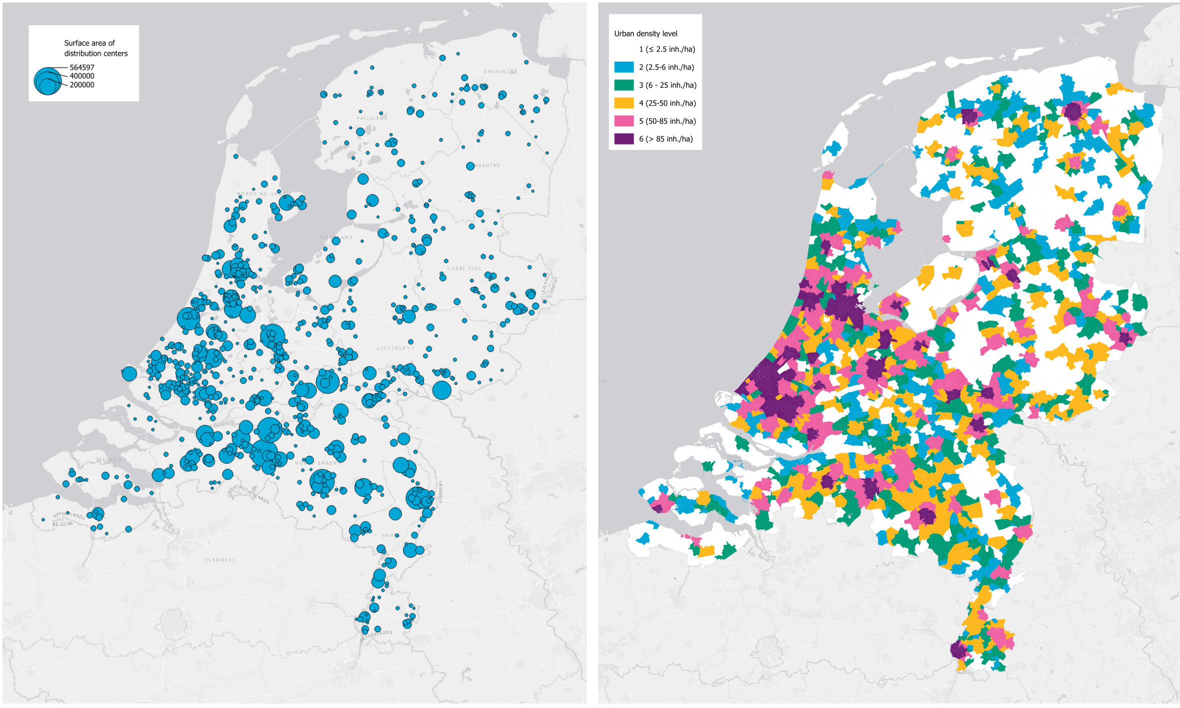
Figure 4.3.



(a) NRM zones (b) Postal code 4 zones

Figure 4.3: Comparison of NRM and postal code 4 zones

Together with the freight tour data, the NRM zones each have their specifications corresponding to the 2018 data, like population size, employment, urban density level and the surface area of a distribution center located in the zone. A total of 20 items. A lot of these items are considered to be explaining factors for freight (trip) production. These variables are tested in the model (chapter 5). Two items, the surface area of distribution centers (in m^2) and urban density levels are visualized in Figure 4.4a and Figure 4.4b, respectively.

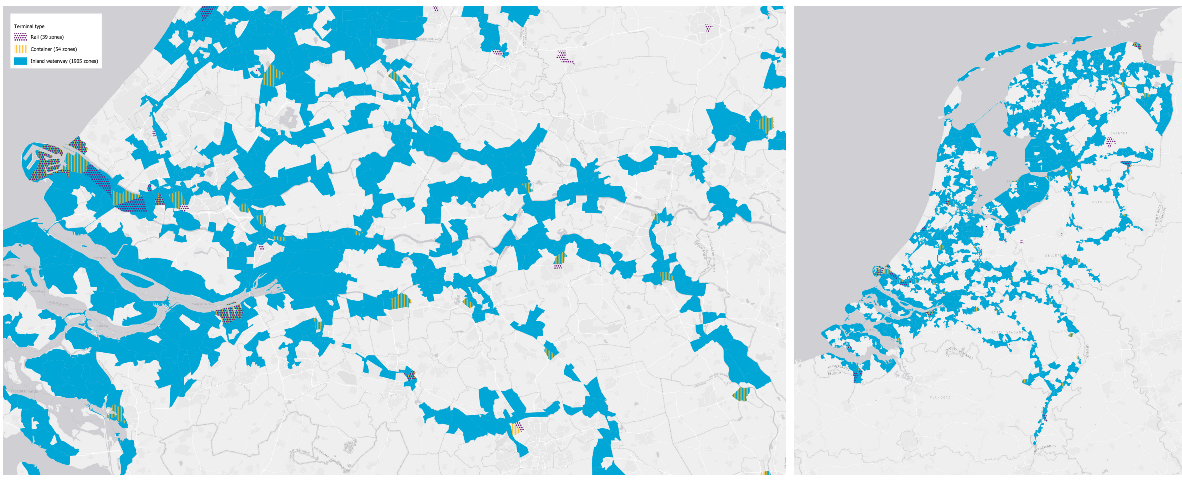


(a) Surface area distribution centers

(b) Urban density levels

Figure 4.4: Distribution of surface area distribution centers and urban density level per NRM zone

Finally, the BasGoed data has additional information regarding terminals. Within the NRM zones, some zones have terminals located within the zone. In total, 54 zones have container terminals, 39 zones have rail terminals and 1905 zones have inland waterway terminals. Within the availability of a container terminal, the number of containers which have been transshipped and what kind of container terminal (rail or water) is available. An overview of the terminals in the Netherlands is depicted in Figure 4.5.



(a) Port of Rotterdam and mid country

(b) The Netherlands

Figure 4.5: Distribution of terminals per NRM zone

CBS Location Data

Next, the CBS location data is described. As mentioned earlier, this consists of squares of 100 metres by 100 metres. Loads of information are paired to these squares. Each square shows information like gender, age distribution, amount of houses and their age and distances to different facilities like pharmacies. In total, there are 130 items per square. A few of these items are taken into account as explaining factors and tested in the regression model. These are more elaborated in the following sections and chapter. Across the Netherlands, a total of 379,138 is distributed.

Dutch Zoning Plan Data

The Dutch zoning plan data is described in this section. The zoning plan data consists of all the zoning plans up to 2024 and is geographically located across the Netherlands in a total of 2,592,204 planes. Each plane has its use, like residential, sports, agricultural, business or industrial. A total of 22 different uses are described by the data.

4.3.2. Relevant Variables

This section presents the relevant variables of the three datasets. First, the BasGoed freight tour data is presented, next, the CBS location data and this section ends with the dataset of the Dutch zoning plans.

BasGoed Freight Tour Data

In the BasGoed freight tour data, the variables shown in Table 4.3 are intended to be relevant. The first 5 variables are used for the dependent variable in the regression model: goods types LS and NST/R, the origin and destination of the trip for trip counts (FTG) and the corresponding weight (FG).

Furthermore, the province in which the NRM zone is located is taken into account. Finally, the population, surface area of distribution centers, employment per zone, availability of a terminal and urban density are taken into account.

No.	Label in dataset	Description
		BasGoed freight tour data
1	`nstr`	The classification of goods based on the NST/R goods type.
2	`logistic_segment`	The classification of goods based on the Logistic Segment.
3	`origin_nrm`	The origin zone of trip.
4	`destination_nrm`	The destination zone of trip.
5	`trip_weight_ton`	The weight transported on the trip.
		NRM base data
6	`PROVINCIE`	The province in which the zone is located.
		BasGoed zonal data
7	`population`	The population size.
8	`surface_distr_meter2`	The surface area of a distribution center.
9	`empl_landbouw`	Employment in the agricultural sector.
10	`empl_industrie`	Employment in the industrial sector.
11	`empl_detail`	Employment in the retail sector.
12	`empl_diensten`	Employment in the services sector.
13	`empl_overheid`	Employment in the governmental sector.
14	`empl_overig`	Employment in the others sector.
15	`terminal_container`	Availability of a container terminal.
16	`terminal_rail`	Availability of a rail terminal.
17	`terminal_water`	Availability of an inland waterway terminal.
18	`urban_density`	The urban density level.

Table 4.3: Variables used from the BasGoed freight tour data

To visualize the

CBS Location Data

The CBS location data is used to gather information on the age distribution per NRM zone. The variables shown in Table 4.4 were therefore considered as relevant.

No.	Label in dataset	Description
1	`INW_014`	The number of inhabitants under the age of 15 on the 1st of January.
2	`INW_1524`	The number of inhabitants aged 15 to 45 on the 1st of January.
3	`INW_2544`	The number of inhabitants aged 25 to 44 on the 1st of January.
4	`INW_2564`	The number of inhabitants aged 45 to 64 on the 1st of January.
5	`INW_65PL`	The number of inhabitants aged 65 or higher on the 1st of January.

Table 4.4: Variables used from the CBS location data

Dutch Zoning Plan Data

Finally, the Dutch zoning plan data was inspected and the relevant variables were determined. In consultation with the chair of this thesis' committee, the following variables were considered relevant and tested in the regression model:

No.	Label in dataset	Description
1	`business`	The surface area of business zoning plans.
2	`industrial`	The surface area of industrial zoning plans.
3	`city center`	The surface area of city center zoning plans.
4	`retail`	The surface area of retail zoning plans.
5	`mixed`	The surface area of mixed zoning plans.
6	`horeca`	The surface area of horeca zoning plans.
7	`office`	The surface area of office zoning plans.
8	`residential`	The surface area of residential zoning plans.
9	`urban development`	The surface area of urban development zoning plans.

Table 4.5: Variables used from the zoning plan data

4.4. Data Preparation

This section describes the steps taken from the raw data provided from the source to a usable dataset which can be applied for regression analysis.

4.4.1. BasGoed Freight Tour Data

Dependent Variables

Regarding the BasGoed freight tour data, the following steps have been taken to prepare the dependent variables:

1. The first step was to filter the data
   - All trips with an origin and destination which had an NRM zone number > 6441 (outside the Netherlands) were left out of the dataset
   - All trips with a vehicle type 8, vans, were also left out of the dataset
     The choice of leaving out vans rests in the generation of van data in the BasGoed freight tour data. According to the manual, the van data is based on a gravity model and not on either

XML or survey data ("BasGoed Functionele Documentatie", 2023). Estimating a model using this type of data would lead to non-realistic conclusions.

2. Next the trips were counted per LS and per NST/R goods type
3. The counted trips were added per NRM zone:
    - If the NRM zone was the destination, it would added to the attraction of the corresponding segment type.
    - If the NRM zone was the origin, it would added to the production of the corresponding segment type.
4. The weight was sorted per NRM zone in the same manner:
    - If the NRM zone was the destination, the weight corresponding to the trip would be added to the attraction of the corresponding segment type.
    - If the NRM zone was the origin, the weight corresponding to the trip would be added to the production of the corresponding segment type.
5. The preparation concludes by adding the values to the corresponding NRM zones in a new table.

Independent Variables

Regarding the BasGoed freight tour data, the following steps have been taken to prepare the independent variables:

1. All trips with an origin and destination which had an NRM zone number > 6441 (outside the Netherlands) were left out of the dataset
2. Next, the independent variables were added to the table with rows of the NRM zones

Terminal Data

Regarding the BasGoed freight tour data, the following steps have been taken to prepare terminal data:

1. For each terminal, if in that zone a terminal was present, the corresponding NRM zone number would get the value 1 in the corresponding terminal column
2. Next, the terminal data was added to the table with rows of the NRM zones

4.4.2. CBS Location Data

To incorporate the CBS location data on the same NRM zone level, a few modifications have been applied using QGIS and Python:

1. The first step was to locate all centroids of the 100 x 100 tiles in QGIS
2. Next, the centroids of the tiles were matched with an underlying NRM zone (intersect)
3. The centroids of the tiles that fell out of the zones concerning the Netherlands (> NRM zone 6441) were replaced inside the nearest NRM zone < 6441
4. The centroids were rematched with their underlying NRM zone
5. And finally, in Python the corresponding attribute table was read and all the information was summed per zone and added to the table with rows of the NRM zones

4.4.3. Dutch Zoning Plan Data

To incorporate the Dutch zoning plan data on the same NRM zone level, a few modifications have been applied using again QGIS and Python:

1. The first step was to split all different land uses

2. Next, the land uses, which were also laid out in space (see Figure 4.2c), were matched with the underlying NRM zone (intersect)
3. The next was to dissolve all land uses within an NRM zone into one big plane
4. This was conducted per different types of land use
5. And finally, in Python the corresponding attribute tables were read and all the information was added to the main table, per land use

4.5. Trip Rates

As the dataset is now prepared in such a way that a model can be fit on it, some first insights in terms of index numbers can be given using trip rates (as seen in Figure 3.1 from (Beagan et al., 2007)). This is given by the sum of trips per category divided by the total number of employees. The results are production and attraction rates per Logistic Segment per employee per workweek and are shown in Table 4.6. The numbers differ significantly from the trip rates shown in Beagan et al., 0.027 - 0.500 trips/employee/workday in Beagan et al.'s study and 0.003 - 0.139 trips/employee/workweek in this study. One of the reasons for the difference being so significant could be that a large number of employees do not cause any trip to happen, for instance, employees working in the services sector. Also, the trip rates shown in Beagan et al. are determined for highway trips only. Moreover, the trip rates the trip rate calculation in this section is a very basic one.

The trip rates for NST/R goods types are shown in Table B.1.

Description	**Prod. trips/Employee/Workweek**	**Attr. trips/Employee/Workweek**
Food (general cargo)	0.040	0.039
Miscellaneous (general cargo)	0.139	0.137
Conditioned transport	0.034	0.033
Facility logistics	0.013	0.012
Construction logistics	0.032	0.032
Waste	0.023	0.023
Parcel (consolidated flows between sorting centres)	0.003	0.003
Hazardous materials	0.020	0.019

Table 4.6: Weekly trip rate of production/attraction per Logistic Segment

4.6. Conclusion

The conclusion of this chapter is given by answering sub-question 1b:

Sub-question 1b: Which models are feasible, given the limited available data?

The BasGoed freight tour data, simulated data from the BasGoed model based on an XML dataset of the Netherlands, consists of both trips made between zones and weight transported, both to (attraction) and from (production) a zone. The availability of this data gives the opportunity to analyse both Freight Trip Generation as Freight Generation. Therefore this is both considered in the model estimations in chapter 5.

Although there are a lot of possibilities in the BasGoed data, the focus of this study will be on the prediction of Logistic Segments and NST/R goods type. Although the NST/R goods types are taken into account in the model development and estimation, considering the practical gain regarding Dutch logistics, only LS will be presented and discussed in the results and conclusion of this study. The results of NST/R will still be made available in this report.

Furthermore, regarding the aggregation level of the model, the finest disaggregated level is the NRM zoning, 6441 zones (aggregated geographic representations) in the Netherlands. Therefore the model estimation will be conducted on that level. An overview of these zones is shown in Figure 4.3.

Finally, regarding the explaining variables, the final set is chosen based on suitability and availability. This leaves a total of 26 independent variables to be tested. A list of these variables can be found in Table 5.3.

5

Model Development and Results

This chapter describes the development of a linear regression model which predicts the Freight Trip Generation and Freight Generation in the Netherlands for each Logistic Segment. The estimations for by NST/R goods type are included in Appendix D. This chapter also gives the results of the model (section 5.3). As is explained in section 4.4, this chapter has as a starting point a table of dependent and independent variables. The development of the model is presented in section 5.1, which also consists of a more in-depth overview of the dependent and independent variables. Moreover, it included the choices made for stratification[1], and the verification and validation of the presented model are described. After a set of independent variables has been found, the final model can be built, statistically tested and validated. The conclusion of the model development is shown in section 5.2 and the results can be found in section 5.3.

5.1. Model Development Process

The linear regression model, which will indicate what explaining factors will be influencing the FTG or FG, will be presented and described in this section. Here the individual steps of the model development and key aspects of the model are described, the preliminary testing and stratification in particular. In short, the model development process and the steps elaborated in this section are:

1. First the data of BasGoed, CBS and the zoning plans are prepared, this is described in section 4.4. The data is merged into one table with the NRM zones as the length of the table and the dependent and independent variables as columns.
    - First the dependent variables are described in subsection 5.1.1.
    - The independent variables are then described in subsection 5.1.2.
2. In the next step some preliminary testing of the data is conducted. This results in the choice of the forward stepwise regression and the omitting of some independent variables, see subsection 5.1.3.
3. The next step is the choice of the stratification of the model. This is explained in subsection 5.1.4.
4. The following step is the verification and validation of the model, see subsection 5.1.5.

[1] In this research: the act of sorting the data in order to have a more fair distribution along the test and train dataset.

5.1.1. Dependent Variables

As the data is sufficiently comprehensive, the dependent variables of this model are freight trips and freight weight. Each dependent variable is divided into production (FTP/FP) and attraction (FTA/FA). In addition, each production and attraction is sorted per Logistic Segment. An overview of the statistics of this dependent variable is shown in Table 5.2, but first Table 5.1 will show a refresher showing which labels belong to which type of logistics segment. The same statistics per NST/R goods type are shown in Table C.2.

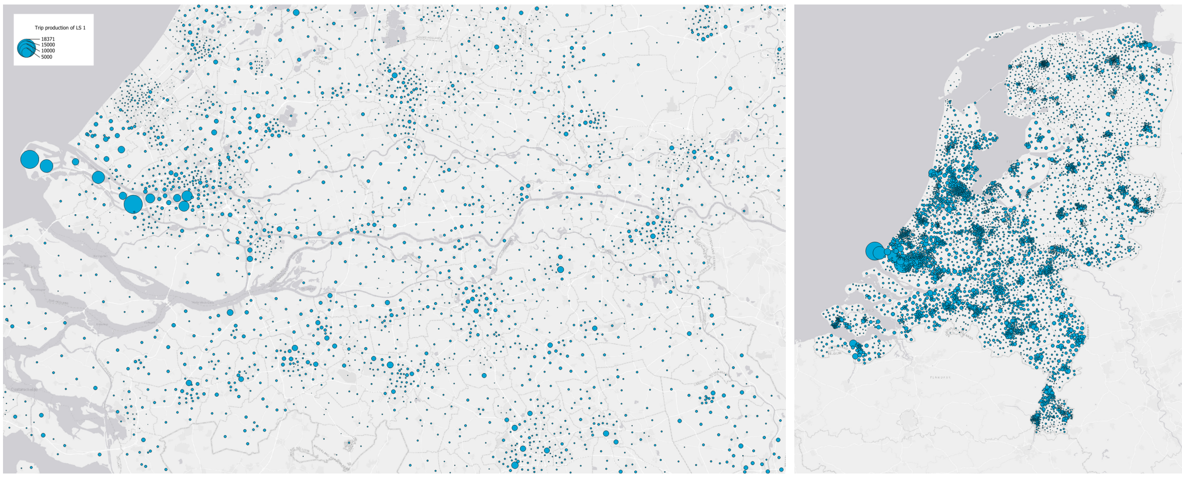
Label	Description
LS 0	Food (general cargo)
LS 1	Miscellaneous (general cargo)
LS 2	Conditioned transport
LS 3	Facility logistics
LS 4	Construction logistics
LS 5	Waste
LS 6	Parcel (consolidated flows between sorting centres)
LS 7	Hazardous materials

Table 5.1: Logistic Segment label description

Variable	Max.	Min.	Avg.	SD	Variable	Max.	Min.	Avg.	SD
Trips/Workweek/Zone					*Tonnes/Workweek/Zone*				
FTP LS 0	3623	0	53.32	105.22	**FP LS 0**	45643.80	0	632.02	1323.65
FTP LS 1	18371	0	187.21	430.58	**FP LS 1**	195913.28	0	1940.90	4668.38
FTP LS 2	2954	0	45.30	91.07	**FP LS 2**	32131.89	0	507.27	1028.91
FTP LS 3	1516	0	16.88	36.37	**FP LS 3**	13485.19	0	148.27	346.64
FTP LS 4	2212	0	42.45	68.99	**FP LS 4**	27351.42	0	470.08	865.02
FTP LS 5	3253	0	30.67	72.65	**FP LS 5**	28223.02	0	244.74	608.44
FTP LS 6	414	0	3.68	8.89	**FP LS 6**	3161.53	0	28.51	71.16
FTP LS 7	5979	0	27.19	92.56	**FP LS 7**	74440.60	0	269.30	1116.48
FTA LS 0	3429	0	52.76	95.84	**FA LS 0**	36206.82	0	626.39	991.20
FTA LS 1	17233	0	184.03	385.57	**FA LS 1**	182978.14	0	1937.60	3822.04
FTA LS 2	2825	0	43.93	82.05	**FA LS 2**	29392.84	0	496.88	857.90
FTA LS 3	1460	0	16.76	32.96	**FA LS 3**	12503.07	0	148.22	283.88
FTA LS 4	2063	0	42.85	62.38	**FA LS 4**	21501.96	0	475.31	604.82
FTA LS 5	3143	0	30.37	69.65	**FA LS 5**	25419.05	0	242.64	532.83
FTA LS 6	395	0	3.65	8.41	**FA LS 6**	3065.95	0	28.43	63.52
FTA LS 7	4820	0	26.11	75.20	**FA LS 7**	37322.98	0	261.23	615.17

Table 5.2: Dependent variable list and characteristics per Logistic Segment

In the statistics, it is clear that some segments have big outliers in the dataset (i.e. FTP LS 1, miscellaneous (general cargo), with a max. of 18371 trips/workweek/zone and a mean of 187.21 trips/workweek/zone). These mainly occur in the harbour district of Rotterdam. To visualize this, an overview of all zones and their share to the trip production of LS 1 is depicted in Figure 5.1, in trips/workweek. The bigger the circle, the higher the amount of trips.



(a) Port of Rotterdam and mid country (b) The Netherlands

Figure 5.1: Distribution of the trip production of LS 1 per NRM zone

5.1.2. Independent Variables

This chapter briefly describes the independent variables used in this model. First, the description for each variable is shown in Table 5.3. Additional information regarding the surface area and what is meant with each term can be found in Table C.3 in Appendix C. Next, the statistics of the independent variables are shown in Table 5.5. Note that several concepts are covered under the term of employment: jobs (both filled jobs and vacancies), employed people and volume of labour (Statistiek, n.d.-b).

Variable	Description
Population	The amount of inhabitants in the zone.
Age 0 - 14	The percentage of inhabitants aged under 14 years old in the zone.
Age 15 - 24	The percentage of inhabitants aged between 15 and 24 years in the zone.
Age 25 - 44	The percentage of inhabitants aged between 25 and 444 years in the zone.
Age 45- 64	The percentage of inhabitants aged between 45 and 64 years in the zone.
Age 65+	The percentage of inhabitants aged above 65 years old in the zone.
Surface area distribution centers	The amount of surface area (in km^2) of distribution centers in the zone.
Employment agricultural	The number of employment in the agricultural sector in the zone.
Employment industrial	The number of employment in the industrial sector in the zone.
Employment retail	The number of employment in the retail sector in the zone.
Employment services	The number of employment in the services sector in the zone.
Employment government	The number of employment in the governmental sector in the zone.
Employment others	The number of employment in any other sector than above in the zone.
Terminal container	The presence of a container terminal in the zone.
Terminal rail	The presence of a rail terminal in the zone.
Terminal inland waterway	The presence of an inland waterway terminal in the zone.
Urban density	The degree of urban density in the zone, see Table 5.4 for an explanation per level.
Surface area business	The amount of surface area (in km^2) of business area in the zone.
Surface area industrial	The amount of surface area (in km^2) of industrial area in the zone.
Surface area city center	The amount of surface area (in km^2) of a city center in the zone.
Surface area retail	The amount of surface area (in km^2) of retail in the zone.
Surface area mixed	The amount of surface area (in km^2) of mixed land-use in the zone.
Surface area horeca	The amount of surface area (in km^2) of horeca in the zone.
Surface area office	The amount of surface area (in km^2) of offices in the zone.
Surface area residential	The amount of surface area (in km^2) of residential in the zone.
Surface area urban development	The amount of surface area (in km^2) of urban development in the zone.

Table 5.3: Description of independent variables

Urban density level	Population density $(inhabitants/ha)$
1	≤ 2.5
2	2.5 - 6
3	6 - 25
4	25 - 60
5	50 - 85
6	> 85

Table 5.4: Description of independent variables

Variable	Type	Max.	Min.	Avg.	SD
Population	Linear	19518	0	2683.16	2846.11
Age 0 - 14	Linear	100	0	13.91	10.38
Age 15 - 24	Linear	100	0	9.23	9.31
Age 25 - 44	Linear	100	0	21.17	12.72
Age 45- 64	Linear	100	0	31.35	16.89
Age 65+	Linear	100	0	17.41	13.18
Surface area distribution centers (m^2)	Linear	564597.00	0	4876.18	23859.16
Employment agricultural	Linear	3287	0	34.37	88.54
Employment industrial	Linear	10333	0	239.46	510.89
Employment retail	Linear	4604	0	120.40	241.58
Employment services	Linear	15291	0	289.80	723.57
Employment government	Linear	18936	0	138.15	461.81
Employment others	Linear	16883	0	521.07	841.10
Terminal container	Dummy	-	-	-	-
Terminal rail	Dummy	-	-	-	-
Terminal inland waterway	Dummy	-	-	-	-
Urban density	Dummy	-	-	-	-
Surface area business (m^2)	Linear	17681732.19	0	76907.76	351524.10
Surface area industrial (m^2)	Linear	14114249.77	0	80194.04	328913.13
Surface area city center (m^2)	Linear	893055.69	0	6005.72	21172.72
Surface area retail (m^2)	Linear	178442.97	0	3195.32	10374.99
Surface area mixed (m^2)	Linear	8334765.18	0	28886.31	227664.02
Surface area horeca (m^2)	Linear	171244.38	0	3141.64	8817.90
Surface area office (m^2)	Linear	218048.16	0	2863.89	12110.06
Surface area residential (m^2)	Linear	2856406.98	0	252186.54	275153.52
Surface area urban development (m^2)	Linear	2189175.62	0	29281.35	122559.19

Table 5.5: Independent variable list and characteristics

5.1.3. Preliminary Testing

After the dependent and independent variables have been determined, some preliminary testing of the model is conducted. The testing, building (and running) of the model is done using the coding language Python. In particular, the Python module Skikit-learn (Pedregosa et al., 2011) and Statsmodels (Seabold & Perktold, 2010). At first, no stratification or method to choose the best independent variables was applied to the model. However, some statistical testing was needed to be added to the model (subsection 5.1.3) and a selection method to optimize the model (Table 5.1.3).

Statistical Testing

To test the model statistically, the use of five different statistical tests were added:

1. R^2
   The R^2 is a parameter which shows how much the dependent variable is explained by an independent variable or variables in this regression model. It is given by the following formula:

$$R^2 = 1 - \frac{SS_{Res}}{SS_{Tot}} \tag{5.1}$$

   where:

   - SS_{Res}, the regression sum of squares, is the sum of the residuals squared (difference between predicted values of y (dependent variable) and observed values of y)
   - SS_{Tot} is the total sum of squares (sum of the distance the data is away from the mean all squared)

2. **Adjusted** R^2
   The adjusted R^2 penalizes the addition of unnecessary variables which are added to increase the initial R^2 (even if they don't improve the model significantly). The adjusted R^2 is calculated using the following formula:

$$Adjusted\ R^2 = 1 - \frac{SS_{Res}/(n-k)}{SS_{Tot}/(n-1)} \tag{5.2}$$

   where:

   - n is the number of data points
   - k is the number of independent variables

3. **Predicted** R^2
   The predicted R^2 is added to determine how well the regression model makes predictions. This statistically helps identify cases where the model provides a good fit for the dependent variable, but is not as good at making predictions (which is a must in this research). Moreover, the predicted R^2 helps determine whether the regression model is overfitting. It is calculated using:

$$Predicted\ R^2 = 1 - \frac{PRESS}{SS_{Tot}} \tag{5.3}$$

   where:

   - $PRESS$ is the predicted residual sum of squares and is used to provide a parameter of the fit of a model to a sample of observations that were not used to estimate the model (X_{test}/y_{test}, also see subsection 5.1.4)
   - The SS_{Tot} here is the total sum of squares of the test dataset

4. t**-test**
   The t-test is added to this regression model to evaluate the significance of an individual independent variable by assessing whether the corresponding coefficient is significantly different from zero. The associated t-value measures the size of the coefficient relative to its standard error. A larger absolute t-value suggests stronger evidence against the null hypothesis that the coefficient is zero. Here a t-value of ≥1.96 is chosen as sufficiently large to cover a 97.5% cumulative probability confidence level of effect on the dependent variable. In other words, it suggests that the relationship between the predictor and the outcome is unlikely to be due to random chance. The coefficient in particular is therefore an important contributor to the model's predictive power. A t-value with a corresponding 97.5% cumulative probability confidence level confirms that this study used a one-tailed t-test. The reasoning behind that choice lies in the search for explaining variables which generate freight (trips), thus an addition to zero. The expected direction of the coefficients is positive.

5. **Variance Inflation Factor (VIF)**
   The VIF value is a measure used to detect multicollinearity in a regression analysis. A high VIF indicates that the associated predictor variable is highly correlated with other predictors in the model. Generally, a VIF value exceeding 10 is considered indicative of problematic multicollinearity. It is calculated using the following formula:

$$VIF_j = \frac{1}{1 - R_j^2} \tag{5.4}$$

   where:

   - R_j^2 is the R^2 obtained by regressing the predictor variable j on all the other predictor variables

Furthermore, the correlation between all the independent variables is observed (except the dummy variables). This is shown in Table 5.6. The table shows no correlation above > 0.50 which indicates that the individual predictors do not correlate with each other (so little change on multicollinearity).

Feature Selection Methods

After securing the statistical testing in the model, the options for how to pick the best predictor variables are next. At first, common sense was used to predict the model outcome. This gave an insight into how the model works and what outcome should be expected. However, time was of the essence and going through 4 different model types (FTP, FTA,FP and FA) and 2 different groups (LS and NST/R) would take a lot of time. Therefore, methods to optimize this model lead to the following possible methods and their (dis)advantages:

1. Test all possible combinations
   This method implies the testing of all possible combinations of predictor variables. This leads to the best possible result (highest R^2) and thus the best fit of the model.
   + This will lead to the best possible result of the model.
   – The number of combinations will be 2^{31}. This will need a lot of computational power and time.
2. Forward selection method
   The forward selection method is a stepwise regression method used to select the most relevant predictors for a regression model. It starts with an empty model and iteratively adds predictors one at a time, choosing the one that improves the model's fit the most according to the best R^2. The iteration continues until no further improvement is observed. This method helps to identify the subset of predictors that best explain the variation in the dependent variable.
   + Takes a lot less time and computation power to implement.
   – Will not lead to the best possible result of the model.
3. Backward elimination method
   This method starts with a model that includes all predictors and removes the least significant predictor at each step until no further improvement is observed.
   + Takes a lot less time and computation power to implement.
   – Less easy to implement; t-value threshold could not be implemented.
   – Will not lead to the best possible result of the model.

After some testing, the choice was made very fast and easily. The forward selection method was defined as the best method given the size and criteria of this research. To test all possible combinations, time and computational power were not in reach. Moreover, the backward elimination method did not give significant differences in R^2, compared to the forward selection method, and was not able to implement the t-value statistic within the hand. The research continues with the use of the forward selection method.

Name	#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Population	**1**	1.00	0.17	0.16	0.29	-0.10	0.09	-0.10	0.04	0.03	0.47	0.21	0.19	0.33	-0.05	-0.09	0.22	0.08	0.03	0.02	0.10	0.46	0.06
Age 0 - 14	**2**	0.17	1.00	0.00	0.20	-0.10	-0.15	-0.06	0.07	-0.02	-0.03	-0.06	-0.04	-0.06	-0.05	-0.04	-0.04	0.01	0.00	0.00	-0.01	0.15	0.06
Age 15 - 24	**3**	0.16	0.00	1.00	0.14	-0.13	-0.15	-0.05	-0.01	0.01	0.11	0.13	0.16	0.11	-0.03	-0.04	0.05	0.00	0.02	-0.02	0.06	0.01	-0.01
Age 25 - 44	**4**	0.29	0.20	0.14	1.00	-0.17	-0.13	-0.03	0.00	0.01	0.20	0.22	0.16	0.18	-0.06	-0.07	0.13	0.02	0.02	-0.02	0.07	0.06	0.03
Age 45 - 64	**5**	-0.10	-0.10	-0.13	-0.17	1.00	0.03	-0.06	0.05	-0.08	-0.10	-0.12	-0.10	-0.15	-0.06	-0.06	-0.06	-0.01	-0.01	0.04	-0.06	0.07	0.00
Age 65+	**6**	0.09	-0.15	-0.15	-0.13	0.03	1.00	-0.11	0.01	-0.09	0.09	-0.08	0.01	0.00	-0.06	-0.11	0.11	0.04	0.01	0.09	0.00	0.17	0.01
Surface area distribution centers	**7**	-0.10	-0.06	-0.05	-0.03	-0.06	-0.11	1.00	0.07	0.40	0.01	0.12	-0.02	0.26	0.20	0.39	-0.03	0.01	0.00	0.01	0.05	-0.07	-0.02
Employment agricultural	**8**	0.04	0.07	-0.01	0.00	0.05	0.01	0.07	1.00	0.05	-0.02	-0.05	-0.06	-0.01	0.08	0.06	0.02	0.08	0.01	0.11	-0.04	0.29	0.06
Employment industrial	**9**	0.03	-0.02	0.01	0.01	-0.08	-0.09	0.40	0.05	1.00	0.12	0.26	0.04	0.31	0.21	0.41	0.02	0.14	0.02	0.03	0.11	0.02	0.01
Employment retail	**10**	0.47	-0.03	0.11	0.20	-0.10	0.09	0.01	-0.02	0.12	1.00	0.35	0.28	0.46	-0.02	0.00	0.51	0.19	0.04	0.03	0.15	0.15	0.01
Employment services	**11**	0.21	-0.06	0.13	0.22	-0.12	-0.08	0.12	-0.05	0.26	0.35	1.00	0.34	0.51	0.02	0.06	0.20	0.05	0.05	0.01	0.43	-0.03	-0.03
Employment government	**12**	0.19	-0.04	0.16	0.16	-0.10	0.01	-0.02	-0.06	0.04	0.28	0.34	1.00	0.34	-0.02	-0.03	0.15	0.02	0.03	-0.01	0.25	-0.01	-0.01
Employment others	**13**	0.33	-0.06	0.11	0.18	-0.15	0.00	0.26	-0.01	0.31	0.46	0.51	0.34	1.00	0.09	0.14	0.25	0.07	0.04	0.08	0.26	0.05	-0.01
Surface area business	**14**	-0.05	-0.05	-0.03	-0.06	-0.06	-0.06	0.20	0.08	0.21	-0.02	0.02	-0.02	0.09	1.00	0.02	0.00	0.05	0.00	0.05	0.00	0.05	0.02
Surface area industrial	**15**	-0.09	-0.04	-0.04	-0.07	-0.06	-0.11	0.39	0.06	0.41	0.00	0.06	-0.03	0.14	0.02	1.00	-0.02	0.04	0.00	0.01	0.02	-0.03	0.02
Surface area city center	**16**	0.22	-0.04	0.05	0.13	-0.06	0.11	-0.03	0.02	0.02	0.51	0.20	0.15	0.25	0.00	-0.02	1.00	0.06	0.01	0.05	0.06	0.19	0.10
Surface area retail	**17**	0.08	0.01	0.00	0.02	-0.01	0.04	0.01	0.08	0.14	0.19	0.05	0.02	0.07	0.05	0.04	0.06	1.00	0.01	0.12	0.06	0.19	0.02
Surface area mixed	**18**	0.03	0.00	0.02	0.02	-0.01	0.01	0.00	0.01	0.02	0.04	0.05	0.03	0.04	0.00	0.00	0.01	0.01	1.00	0.00	0.02	0.04	0.01
Surface area horeca	**19**	0.02	0.00	-0.02	-0.02	0.04	0.09	0.01	0.11	0.03	0.03	0.01	-0.01	0.08	0.05	0.01	0.05	0.12	0.00	1.00	0.05	0.28	0.02
Surface area office	**20**	0.10	-0.01	0.06	0.07	-0.06	0.00	0.05	-0.04	0.11	0.15	0.43	0.25	0.26	0.00	0.02	0.06	0.06	0.02	0.05	1.00	0.02	0.00
Surface area residential	**21**	0.46	0.15	0.01	0.06	0.07	0.17	-0.07	0.29	0.02	0.15	-0.03	-0.01	0.05	0.05	-0.03	0.19	0.19	0.04	0.28	0.02	1.00	0.01
Surface area urban development	**22**	0.06	0.06	-0.01	0.03	0.00	0.01	-0.02	0.06	0.01	0.01	-0.03	-0.01	-0.01	0.02	0.02	0.10	0.02	0.01	0.02	0.00	0.01	1.00

Table 5.6: Correlation matrix of independent variables

5.1.4. Stratification

The final part of optimizing the model is the stratification of the test and train dataset. First of all, the dataset is divided into these two sets to estimate the predicted R^2. Having that R^2 helps show the predictive of a model; it shows how well the model will perform on new, unseen data (the test data). It therefore tells something about the generalization of the model. Ideally, the R^2 and the predicted R^2 should not differ that much, which suggests a good generalization. In pursuit of this, different manners of composing the test and train data have been investigated. The stratification option was integrated into the same module mentioned before, Skikit-learn. The following 4 options have been taken into account and have been tested on the FTP of both LS and NST/R. Their results are seen in Table 5.9:

1. A test/train split of 20/80 with no stratification
   The dataset is split according to the default settings: 20% of the data is chosen to be the test set and 80 % is chosen to be the training set. With no stratification in place, the picking of which parts are set to the test set or train set are picked randomly.
2. A test/train split of 20/80 with a stratification in level of urban density
   The dataset is again split according to the default settings: 20% of the data is chosen to be the test set and 80 % is chosen to be the training set. As a stratification rule, the level of urban density is chosen. The reasoning behind it is that if a model needs to represent an evenly geographical spread, all parts of the Netherlands - from the countryside to the dense cities, the urban density level could be a good measure. The result of this rule is an evenly distributed test and train set where the ratio of urban density level is the same in the test and train dataset, as it is in the full dataset. In this dataset, the distribution of the level of urban density is shown in Table 5.7.

Urban density level	No. of zones	Percentage
1	1020	15.84
2	480	7.45
3	678	10.53
4	1280	19.87
5	1687	26.19
6	1296	20.12
Total	6441	100.00

Table 5.7: Distribution of urban density level

3. A test/train split of 20/80 with a stratification per province
   The dataset is again split according to the default settings: 20% of the data is chosen to be the test set and 80 % is chosen to be the training set. As a stratification rule, the distribution of province zones is chosen. Again, the reasoning behind it is that if a model needs to represent an evenly geographical spread, all parts of the Netherlands - from the countryside to the dense cities, the distribution of the provinces could be a good measure for this. The result of this rule is an evenly distributed test and train set where the ratio of province distribution is the same in the test and train dataset, as it is in the full dataset. In this dataset, the distribution of provinces is shown in Table 5.8.

Province	No. of zones	Percentage
Drenthe	353	5.48
Flevoland	168	2.61
Friesland	641	9.95
Gelderland	790	12.27
Groningen	387	6.01
Limburg	492	7.64
Noord-Brabant	884	13.72
Noord-Holland	735	11.41
Overijssel	566	8.79
Utrecht	376	5.84
Zeeland	253	3.93
Zuid-Holland	796	12.36
Total	6441	100.00

Table 5.8: Distribution of provinces

4. A test/train split of 50/50 with no stratification
   Finally, the dataset is split according to a 50/50 split: 50% of the data is chosen to be the test set and 50 % is chosen to be the training set. With no stratification in place, the picking of which parts are set to the test set or train set are picked randomly. One of the reasons for this is that if both sets have an equal number of samples, the evaluation of the model's performance is likely to be more reliable.

Test/Train	20/80						50/50	
Stratification	None		Urban density		Provinces		None	
	R^2	*Pred. R^2*	*R^2*	*Pred. R^2*	*R^2*	*Pred. R^2*	*R^2*	*Pred. R^2*
LS 0	0.60	-0.01	0.63	0.51	0.56	0.75	0.60	0.48
LS 1	0.51	0.42	0.68	0.46	0.50	0.82	0.53	0.69
LS 2	0.58	0.56	0.63	0.47	0.53	0.79	0.58	0.55
LS 3	0.57	0.37	0.65	0.46	0.55	0.76	0.57	0.48
LS 4	0.65	0.51	0.65	0.57	0.63	0.59	0.61	0.54
LS 5	0.60	0.74	0.71	0.48	0.58	0.77	0.55	0.66
LS 6	0.53	0.67	0.63	0.41	0.53	0.57	0.51	0.50
LS 7	0.36	-1.31	0.54	0.22	0.35	0.14	0.41	-0.83
NST/R 0	0.61	0.79	0.61	0.45	0.53	0.74	0.58	0.52
NST/R 1	0.57	-0.11	0.65	0.50	0.54	0.81	0.58	0.59
NST/R 2	0.42	0.27	0.43	0.32	0.42	0.19	0.43	0.21
NST/R 3	0.22	-6.65	0.32	0.11	0.20	-0.51	0.25	-2.44
NST/R 4	0.54	0.76	0.65	0.44	0.51	0.78	0.50	0.66
NST/R 5	0.45	-0.17	0.49	0.36	0.43	0.47	0.46	0.07
NST/R 6	0.50	0.27	0.56	0.39	0.49	0.61	0.50	0.45
NST/R 7	0.31	-1.36	0.51	0.43	0.25	0.48	0.31	0.51
NST/R 8	0.46	0.04	0.62	0.34	0.44	0.77	0.49	0.29
NST/R 9	0.59	0.37	0.67	0.44	0.57	0.57	0.58	0.35
NST/R -1	0.27	0.71	0.51	0.41	0.23	0.30	0.29	0.34
Difference	14.68		3.37		3.94		6.08	

Table 5.9: Performance of different stratification tactics

The results of the first tests of stratification and choice of test/train split lead to some interesting results. All different models were tested on their R^2 and predicted R^2 and from the outcomes, the sum of absolute differences between both, which shows the generalization and predictive power of the model, is calculated. The results show that the test/train split of 20/80 and 50/50 with no stratification give the worst outcomes and a test/train split of 20/80 with both a stratification choice of urban density or provinces gives less differences between both R^2, 3.37 and 1.56 absolute difference respectively. Since only the FTP was used in the first tests, a more extensive test is conducted using all four different models on both LS and NST/R. The results are shown in Table 5.10.

Type	FTP				FTA				FP				FA			
Stratification	Urban density		Provinces		Urban density		Provinces		Urban density		Provinces		Urban density		Provinces	
	R^2	*Pred.* R^2	R^2	*Pred.* R^2	R^2	*Pred.* R^2	R^2	*Pred.* R^2	R^2	*Pred.* R^2	R^2	*Pred.* R^2	R^2	*Pred.* R^2	R^2	*Pred.* R^2
LS 0	0.63	0.51	0.56	0.75	0.58	0.51	0.58	0.18	0.58	0.44	0.49	0.72	0.56	0.48	0.57	0.11
LS 1	0.68	0.46	0.50	0.82	0.55	0.44	0.51	0.41	0.62	0.39	0.41	0.81	0.57	0.45	0.53	0.32
LS 2	0.63	0.47	0.53	0.79	0.62	0.49	0.58	0.55	0.57	0.43	0.46	0.75	0.61	0.49	0.58	0.45
LS 3	0.65	0.46	0.55	0.76	0.61	0.48	0.58	0.30	0.59	0.44	0.48	0.76	0.60	0.50	0.59	0.21
LS 4	0.65	0.57	0.63	0.59	0.66	0.58	0.66	0.46	0.54	0.50	0.53	0.52	0.67	0.57	0.67	0.37
LS 5	0.71	0.48	0.58	0.77	0.72	0.48	0.61	0.70	0.64	0.43	0.51	0.77	0.73	0.47	0.60	0.72
LS 6	0.63	0.41	0.53	0.57	0.65	0.43	0.55	0.57	0.49	0.35	0.43	0.39	0.62	0.43	0.55	0.50
LS 7	0.54	0.22	0.35	0.14	0.49	0.19	0.35	-1.48	0.50	0.24	0.34	0.44	0.46	0.16	0.33	-3.13
NST/R 0	0.61	0.45	0.53	0.74	0.68	0.53	0.62	0.78	0.58	0.41	0.44	0.77	0.68	0.50	0.59	0.83
NST/R 1	0.65	0.50	0.54	0.81	0.54	0.50	0.55	0.01	0.63	0.48	0.47	0.74	0.48	0.49	0.51	-0.18
NST/R 2	0.43	0.32	0.42	0.19	0.44	0.33	0.43	0.23	0.16	0.14	0.19	-0.20	0.52	0.38	0.49	0.34
NST/R 3	0.32	0.11	0.20	-0.51	0.22	0.09	0.21	-15.29	0.48	0.28	0.25	0.76	0.18	0.07	0.20	-23.27
NST/R 4	0.65	0.44	0.51	0.78	0.66	0.46	0.54	0.74	0.59	0.42	0.40	0.76	0.62	0.52	0.51	0.76
NST/R 5	0.49	0.36	0.43	0.47	0.40	0.39	0.47	-1.86	0.46	0.33	0.35	0.72	0.36	0.38	0.45	-2.71
NST/R 6	0.56	0.39	0.49	0.61	0.53	0.40	0.51	0.21	0.34	0.23	0.24	0.54	0.57	0.44	0.55	0.22
NST/R 7	0.51	0.43	0.25	0.48	0.24	0.30	0.29	-1.51	0.49	0.44	0.21	0.42	0.22	0.30	0.28	-2.01
NST/R 8	0.62	0.34	0.44	0.77	0.53	0.31	0.45	0.07	0.61	0.35	0.38	0.82	0.49	0.31	0.43	-0.03
NST/R 9	0.67	0.44	0.57	0.57	0.67	0.44	0.60	0.24	0.61	0.40	0.49	0.67	0.66	0.45	0.60	0.16
NST/R -1	0.51	0.41	0.23	0.30	0.35	0.43	0.26	0.88	0.51	0.41	0.23	0.30	0.35	0.43	0.26	0.88
Difference	3.37		3.94		2.63		25.35		2.88		5.01		2.49		37.22	

Table 5.10: Performance of 20/80 models using a stratification in urban density and provinces

The results of this final test show that the stratification choice of provinces gives a total absolute difference of 71.52 whereas urban density gives a total absolute difference of 11.37. Moreover, in some cases using provinces as a stratification, the predicted R^2 shows signs of a very misfit model (values far below zero). In terms of regression analysis, the prediction using test data shows results which are even worse than a horizontal line (the mean of the dataset). The level of urban density, however, shows only positive results and the absolute difference is also a lot smaller. This concludes that the stratification based on the level of urban density is superior since it gives a more generalized result (small absolute difference) and will be used in the models predicted from this point onwards.

5.1.5. Verification & Validation

Now that the model is ready to be discussed for the estimation results, first, the verification and validation of the model are covered, subsection 5.1.5 and subsection 5.1.5 respectively. The estimated parameters will be discussed in section 5.3.

Verification

This model rests on estimation methods which have been commonly used. Also for this specific method, it is a trusted method (also see Table 3.1). Therefore any further verification of this model is not addressed in this research.

Validation

The validation of the model is addressed using known explanatory factors found in the literature (chapter 3), as those factors are known to be independent and significant in those models. The following three findings are discussed, which are found as well in the literature, as in the model results discussed in this study:

- Employment, surface area and commodity type are commonly significant explaining factors in FTG/FG (Table 3.1)
  In the results of the models, employment and surface area are shown to be the most significant explaining factors. This is also observed in the literature. Moreover, the different types of commodities have very different influencing factors on the total trips or weight produced or attracted. This shows that commodity type as well is a very important factor in the production or attraction. This is also found in the literature.
- Employment is more suitable for representing FP while surface area explains FA better (Sahu & Pani, 2020)
  The results from this model confirm the observation. Although in both FP and FA are employment as surface area significant explaining factors. In FP, employment has a larger value of the standardized coefficient (more about that in section 5.3) whereas in FA, the surface area has a larger value of the standardized coefficient. This is in line with the observation from the literature since a larger standardized value indicated a stronger relationship between the predictor variable and the outcome variable.
- Employment is invariably considered the most preferred causal variable in FG/FTG models (Pani et al., 2018)
  75 out of 76 models show that any kind of employment is a significant explaining factor to the FTG or FG. Only in the FP for NST/R goods type 7 (fertilizers) explaining factors in employment lack. This shows a very large preferred causal link with employment and confirms the observation from the literature.
- The trip rates found in Table 4.6 do correspond with the results from the models, although both rates are in trips/workweek, the model results are 5-10 times higher. It could only explained by the better prediction power of the model which incorporates other factors as well instead of just counting and dividing to calculate the trip rates.

Sub-question 1c: How strong is the explanatory power of the model?

The model results show that the explanatory power of the model is robust, demonstrating high significance for the predicting factors of freight production and attraction by having high t-values. The results show accurate findings which meet the standards found in other literature and the model effectively uses variables such as the surface area of distribution centers, employment in various sectors, the presence of terminals in a zone, and urban density levels. This extensive set of explanatory variables ensures that the model can predict logistics movements with a high degree of accuracy. Moreover, the model has a high generalization power, it shows approximately similar results with unseen as with seen data. However, on average, ≈50-60% of the dataset is fitted correctly. This means that there should be kept in mind that the possibility of over- or underfitting exists.

5.2. Conclusion of Model Estimation

This model estimation section is concluded with a summary of what choices are made in the model development which led to the final model estimation process presented in Figure 5.2. The results from the model will be discussed in the following section.

- The model will consist of a multiple linear regression analysis.
- The number of dependent variables and thus models to be tested are:
  $(8\ \text{LS} + 11\ \text{NST/R}) \cdot (\text{FTP} + \text{FTA} + \text{FP} + \text{FA}) = 19 \cdot 4 = 76.$
- The age distribution as an independent variable has been left out of the model, leaving a total of 21 independent variables to be tested. 9 of them (3 terminal variables and 6 urban density levels)

are dummy variables.

- A forward selection method is chosen as the method to determine the best features in the model.
- The statistical tests used in this model are the R^2, adjusted R^2, predicted R^2, one-tailed t-test and the VIF.
- The dataset of 6441 NRM zones is split into a test/train set of 20% and 80% using a stratification according to the distribution of the urban density level.

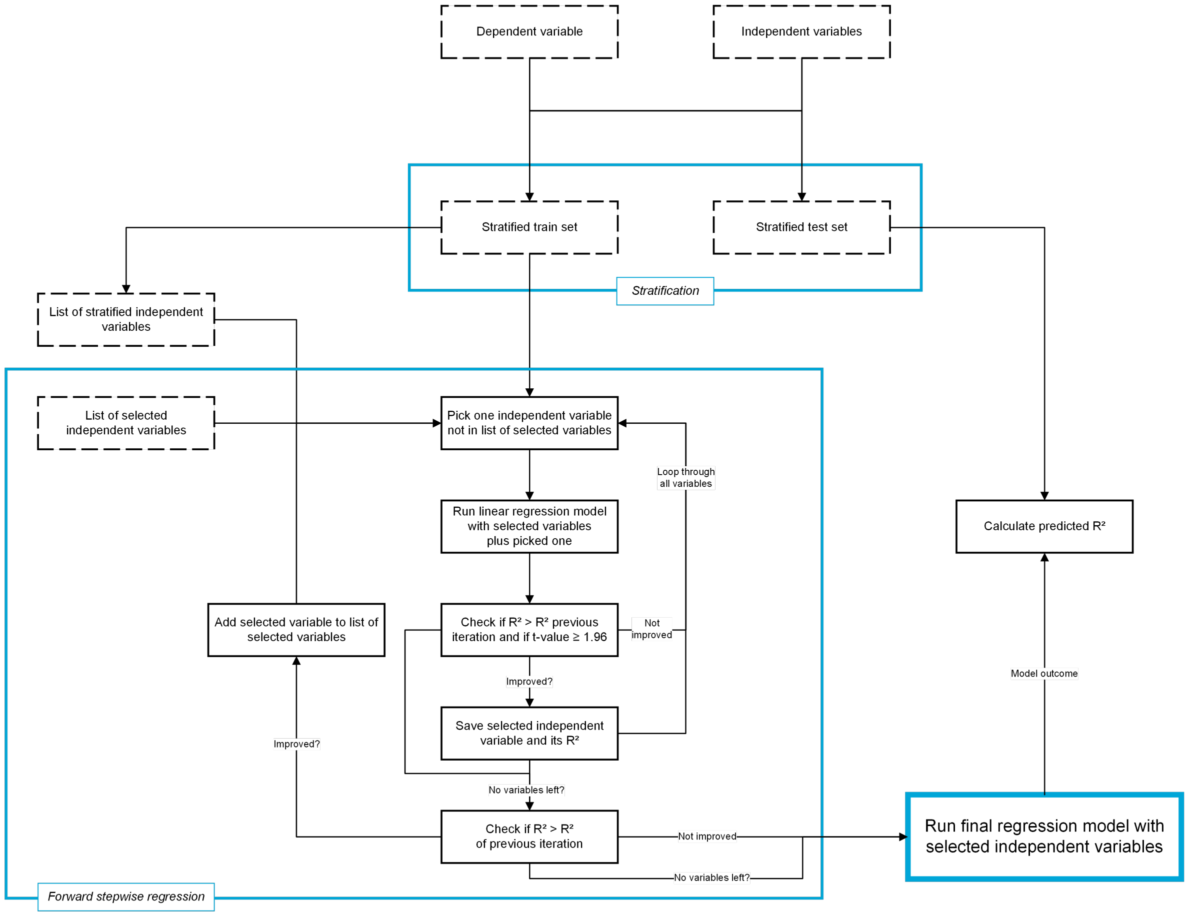


Figure 5.2: Model estimation process

5.3. Model Results

This section describes the results found from the 76 regression models. First, some general notes:

- All results shown in the table are found to be significant[2] coefficients for the Netherlands. The cumulative probability confidence level is at least 97.5% and may even be upped to a confidence level of 99.999% (where $t - value \geq 3.090$).
- All coefficients shown in the table are positive, meaning that the coefficient will increase the amount of trips or weight.
- The R^2 of all 76 models are distributed from $0.67 \leq R^2 \leq 0.19$ and have a mean of 0.46, this means that two-thirds to one-fifth of the variation in the dependent variables is accounted for by the significant independent variables.

[2] With exceptions to some of the intercepts.

- The absolute difference between the R^2 and the predicted R^2 is distributed from $0.01 \leq \Delta R^2 \leq 0.33$ with a mean of 0.15. This shows that the generalization and the predictive power of the models are very high.
- The results of the regression model, showing only significant outcomes to the 'population' for the Netherlands can be read using the following formula:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \tag{5.5}$$

  where:
  - $\hat{Y}$ is the predicted value of the dependent variable (trips or weight)
  - β_0 is the intercept
  - $\beta_1, \beta_2, \dots, \beta_n$ are the coefficients for the significant independent variables $X_1, X_2, \dots, X_n$ respectively
  - $X_1, X_2, \dots, X_n$ are the values of the significant independent variables
- The intercept in the results can be seen as a fixed correction factor, however, when none of the other coefficients has any value, the intercept should not be interpreted.
- In the result tables, 'St. coefficient' is shown as an extra coefficient. This means the standardized value of the coefficient. It shows the change in the dependent variable for a one standard deviation change in the independent variable while holding all other variables constant. As mentioned before, a larger value of a standardized coefficient indicates a stronger magnitude between the independent and the dependent variable. The row in the table is also marked with blue colours. The darker the blue, the higher the standardized coefficient, and the stronger the relationship with the dependent variable.
- The results should be put into perspective: some high coefficients would suggest a very high influence on production or attraction. However, the maximum values of the independent variables (and thus the different ratios between each other, see Table 5.5) differ significantly and should be taken into account. For example, the amount of m^2 industrial or business zoning would be significantly higher in a particular zone than the amount of m^2 distribution centers. Therefore, per m^2 (or per km^2 in the result tables), the coefficient of the surface area of distribution centers is most likely to have a higher number to indicate its high influencing factor in freight (trip) generation. The standardized coefficients will transform these different units to a unitless and comparable number.

The following parts of this section discuss the model outcomes per freight (trip) generation model. However, this section discusses only 24 of the 76 models. Only the models per LS are discussed in this section, the results per NST/R goods type can be found in Appendix D. First, the FTG per LS is discussed in subsection 5.3.1, here the production, attraction and comparison of both is discussed. Next, the FG is discussed with also the production, attraction and comparison between both in subsection 5.3.2. This section concludes with a summary of the results.

5.3.1. Freight Trip Generation

First, the generation of trips is discussed. This is conducted by investigating the production and attraction and searching the comparisons and differences between both.

Freight Trip Production

For the Freight Trip Production of the different Logistic Segments, which can be found in Table 5.11, the following explaining factors turned out to be significant:

- Significant:

  - The surface area of distribution centers
  - The employment in the agricultural, industrial, retail and services sector
  - The presence of rail or inland waterway terminals
  - The urban density level of 3, 4, 5 and 6
  - The surface area of business, industrial and offices
- Not significant:
  - The population
  - The age distribution
  - The employment in the governmental and other sector
  - The presence of a container terminal
  - The urban density level of 1 and 2 (not dense)
  - The surface area of city centers, retail, mixed, horeca, residential and urban development.

In most segments, the industrial employment in the zone is found to be the most influencing factor of trip production, as second best the surface area of business in the zone shows a high influence and in some cases, it is a tie in terms of magnitude with the surface area of distribution centers in the zone. An increase in employment for the industrial sector shows the largest increase for the Logistic Segment of miscellaneous (general cargo), 261 trips/1000 employment. General food cargo and construction logistics are second with half to a third of the trip increase of miscellaneous cargo. Moreover, in the same miscellaneous cargo segment, the employment in the agricultural sector shows an increase which is twice the size of industrial employment (419 trips/1000 employment).

Furthermore, the urban density level only has a minor effect on the total amount of trips. The level itself increases the amount with a maximum of 50, while the increase of 1000 in employment increases the amount of trips with 261 in that same segment.

What is interesting in these results, is that the addition of an inland waterway terminal increases the consolidated parcel flows with ≈1 trip per workweek. The significance of this explaining factor shows that a small amount of parcels are already transported using inland waterways. This, however, is still a small amount. Another explanation perhaps could be a correlation with the location of inland waterways. Zones with inland waterways are bound to be in industrial locations, well accessible by other modes so also a perfect place for parcel depots to be located. The surface area of distribution centers is found to be a predictor of the FTP of almost all Logistic Segments, which is a very straightforward result. The presence of a distribution center will in many cases lead to an increase of freight trips in that area. What is interesting, is that this predictor is not the most influencing for the generation of trips, employment and surface area of zoning plans exceed that.

Freight Trip Attraction

For the Freight Trip Attraction of the different Logistic Segments, which can be found in Table 5.12, the following explaining factors turned out to be significant:

- Significant:
  - Surface area distribution centers
  - The employment in the agricultural, industrial, services and other sector
  - The presence of rail or inland waterway terminals
  - The urban density level of 2, 3, 4, 5 and 6
  - The surface area of business, industrial and offices
- Not significant:
  - The population

- The age distribution
- The employment in the retail and governmental sector
- The presence of a container terminal
- The urban density level of 1 (not dense)
- The surface area of city centers, retail, mixed, horeca, residential and urban development

In the attraction of trips, the surface area of business has a slightly higher influencing power than the employment in the industrial sector, but both turn out to be highly influencing. What else is interesting about these results, is that the increase in office spaces leads to an increase in trips attracted to the zone in the construction logistics. Each km^2 gives an increase of 218 trips. One could argue that this is related to the construction of those offices. Another reason could be that many real estate developments are in and around office sites. Also, more regular maintenance activities around office buildings are likely in a competitive real estate sector with short renting periods. Moreover, in all segments, the trip attraction decreases if the urban density level is increased from 5 to 6.

Comparison of Production and Attraction

When comparing the production and attraction of trips for each LS, the following observations are noticed:

- Freight trips are (in small amounts) produced in the retail sector although they are not attracted there. They, however, are (in small amounts) attracted to the 'others' sector but not produced there.
- Having a rail terminal in a zone causes freight to be attracted to it in terms of trips, whereas the production of trips caused by a rail terminal in those zones is rather low. This could mean that the transhipment of rail-to-truck is larger than truck-to-rail.
- An urban density level which is lower is more likely to attract freight trips than producing it.
- In terms of surface area (distribution centers and zoning plans), the production of trips is one-fourth lower than the attraction of trips (3332 vs. 4188). However, if you only look at the zoning plans, the production and attraction are evened out.
- The trip prediction decreases only in attraction if the urban density level is raised from 5 to 6.

5.3.2. Freight Generation

This section discusses the generation of tonnes. This is also conducted by investigating the production and attraction and searching the comparisons and differences between both.

Freight Production

For the Freight Production of the different Logistic Segments, which can be found in Table 5.13, the following explaining factors turned out to be significant:

- Significant:
  - The surface area of distribution centers
  - The employment in the agricultural, industrial, retail and services sector
  - The presence of rail or inland waterway terminals
  - The urban density level of 3, 4, 5 and 6
  - The surface area of business, industrial, retail and offices
- Not significant:
  - The population
  - The age distribution

Table 5.11: Estimation results for explaining Freight Trip Production per LS

Trips/Workweek/Zone *N = 5152*			Intercept		Surface area distribution centers	Employment: Agricultural	Employment: Industrial	Employment: Retail	Employment: Services	Terminal: Rail	Terminal: Inland waterway	Urban density: 3	Urban density: 4	Urban density: 5	Urban density: 6	Surface area: Business	Surface area: Industrial	Surface area: Office	R^2	Adj. R^2	Predicted R^2
		Unit	-		*km²*	*x1000*				-	-	-	-	-	-	*km²*					
LS 0		**Coefficient**	**12.03**		**229.71**		**105.48**	**7.48**			**8.86**		**5.88**	**7.93**		**110.42**			**0.58**	*0.58*	*0.51*
	Food (general cargo)	St. coefficient		53.57	5.55		53.70	1.84			3.95		2.35	3.48		40.66					
		T-value	*8.17*	*58.72*	*5.55*		*53.27*	*2.00*			*4.26*		*2.44*	*3.61*		*43.24*					
		VIF			*1.26*		*1.51*	*1.19*			*1.22*		*1.10*	*1.15*		*1.12*					
LS 1		**Coefficient**	**18.18**		**430.99**	**418.65**	**260.79**				**43.04**		**29.50**	**45.18**	**50.27**	**592.59**	**62.18**		**0.55**	*0.55*	*0.44*
	Miscellaneous (general cargo)	St. coefficient		187.97	10.41	38.58	132.76				19.19		11.77	19.86	20.16	218.23	21.42				
		T-value	*2.44*	*50.63*	*2.45*	*10.03*	*30.80*				*5.07*		*2.77*	*4.51*	*4.65*	*56.54*	*5.01*				
		VIF			*1.37*	*1.09*	*1.70*				*1.29*		*1.10*	*1.12*	*1.12*	*1.14*	*1.41*				
LS 2		**Coefficient**	**-1.02**		**175.35**	**329.08**	**40.32**				**11.06**	**7.20**	**10.87**	**15.46**	**14.91**	**140.86**	**9.64**		**0.62**	*0.62*	*0.49*
	Conditioned transport	St. coefficient		45.59	4.24	30.32	20.53				4.93	2.21	4.34	6.80	5.98	51.87	3.32				
		T-value	*-0.56*	*58.37*	*4.75*	*37.48*	*22.61*				*6.19*	*2.48*	*4.48*	*6.71*	*6.06*	*63.85*	*3.70*				
		VIF			*1.37*	*1.12*	*1.70*				*1.32*	*1.08*	*1.11*	*1.13*	*1.13*	*1.15*	*1.41*				
LS 3		**Coefficient**	**3.41**		**98.19**		**34.72**		**1.05**	**22.60**				**3.10**	**2.84**	**37.07**			**0.61**	*0.61*	*0.48*
	Facility logistics	St. coefficient		16.91	2.37		17.68		0.77	1.78				1.36	1.14	13.65					
		T-value	*8.22*	*58.53*	*7.51*		*54.56*		*2.47*	*5.94*				*4.46*	*3.58*	*44.64*					
		VIF			*1.26*		*1.52*		*1.36*	*1.07*				*1.11*	*1.23*	*1.17*					
LS 4		**Coefficient**	**11.60**		**124.54**		**90.68**					**4.99**	**6.27**	**8.35**	**6.27**	**40.14**		**106.87**	**0.66**	*0.66*	*0.58*
	Construction logistics	St. coefficient		42.56	3.01		46.16					1.53	2.50	3.67	2.51	14.78		1.32			
		T-value	*10.49*	*81.06*	*5.24*		*79.99*					*2.55*	*3.88*	*5.51*	*3.85*	*27.32*		*2.48*			
		VIF			*1.26*		*1.51*					*1.02*	*1.06*	*1.10*	*1.10*	*1.11*		*1.09*			
LS 5		**Coefficient**	**2.17**		**924.45**		**47.71**		**1.83**	**59.18**	**8.49**		**3.85**	**6.13**	**6.22**	**61.76**	**11.42**		**0.72**	*0.72*	*0.48*
	Waste	St. coefficient		30.79	22.33		24.29		1.34	4.65	3.78		1.54	2.69	2.49	22.74	3.93				
		T-value	*2.50*	*67.01*	*42.54*		*44.35*		*2.70*	*9.45*	*8.07*		*2.94*	*5.02*	*4.56*	*45.90*	*7.19*				
		VIF			*1.37*		*1.79*		*1.37*	*1.13*	*1.27*		*1.10*	*1.14*	*1.29*	*1.22*	*1.48*				
LS 6		**Coefficient**	**0.50**		**103.66**		**6.25**		**0.22**	**6.42**	**0.97**			**0.48**	**0.53**	**6.12**	**0.94**		**0.65**	*0.65*	*0.43*
	Parcel (consolidated flows between sorting centres)	St. coefficient		3.67	2.50		3.18		0.16	0.50	0.43			0.21	0.21	2.26	0.32				
		T-value	*5.22*	*59.77*	*35.78*		*43.58*		*2.44*	*7.68*	*6.93*			*3.25*	*3.16*	*34.15*	*4.43*				
		VIF			*1.37*		*1.76*		*1.37*	*1.13*	*1.23*			*1.13*	*1.27*	*1.22*	*1.48*				
LS 7		**Coefficient**	**5.89**				**58.34**						**4.63**	**3.53**		**65.22**			**0.49**	*0.49*	*0.19*
	Hazardous materials	St. coefficient		26.76			29.70						1.85	1.55		24.02					
		T-value	*6.96*	*45.16*			*49.11*						*2.97*	*2.49*		*39.87*					
		VIF					*1.22*						*1.05*	*1.08*		*1.09*					

Table 5.12: Estimation results for explaining Freight Trip Attraction per LS

Trips/Workweek/Zone N = 5152					Employment				Terminal		Urban density					Surface area					
		Intercept		Surface area distribution centers	Agricultural	Industrial	Services	Others	Rail	Inland waterway	2	3	4	5	6	Business	Industrial	Office	R²	Adj. R²	Predicted R²
	Unit	-		km²		x1000			-	-	-	-	-	-	-		km²				
LS 0	Coefficient	13.65		364.80		81.33			37.44	7.90			8.46	11.54	6.74	124.70			0.63	0.63	0.51
Food (general cargo)	St. coefficient		53.09	8.81		41.40			2.94	3.52			3.37	5.07	2.70	45.92					
	T-value	9.39	68.75	10.43		48.66			3.68	4.48			3.84	5.66	3.08	56.02					
	VIF			1.26		1.51			1.07	1.26			1.09	1.12	1.12	1.18					
LS 1	Coefficient	20.45		957.04	287.92	195.61		12.43	247.72	36.63		20.44	36.49	53.54	51.67	608.41	44.45		0.68	0.68	0.46
Miscellaneous (general cargo)	St. coefficient		184.85	23.12	26.53	99.58		10.64	19.46	16.33		6.27	14.56	23.54	20.72	224.05	15.31				
	T-value	3.19	67.11	7.25	9.29	30.33		3.46	6.60	5.81		1.99	4.25	6.53	5.71	75.33	4.67				
	VIF			1.40	1.12	1.78		1.73	1.13	1.32		1.09	1.14	1.20	1.37	1.23	1.49				
LS 2	Coefficient	3.25		228.23	226.73	33.30		2.27	35.26	9.69		7.01	9.97	14.60	13.40	132.34	5.28		0.63	0.62	0.47
Conditioned transport	St. coefficient		44.19	5.51	20.89	16.95		1.94	2.77	4.32		2.15	3.98	6.42	5.37	48.73	1.82				
	T-value	2.03	64.43	6.95	29.38	20.73		2.53	3.77	6.17		2.74	4.66	7.15	5.94	65.81	2.22				
	VIF			1.40	1.12	1.78		1.73	1.13	1.32		1.09	1.14	1.20	1.37	1.23	1.49				
LS 3	Coefficient	2.41		139.06		25.20	0.90		27.37	1.26	2.50	2.87	3.21	5.76	5.46	40.76			0.65	0.65	0.46
Facility logistics	St. coefficient		16.81	3.36		12.83	0.66		2.15	0.56	0.66	0.88	1.28	2.53	2.19	15.01					
	T-value	3.76	69.08	12.61		46.61	2.50		8.54	2.26	2.31	2.96	3.87	7.29	6.42	58.01					
	VIF			1.26		1.60	1.37		1.07	1.33	1.04	1.06	1.11	1.15	1.30	1.18					
LS 4	Coefficient	14.17		217.45		74.15						5.29	6.13	9.92	7.93	48.16		105.81	0.65	0.65	0.57
Construction logistics	St. coefficient		43.01	5.25		37.75						1.62	2.45	4.36	3.18	17.73		1.31			
	T-value	13.89	88.77	9.91		70.87						2.94	4.11	7.09	5.28	35.52		2.66			
	VIF			1.26		1.51						1.02	1.06	1.10	1.10	1.11		1.09			
LS 5	Coefficient	3.56		914.63		37.06	1.82		59.34	8.68			4.25	6.99	6.62	65.88	12.89		0.71	0.70	0.48
Waste	St. coefficient		30.46	22.10		18.87	1.34		4.66	3.87			1.69	3.07	2.65	24.26	4.44				
	T-value	4.18	67.40	42.78		35.02	2.74		9.63	8.39			3.29	5.82	4.93	49.77	8.24				
	VIF			1.37		1.79	1.37		1.13	1.27			1.10	1.14	1.29	1.22	1.48				
LS 6	Coefficient	0.63		104.67		4.77	0.21		6.12	1.02			0.36	0.69	0.67	6.44	1.20		0.63	0.63	0.41
Parcel (consolidated flows between sorting centres)	St. coefficient		3.65	2.53		2.43	0.15		0.48	0.45			0.14	0.30	0.27	2.37	0.41				
	T-value	5.59	61.45	37.22		34.26	2.36		7.56	7.47			2.12	4.35	3.79	36.97	5.84				
	VIF			1.37		1.79	1.37		1.13	1.27			1.10	1.14	1.29	1.22	1.48				
LS 7	Coefficient	8.08				44.15			24.32				4.70	3.99		65.62			0.54	0.54	0.22
Hazardous materials	St. coefficient		25.88			22.48			1.91				1.87	1.75		24.16					
	T-value	12.23	55.85			47.51			3.99				3.86	3.60		49.80					
	VIF					1.22			1.07				1.05	1.08		1.15					

  - The employment in the governmental and other sector
  - The presence of a container terminal
  - The urban density level of 1 and 2 (not dense)
  - The surface area of city centers, mixed, horeca, residential and urban development

For freight production, employment in the industrial sector is one of the most influencing factors (with the highest t-values). Moreover, the biggest impact on the amount of tonnes is found in the miscellaneous segment.

Freight Attraction

For the Freight Attraction of the different Logistic Segments, which can be found in Table 5.14, the following explaining factors turned out to be significant:

- Significant:
  - The population
  - Surface area distribution centers
  - The employment in the agricultural, industrial, services and other sector
  - The presence of container, rail or inland waterway terminals
  - The urban density level of 2, 3, 4, 5 and 6
  - The surface area of business, industrial and offices
- Not significant:
  - The population
  - The age distribution
  - The employment in the retail and governmental sector
  - The urban density level of 1 (not dense)
  - The surface area of city centers, retail, mixed, horeca, residential and urban development

In most segments, the surface area of business in the zone is found to be the most influencing factor. As second best, employment in the industrial sector is found to be very influential. Moreover, what has been noticed is that the weight attraction of all segments decreases if the urban density level is increased from 5 to 6, this, unlike production, is only seen in the attraction.

verschil tussen FP en FTP

Comparison of Production and Attraction

When comparing the production and attraction of weight for each LS, the following observations are noticed:

- Distribution centers in the LSs mainly cause freight weight to be attracted to a zone, the production of the weight is rather small (23,727 vs. 6,144 tonnes/workweek/zone/km^2 distribution center).
- Freight weight is mostly produced in the services sector and is attracted by employment in the 'others' sector.
- Having a rail terminal in a zone causes freight to be attracted to it in terms of tonnes, whereas the production of weight caused by a rail terminal in those zones is rather low. The explanation for this is found in the transhipment opportunity in these zones.
- An urban density level which is lower (≤ 4) is more likely to attract freight weight and not produce it.
- An increase in surface area in retail is more likely to produce freight weight instead of attracting it.

- The weight prediction decreases only in attraction if the urban density level is raised from 5 to 6.

5.4. Conclusion of Model Results

In conclusion, the following observations can be found from all four different models:

- Distribution centers in the LSs mainly cause freight weight to be attracted to a zone, the production of the weight is rather small. However, in terms of trips, both production and attraction are equally distributed (what goes in, also goes out).
- Rail terminals increase the attraction of trips and weight, but do not cause (a lot of) production. This could mean that the transhipment of rail to truck is larger than truck to rail.
- An urban density level which is lower is more likely to attract freight than producing it.
- Only in the freight (trip) attraction, urban density level 2 is significant.
- The attraction of both trips and weight decreases when the urban density level is increased from 5 to 6.

Moreover, the conclusion in terms of explaining factors is addressed in the answering of the following sub-question:

Sub-question 1: What are the explaining factors for freight (trip) generation of logistics in the Netherlands?

The answering of this sub-question is a very straightforward one, the explaining factors for freight (trip) generation of logistics in the Netherlands are:

- Freight Trip Generation
  - The surface area of distribution centers
  - The employment in the agricultural, industrial, retail and services and other sector
  - The presence of rail or inland waterway terminals
  - The urban density level of 2, 3, 4, 5 and 6
  - The surface area of business, industrial and offices
- Freight Generation
  - The population
  - The surface area of distribution centers
  - The employment in the agricultural, industrial, retail and services and other sector
  - The presence of container, rail or inland waterway terminals
  - The urban density level of 2, 3, 4, 5 and 6
  - The surface area of business, industrial, retail and offices

  However, it should be noted that the surface area of distribution centers, employment in agricultural and industrial sectors and the surface area of business zoning in zones have the biggest influence on generating trips and weight. Adjusting these parameters in a zone will lead to a significant increase (or decrease) of freight in a zone.

Table 5.13: Estimation results for explaining Freight Production per LS

Tonnes/Workweek/Zone *N = 5152*						Employment				Terminal		Urban density				Surface area						
			Intercept		Surface area distribution centers	Agricultural	Industrial	Retail	Services	Rail	Inland waterway	3	4	5	6	Business	Industrial	Retail	Office	R²	Adj. R²	Predicted R²
		Unit	-		*km²*		*x1000*			-	-	-	-	-	-		*km²*					
LS 0		**Coefficient**	**132.94**				**1527.01**				**98.40**			**68.35**		**1044.13**		**2855.22**		**0.56**	*0.56*	*0.48*
Food (general cargo)		St. coefficient		633.68			777.36				43.88			30.05		384.51		30.58				
		T-value	*8.18*	*54.80*			*64.82*				*3.74*			*2.58*		*32.61*		*2.61*				
		VIF					*1.28*				*1.17*			*1.11*		*1.10*		*1.08*				
LS 1		**Coefficient**	**-48.60**			**5048.79**	**3539.06**		**127.43**		**775.44**		**247.42**	**385.50**	**454.07**	**5758.23**	**431.68**			**0.57**	*0.56*	*0.45*
Miscellaneous (general cargo)		St. coefficient		1945.86		465.24	1801.65		93.56		345.80		98.74	169.48	182.06	2120.53	148.71					
		T-value	*-0.62*	*49.44*		*11.43*	*39.37*		*2.20*		*8.61*		*2.19*	*3.63*	*3.81*	*52.51*	*3.42*					
		VIF				*1.09*	*1.70*		*1.37*		*1.29*		*1.11*	*1.14*	*1.29*	*1.12*	*1.31*					
LS 2		**Coefficient**	**-31.28**			**3961.79**	**609.54**				**141.66**	**91.09**	**114.08**	**154.77**	**142.08**	**1431.54**	**84.67**			**0.61**	*0.60*	*0.49*
Conditioned transport		St. coefficient		509.47		365.07	310.30				63.17	27.95	45.52	68.04	56.97	527.18	29.17					
		T-value	*-1.51*	*57.15*		*39.58*	*30.71*				*6.94*	*2.75*	*4.13*	*5.90*	*5.06*	*57.63*	*2.96*					
		VIF				*1.12*	*1.61*				*1.32*	*1.08*	*1.11*	*1.13*	*1.13*	*1.12*	*1.30*					
LS 3		**Coefficient**	**16.44**				**369.07**		**12.59**	**171.42**				**25.21**	**25.87**	**331.52**			**514.86**	**0.60**	*0.60*	*0.50*
Facility logistics		St. coefficient		148.47			187.88		9.25	13.47				11.08	10.37	122.09			6.38			
		T-value	*4.03*	*52.17*			*62.67*		*2.74*	*4.58*				*3.68*	*3.32*	*40.88*			*2.01*			
		VIF					*1.32*		*1.62*	*1.07*				*1.11*	*1.23*	*1.15*			*1.30*			
LS 4		**Coefficient**	**118.65**				**1206.81**		**26.33**					**53.14**	**42.12**	**373.11**		**1332.80**		**0.67**	*0.67*	*0.57*
Construction logistics		St. coefficient		470.66			614.36		19.33					23.36	16.89	137.40		14.28				
		T-value	*12.57*	*72.67*			*89.31*		*2.76*					*3.41*	*2.37*	*20.83*		*2.18*				
		VIF					*1.35*		*1.36*					*1.12*	*1.22*	*1.09*		*1.08*				
LS 5		**Coefficient**	**1.72**		**5569.85**		**502.96**		**18.99**	**389.18**	**57.93**		**22.91**	**41.17**	**50.97**	**497.76**	**54.10**		**776.18**	**0.73**	*0.73*	*0.47*
Waste		St. coefficient		244.58	134.57		256.05		13.94	30.58	25.83		9.14	18.10	20.44	183.30	18.64		9.62			
		T-value	*0.25*	*66.86*	*32.18*		*58.71*		*3.21*	*7.81*	*6.91*		*2.19*	*4.23*	*4.69*	*46.46*	*4.27*		*2.35*			
		VIF			*1.37*		*1.79*		*1.63*	*1.13*	*1.27*		*1.10*	*1.15*	*1.29*	*1.22*	*1.48*		*1.30*			
LS 6		**Coefficient**	**1.46**		**574.00**		**61.20**	**3.84**	**3.51**	**39.35**	**6.52**			**3.51**	**4.48**	**47.77**		**111.13**		**0.62**	*0.62*	*0.43*
Parcel (consolidated flows between sorting centres)		St. coefficient		28.37	13.87		31.15	0.95	2.58	3.09	2.91			1.54	1.80	17.59		1.19				
		T-value	*1.77*	*54.70*	*24.41*		*52.76*	*1.67*	*4.40*	*5.76*	*5.51*			*2.79*	*3.12*	*31.96*		*2.23*				
		VIF			*1.26*		*1.60*	*1.48*	*1.50*	*1.07*	*1.23*			*1.18*	*1.32*	*1.18*		*1.14*				
LS 7		**Coefficient**	**45.41**				**711.96**									**597.94**				**0.46**	*0.46*	*0.16*
Hazardous materials		St. coefficient		262.24			362.44									220.20						
		T-value	*5.95*	*38.21*			*51.99*									*31.59*						
		VIF					*1.09*									*1.09*						

Table 5.14: Estimation results for explaining Freight Attraction per LS

Tonnes/Workweek/Zone N = 5152		Intercept		Population	Surface area distribution centers	Employment: Agricultural	Employment: Industrial	Employment: Services	Employment: Others	Terminal: Container	Terminal: Rail	Terminal: Inland waterway	Urban density: 2	Urban density: 3	Urban density: 4	Urban density: 5	Urban density: 6	Surface area: Business	Surface area: Industrial	Surface area: Office	R²	Adj. R²	Predicted R²
	Unit	-		x1000	km²		x1000			-	-	-	-	-	-	-	-		km²				
LS 0	Coefficient	224.52			3139.46		652.92				253.34	90.69		63.95	128.41	164.24	89.49	1423.90	61.17		0.58	0.58	0.44
Food (general cargo)	St. coefficient		631.93		75.85		332.38				19.90	40.44		19.62	51.24	72.20	35.88	524.37	21.07				
	T-value	11.87	74.04		7.78		33.45				2.18	4.64		2.01	4.88	6.63	3.39	56.97	2.07				
	VIF				1.37		1.71				1.13	1.31		1.05	1.10	1.13	1.13	1.22	1.48				
LS 1	Coefficient	476.71			9290.33	2512.44	999.89		122.60		2285.52	142.06	270.05	327.55	499.80	690.65	600.49	6258.59	433.08		0.62	0.62	0.39
Miscellaneous (general cargo)	St. coefficient		1944.54		224.46	231.52	509.02		105.01		179.56	63.35	70.92	100.50	199.45	303.63	240.77	2304.79	149.19				
	T-value	6.03	67.29		6.71	7.72	14.77		3.25		5.80	2.15	2.10	2.84	5.03	7.20	5.77	73.81	4.33				
	VIF				1.40	1.15	1.78		1.73		1.13	1.35	1.07	1.10	1.15	1.21	1.38	1.23	1.49				
LS 2	Coefficient	109.18			1672.01	1801.66	272.87		24.70		323.43	95.65		73.48	106.78	155.29	140.06	1428.83	66.88		0.57	0.57	0.43
Conditioned transport	St. coefficient		500.50		40.40	166.02	138.91		21.15		25.41	42.65		22.54	42.61	68.27	56.16	526.18	23.04				
	T-value	6.07	64.77		4.52	20.72	15.08		2.45		3.07	5.41		2.55	4.43	6.75	5.52	63.06	2.50				
	VIF				1.40	1.12	1.78		1.73		1.13	1.32		1.09	1.14	1.20	1.37	1.23	1.49				
LS 3	Coefficient	27.83			722.30		149.00		8.06		212.99	11.14	27.10	27.87	35.93	59.77	58.20	410.49	18.65		0.59	0.59	0.44
Facility logistics	St. coefficient		148.62		17.45		75.85		6.90		16.73	4.97	7.12	8.55	14.34	26.28	23.34	151.17	6.42				
	T-value	4.60	64.86		6.59		27.78		2.70		6.82	2.12	2.66	3.05	4.59	7.99	7.19	61.10	2.35				
	VIF				1.40		1.78		1.73		1.13	1.34	1.05	1.07	1.14	1.21	1.38	1.22	1.49				
LS 4	Coefficient	197.01		5.36	2171.05		571.22				155.83			60.60	73.06	120.56	92.61	558.37		965.92	0.54	0.54	0.50
Construction logistics	St. coefficient		477.05	15.21	52.45		290.80				12.24			18.59	29.16	53.00	37.13	205.62		11.97			
	T-value	17.00	89.36	2.66	8.91		49.50				2.22			3.04	4.41	7.63	5.33	36.32		2.20			
	VIF			2.04	1.28		1.51				1.07			1.09	1.21	1.43	1.54	1.17		1.10			
LS 5	Coefficient	30.29			6097.72		182.90		9.67		511.47	63.32	45.70	35.97	62.54	84.22	78.57	577.95	130.84		0.64	0.64	0.43
Waste	St. coefficient		243.33		147.32		93.11		8.29		40.18	28.24	12.00	11.04	24.96	37.02	31.50	212.83	45.07				
	T-value	3.04	64.48		33.76		20.71		1.97		9.95	7.33	2.72	2.39	4.85	6.83	5.89	52.23	10.03				
	VIF				1.40		1.78		1.73		1.13	1.34	1.05	1.07	1.14	1.21	1.38	1.22	1.49				
LS 6	Coefficient	9.70			633.61		23.59	2.45		15.29	57.79	6.67				4.62	4.05	49.99	12.90		0.49	0.49	0.35
Parcel (consolidated flows between sorting centres)	St. coefficient		28.51		15.31		12.01	1.80		1.42	4.54	2.98				2.03	1.62	18.41	4.44				
	T-value	11.71	53.76		25.01		19.03	3.14		2.37	7.92	5.49				3.60	2.77	31.28	6.91				
	VIF				1.40		1.77	1.37		1.24	1.14	1.24				1.13	1.27	1.28	1.53				
LS 7	Coefficient	119.34					272.05				277.06				49.43	44.88		671.37			0.50	0.50	0.24
Hazardous materials	St. coefficient		259.96				138.49				21.77				19.73	19.73		247.24					
	T-value	19.27	59.90				31.26				4.86				4.33	4.32		54.40					
	VIF						1.22				1.07				1.05	1.08		1.15					

6
Application

This chapter elaborates on the application of the explanatory model. The application is conducted to show the practical use of the model, to what extent the model is feasible to use in spatial planning scenarios and what results can be expected. Moreover, the application shows how this model can benefit, amongst others, policymakers in evaluating the impact of spatial scenarios. As shown in the previous chapters, the model can simultaneously explain freight trips and freight weight. In the application, the emphasis is on freight trips since these are easier to understand and picture for policymakers. The results and the consequential impact are therefore defined in the change of freight trips.

This application chapter will first discuss to what extent the model can be applied, and what spatial planning scenarios match with the application of the explanatory model (section 6.1). Next, a use case is described in which the model is applied (section 6.2). The outlines of the use case will first be described, after which its corresponding parameters from the base model are defined. Next, the increase of those parameters by the considered spatial planning scenario is estimated and the use case ends with the model application and the discussion of the results. This chapter ends with a conclusion and learning goals of the application (section 6.3).

6.1. Feasible Applications

As the explaining factors for the prediction of logistic movement are known due to the regression model (section 5.3), feasible applications for the prediction of logistic movements are to be found in line with the explaining factors. The explaining factors mainly show terms of surface area, employment, terminals and the urban density level as influencing factors for logistic movements.

Sub-question 2a: Which spatial planning scenarios can be considered using this model?

Within mind that the explaining factors for the prediction of logistic movements mainly concern surface area, employment, terminals and the urban density level, the following spatial planning scenarios can be feasible with this prediction model:

1. Land use
   A spatial planning scenario which explores different ways of allocating and utilizing land within a province, municipality or region would fit within the feasibility of this prediction model. The surface areas of the different uses would change and this model can show the impact of that in terms of logistic movements.

2. Economic development
   A spatial planning scenario like economic development, which stimulates economic growth and creates employment, would be feasible for this prediction model. The employment of different areas would change and this model can show the impact of that in terms of logistic movements.

3. Urban growth
   An urban growth scenario, which would explore new areas for cities or municipalities to grow, would fit in the feasibility of this model. The urban density level of the model in combination with its employment and surface area factors could show the impact of urban growth. However, the impact of logistic movements would be more visible in an urban growth scenario which improves employment and increases logistical facilities like terminals or distribution centers.

6.2. Use Case: A12 Zone

In this section, the use case of the A12 zone is elaborated. The A12 zone project near Utrecht is a large-scale urban and infrastructural development initiative aimed at addressing the housing shortage and improving connectivity in the region. This project involves the collaboration of local municipalities, including Utrecht, Nieuwegein, and Houten, as well as regional and national authorities.

Next to the use case of the A12 zone, another use case has been conducted in light of this study: Stadshavens Rotterdam. This use case, however, turned out to be less relevant and is therefore not described in the main report. The main finding from this use case is that the model is not developed to predict seaport areas. Although the impact of the spatial planning scenario considered in the use case was predicted, the comparison between the base dataset and the explanatory model differed $\approx \frac{2}{3}^{th}$ trips (≈20000). Therefore another use case has been found which will be described hereafter. The use case regarding Stadshavens Rotterdam can be found in Appendix F.

6.2.1. Background

To ensure suitable and affordable housing in the Utrecht region, the construction of new housing is very important. The A12 zone is the area on both sides of the A12 motorway between junction Oudenrijn and Lunetten. This area has been designated as a site with many opportunities for new housing, businesses, and nature and recreational areas after 2030 (Gemeente Nieuwegein, n.d.). An overview of the A12 zone project, the areas of housing and employment increase and NRM zones corresponding to the project area are shown in Figure 6.1.

(a) Increase of housing (b) Increase of employment (c) Corresponding NRM zones

Figure 6.1: Overview of the A12 zone and the corresponding NRM zones

In total, 15250 new houses are planned to be built and 9300 new jobs are realised. These are allocated to the different development areas. What increase in housing and employment is realised per development area is depicted in Table 6.1. The purple areas correspond to the ones found in Figure 6.1. Moreover, the corresponding NRM zone allocation is also shown in this table. This means that, for example, when one development area is 50/50 divided into two NRM zones, the increase of housing and employment per NRM zone is also 50/50 divided.

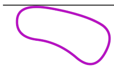
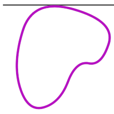
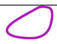
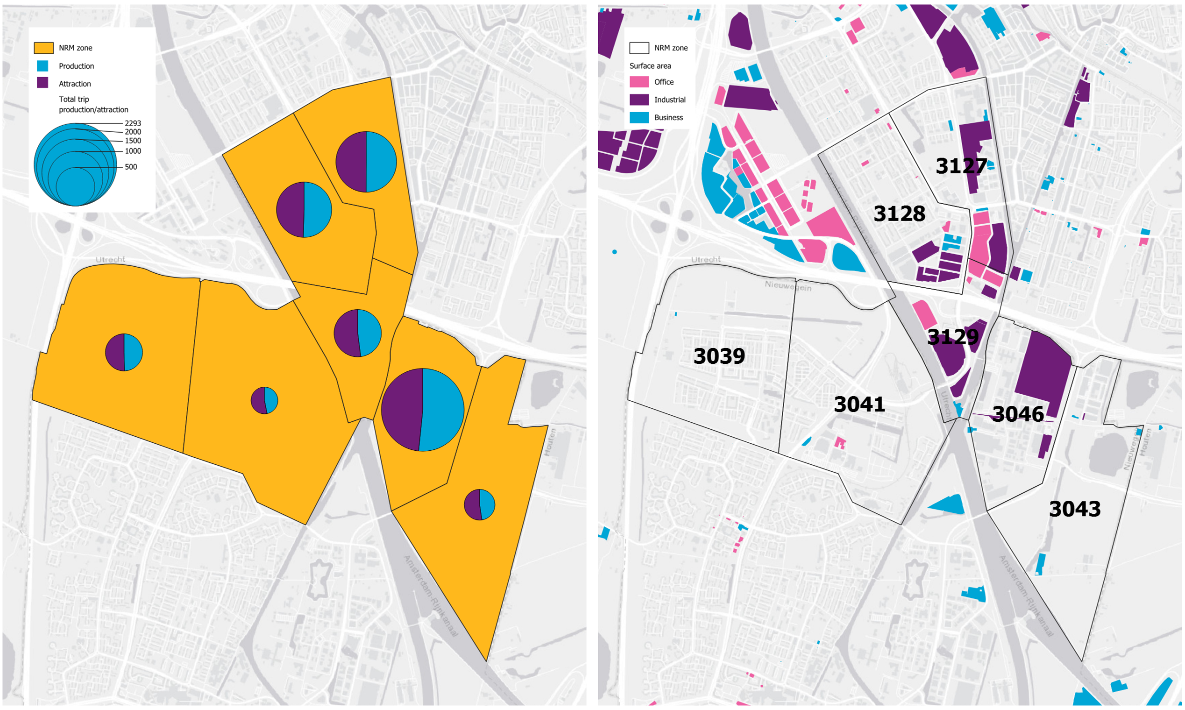
Zone	Name	Increase housing	Increase employment	Corresponding NRM zone allocation
	Galecopperzoom	5750	1300	3039 (33,3%), 3041 (67,7%)
	Tramremise	2100	4200	3041
	Westraven	3300	0	3129
	Liesbosch / Laagraven	750	650	3043 (16.7%), 3046 (83.3%)
	Woonboulevard	2200	2150	3128
	Merwedekanaalzone (subarea no. 6)	1150	1000	3127 (50%), 3129 (50%)
		15,250	9,300	

Table 6.1: Development plans of the A12 zone

6.2.2. Area Parameters

From the BasGoed freight tour data, which is used to predict the freight trips, the initial observed situation is determined regarding the trip production and attraction of the individual zones. These are distributed over the different NRM zones according to Figure 6.2a. Moreover, from the zonal plans the different land uses which are relevant to FTG per LS are depicted in Figure 6.2b.



(a) Current trip generation in A12 zone area (b) Current surface area distribution in A12 zone area

Figure 6.2: A12 zone area visualized using the model base dataset

De total FTP and FTA of the different zones are the combination of different trips per Logistic Segment. How the distribution of segments over the total amount of production and attraction per zone is defined is shown in Table 6.2. The table shows that LS 1 is most prominently present in the zones with ≈45% of all the trips. This segment contains general cargo, which could explain the size of trips containing this segment; a lot of not-specific cargo coming into and going out of the area for offices or industries. Moreover, the total production and attraction are both the same size, meaning all the incoming trips also leave the area.

Production (trips/workweek)									
NRM zone	**LS 0**	**LS 1**	**LS 2**	**LS 3**	**LS 4**	**LS 5**	**LS 6**	**LS 7**	**Total**
3039	25	101	21	21	29	20	1	11	229
3041	16	39	11	8	20	12	3	4	113
3043	10	75	15	13	17	14	1	6	151
3046	140	518	61	68	150	150	22	75	1184
3127	63	264	51	27	69	119	13	12	618
3128	67	218	42	42	70	38	3	37	517
3129	33	222	26	15	25	21	4	14	360
Total	354	1437	227	194	380	374	47	159	3172
Attraction (trips/workweek)									
NRM zone	**LS 0**	**LS 1**	**LS 2**	**LS 3**	**LS 4**	**LS 5**	**LS 6**	**LS 7**	**Total**
3039	25	104	22	20	30	21	1	10	233
3041	18	48	13	9	21	12	3	4	128
3043	10	83	13	16	19	14	1	10	166
3046	131	484	63	64	144	133	19	71	1109
3127	62	262	51	23	69	121	13	15	616
3128	60	219	47	41	71	35	1	32	506
3129	34	236	29	17	29	24	5	16	390
Total	340	1436	238	190	383	360	43	158	3148

Table 6.2: Current FTG of the A12 zone per LS

Considering the explaining factors found in the results (section 5.3), the explaining factors for the prediction of logistic movements are the surface area of distribution centers, employment in the agricultural, industrial, retail, services and others sector, the presence of a rail or inland waterway terminal, the urban density level of 2, 3, 4, 5 and 6 and the surface area of business, industrial and offices in the area. From the base dataset, the BasGoed freight tour data, CBS and zonal data, the parameters of the corresponding NRM zones for these variables are presented. This is shown in Table 6.3.

NRM zone	**3039**	**3041**	**3043**	**3046**	**3127**	**3128**	**3129**	**Total**
Surface area distribution centers (m^2)	0	0	0	49,029	30,150	0	0	79,179
Employment agricultural	7	0	14	1	0	0	0	22
Employment industrial	65	34	22	1,400	203	307	40	2,071
Employment retail	134	21	69	346	79	923	1	1,573
Employment services	188	329	13	2,750	4,632	548	764	9,224
Employment other	375	769	404	3,244	4,280	1,418	427	10,917
Terminal rail	0	0	0	0	0	0	0	0
Terminal inland waterway	0	0	0	0	1	0	1	2
Urban density level	6	6	6	6	6	6	6	7x6
Surface area business (m^2)	577	2,739	13,294	1,635	15,003	15038	5,816	54,103
Surface area industrial (m^2)	0	0	701	23,7964	146,407	77424	132,649	595,145
Surface area office (m^2)	0	5,903	0	0	59,415	7,708	62,584	135,610

Table 6.3: Current situation A12 zone

6.2.3. Increase Estimation

The project specifications, 15250 new houses and 9300 new jobs, are transformed into a change of parameters for each corresponding NRM zone. Regarding the 15250 new houses, all 7 NRM zones already have the highest urban density level. All other explaining factors for FTG are not relevant for housing, therefore these are not taken into account in this use case. However, the implications of 9300 new jobs are taken into account and resulted in an increase of employment per zone and an increase of land use. Since distribution centers, business, industrial and offices all have a relation with employment, an increase of surface area per NRM zone is considered and is calculated using Equation 6.1.

$$\text{Increase surface area}_i = \text{Increase employment}_i \cdot \frac{\text{Current surface area}_i}{\text{Current employment}_i} \tag{6.1}$$

where:

- Increase surface area$_i$ is the increase in total surface area of distribution centers, business, industrial and offices for zone i.
- Increase employment$_i$ is the increase in total employment of agricultural, industrial, retail, services and other for zone i.
- Current surface area$_i$ is the total surface area of distribution centers, business, industrial and offices according to the base model in zone i.
- Current employment$_i$ is the total employment of agricultural, industrial, retail, services and other according to the base model in zone i.

This formula shows that the increase in surface area is related to the increase in employment in that same zone. The increase in employment itself is defined by the increase of the development area (see Table 6.1) and divided over the NRM zones (shown in the same table). The resulting increase of surface area is shown in Table 6.4.

NRM zone	Increase housing	Increase employment	Increase surface area
3039	1,917	433	$433 \cdot \frac{577}{769} = 325\,m^2$
3041	5,933	5,067	$5,067 \cdot \frac{8,642}{1,153} = 37,977\,m^2$
3043	125	108	$108 \cdot \frac{13,995}{522} = 2,896\,m^2$
3046	625	542	$542 \cdot \frac{288,628}{7,741} = 20,209\,m^2$
3127	575	500	$500 \cdot \frac{250,975}{9,194} = 13,649\,m^2$
3128	2,200	2,150	$2,150 \cdot \frac{100,170}{3,196} = 67,386\,m^2$
3129	3,875	500	$500 \cdot \frac{201,050}{1,232} = 81,595\,m^2$
Total	15,250	9,300	224,036 m^2

Table 6.4: Increase of employment and surface area per NRM zone

Both the increase in surface area and employment are finally added to the designated variable (i.e. surface area business) according to the total weight of the employment or surface area in that zone. The results are shown in Table 6.5. Due to an increase of the surface area of distribution centers, the biggest increase in trips is expected to be in zone 3128.

NRM zone	3039	3041	3043	3046	3127	3128	3129	Total
Surface area distribution centers (m^2)	0	0	0	3433	1,640	5,073	0	0
Employment agricultural	4	0	3	0	0	7	0	0
Employment industrial	37	149	5	98	11	522	207	16
Employment retail	75	92	14	24	4	832	621	0
Employment services	106	1446	3	193	252	369	310	2,678
Employment other	211	3379	84	227	233	954	173	5,261
Surface area business (m^2)	325	12036	2751	114	816	10,116	2361	28,518
Surface area industrial (m^2)	0	0	145	16,661	7,962	52,085	53,835	130,688
Surface area office (m^2)	0	25,941	0	0	3,231	5,185	25,399	59,757

Table 6.5: Expected increase of employment and surface area parameters

6.2.4. Model Application

After the initial parameters and the expected parameters have been defined, the model and its explaining factors are applied to the study area. For this, two different tables have been generated which show a brief overview of the production, attraction and their explaining factors per LS for FTG and FG, in Table 6.6 and Table 6.7 respectively, the results per NST/R goods type can be found in Appendix E.

Table 6.6: Coefficients for explanatory variables of Freight Trip Generation per LS

Trips/Workweek/Zone *N = 5152*		Intercept	Surface area distribution centers	Employment: Agricultural	Employment: Industrial	Employment: Retail	Employment: Services	Employment: Others	Terminal: Rail	Terminal: Inland waterway	Urban density: 2	Urban density: 3	Urban density: 4	Urban density: 5	Urban density: 6	Surface area: Business	Surface area: Industrial	Surface area: Office
	Unit	-	*km²*	*x1000*	*x1000*	*x1000*	*x1000*	*x1000*	-	-	-	-	-	-	-	*km²*	*km²*	*km²*
Production	Food (general cargo)	12	230		105	7				9			6	8		110		
Production	Miscellaneous (general cargo)	18	431	419	261					43			30	45	50	593	62	
Production	Conditioned transport	-1	175	329	40					11		7	11	15	15	141	10	
Production	Facility logistics	3	98		35		1		23					3	3	37		
Production	Construction logistics	12	125		91							5	6	8	6	40		107
Production	Waste	2	924		48		2		59	8			4	6	6	62	11	
Production	Parcel (consolidated flows between sorting centres)	0	104		6		0		6	1				0	1	6	1	
Production	Hazardous materials	6			58								5	4		65		
Attraction	Food (general cargo)	14	365		81				37	8			8	12	7	125		
Attraction	Miscellaneous (general cargo)	20	957	288	196			12	248	37		20	36	54	52	608	44	
Attraction	Conditioned transport	3	228	227	33			2	35	10		7	10	15	13	132	5	
Attraction	Facility logistics	2	139		25		1		27	1	2	3	3	6	5	41		
Attraction	Construction logistics	14	217		74							5	6	10	8	48		106
Attraction	Waste	4	915		37		2		59	9			4	7	7	66	13	
Attraction	Parcel (consolidated flows between sorting centres)	1	105		5		0		6	1			0	1	1	6	1	
Attraction	Hazardous materials	8			44				24				5	4		66		

Table 6.7: Coefficients for explanatory variables of Freight Generation per LS

Trips/Workweek/Zone *N = 5152*		Intercept	Population	Surface area distribution centers	Employment: Agricultural	Employment: Industrial	Employment: Retail	Employment: Services	Employment: Others	Terminal: Container	Terminal: Rail	Terminal: Inland waterway	Urban density: 2	Urban density: 3	Urban density: 4	Urban density: 5	Urban density: 6	Surface area: Business	Surface area: Industrial	Surface area: Retail	Surface area: Office
	Unit	-	*x1000*	*km²*	*x1000*	*x1000*	*x1000*	*x1000*	*x1000*	-	-	-	-	-	-	-	-	*km²*	*km²*	*km²*	*km²*
Production	Food (general cargo)	133				1527						98				68		1044		2855	
Production	Miscellaneous (general cargo)	-49			5049	3539		127				775			247	386	454	5758	432		
Production	Conditioned transport	-31			3962	610						142		91	114	155	142	1432	85		
Production	Facility logistics	16				369		13			171					25	26	332			515
Production	Construction logistics	119				1207		26								53	42	373		1333	
Production	Waste	2		5570		503		19			389	58			23	41	51	498	54		776
Production	Parcel (consolidated flows between sorting centres)	1		574		61	4	4			39	7				4	4	48		111	
Production	Hazardous materials	45				712												598			
Attraction	Food (general cargo)	225		3139		653					253	91		64	128	164	89	1424	61		
Attraction	Miscellaneous (general cargo)	477		9290	2512	1000			123		2286	142	270	328	500	691	600	6259	433		
Attraction	Conditioned transport	109		1672	1802	273			25		323	96		73	107	155	140	1429	67		
Attraction	Facility logistics	28		722		149			8		213	11	27	28	36	60	58	410	19		
Attraction	Construction logistics	197	5	2171		571					156			61	73	121	93	558			966
Attraction	Waste	30		6098		183			10		511	63	46	36	63	84	79	578	131		
Attraction	Parcel (consolidated flows between sorting centres)	10		634		24		2		15	58	7				5	4	50	13		
Attraction	Hazardous materials	119				272					277				49	45		671			

As a reference, the FTG is calculated using the input parameters from Table 6.3 and the explaining factors found in Table 6.6. This results in a total production of 2759 trips/workweek and a total attraction of 2809 trips/workweek of the study area. The prediction from the model shows an error of 413 and 339 trips/workweek for production and attraction, respectively. This shows a margin error of 10-13% and is within the expectancy of the model outcome, considering the model R^2 is 0.72 at best, meaning that 72% of the dataset is explained correctly and thus 28% is not.

6.2.5. Results

The expected increase of the parameters is applied to the current parameters. Those are consequently applied to the explaining coefficients which results in a future freight trip scenario. The impact is calculated using the additive growth method, meaning that the predicted impact is added to the base scenario (base dataset), defined by the following two equations: the predicted freight trip impact of the scenario is calculated using Equation 6.3 and the total predicted future trips is calculated using Equation 6.3.

$$\text{Predicted freight trip impact}_i = \text{Model outcome future}_i - \text{Model outcome current}_i \tag{6.2}$$

where:

- Predicted freight trip impact is the predicted increase of freight trips due to the spatial scenario in zone i.
- Model outcome future is the model outcome of the explanatory model considering the future scenario in zone i.
- Model outcome current is the model outcome of the explanatory model considering the current scenario in zone i.

$$\text{Predicted future freight trips}_i = \text{Observed trips}_i + \text{Predicted freight trip impact}_i \tag{6.3}$$

where:

- Predicted future freight trips are the predicted total freight trips after applying the spatial scenario in zone i.
- Observed trips are the freight trips currently observed in zone i according to the BasGoed freight tour data.
- Predicted freight trip impact is the predicted increase of freight trips due to the spatial scenario in zone i.

The application of the explaining coefficients resulted in a predicted freight trip production increase of 414 trips/workweek and a predicted attraction increase of 408 trips/workweek of the study area (also see Table 6.8), the total predicted future production/attraction trips are 3586 trips/workweek and 3556 trips/workweek, respectively. Considering the initial production/attraction trips of 3172 and 3148, respectively, this is a significant increase of ≈13% trips in the study area.

Production (trips/workweek)								
NRM zone	**3039**	**3041**	**3043**	**3046**	**3127**	**3128**	**3129**	**Total**
Observed trips	229	113	151	1184	618	517	360	3172
Model outcome current	183	160	173	1172	449	363	258	2759
Model outcome future	211	277	181	1244	463	518	280	3174
Predicted freight trip impact	28	117	8	73	13	154	21	414
Predicted future freight trips	257	230	159	1257	631	671	381	3586
Attraction (trips/workweek)								
NRM zone	**3039**	**3041**	**3043**	**3046**	**3127**	**3128**	**3129**	**Total**
Observed trips	233	128	166	1109	616	506	390	3148
Model outcome current	201	191	197	1069	521	355	274	2809
Model outcome future	225	335	205	1133	537	488	294	3217
Predicted freight trip impact	24	144	8	64	16	132	20	408
Predicted future freight trips	257	272	174	1173	632	638	410	3556

Table 6.8: Results of the use case

The highest increase, being a production/attraction increase of 103/112%, respectively, is found in zone 3041. This zone also meets the highest increase of employment and has little amount of trips in the current situation. The amount of trips in that zone is more than doubled due to the spatial planning scenario (≈121 to ≈251 trips/workweek). Zone 3128 has approximately the same increase of trips, whereas the increase in relation to the observed trips is only 30%. A visualized overview of the increase in production/attraction per NRM zone is shown in Figure 6.3.

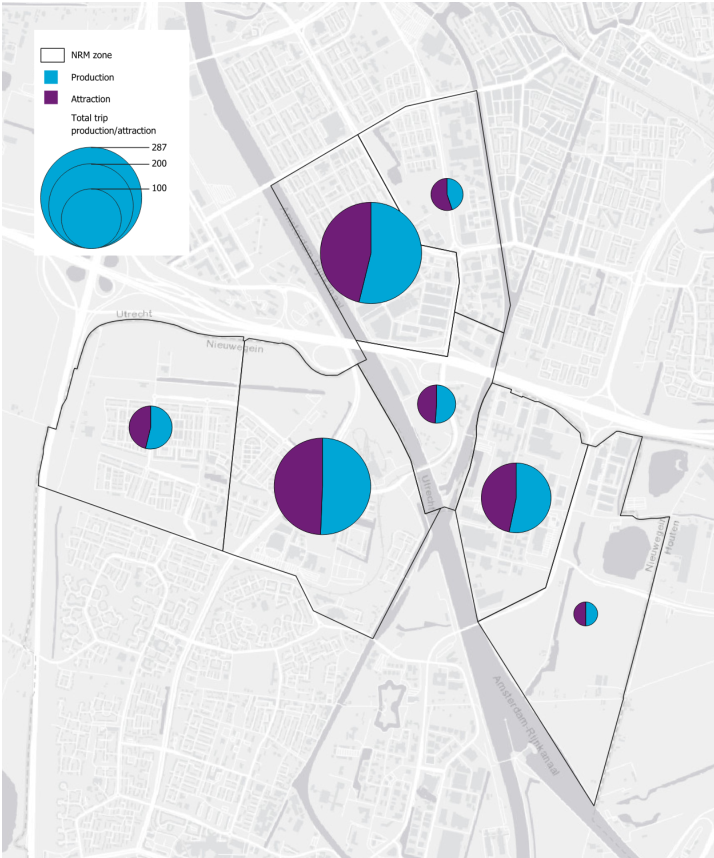


Figure 6.3: Trip increase in study area

Sub-question 2b: What is the impact in terms of logistics movements?

Although some assumptions have been made in this use case, the impact of a spatial planning scenario is roughly estimated using the explanatory model. The results show movements in absolute form, but the impact can also be put in perspective; roughly 13% increase regarding the current number of movements. In this use case, the impact shows a total increase of 414 outgoing and 408 incoming trips per workweek. However, as the impact of this use case is depicted in a total number, the different logistic segments are not taken into account.

6.3. Conclusion of Application

This section concludes the application with the answering of two sub-questions.

Sub-question 2c: What policy implications can be learned from this use case?

With the knowledge of the increase of logistics movements in a to-be-developed area, new implications for policymakers can be thought of. After applying the explanatory model, and defining how severe the impact of the spatial planning scenario will be in terms of movements, questions like *How will these movements go through the network and what will their entry and exit points be?*, *Will the increase of movements fit on the current road infrastructure?*, *How can the movements to this area be regulated?*, *What will the optimal solution for supplying the area be?* and *Should another modality be considered for movements towards this area?* could arise. Some of these implications can be addressed with the knowledge from this model: the quantification of logistics.

Sub-question 2: What is the impact of applying the model to a use case?

The impact of applying this explanatory model for logistics movements to a use case, created insights for the spatial planning scenario applied. Although not every spatial planning scenario is feasible with this model and areas with a large trip generation like seaports are bound to have a large explanatory error, in scenarios like urban growth, economic development and land use the increase of movements can create new implications for policymakers to address before the implementation phase of the scenario. Moreover, the results of this application can be input for other models, like traffic flow models.

7 Discussion

The goal of this research is to provide a comprehensive understanding of the factors influencing freight (trip) generation in urban areas in the Netherlands, which is essential for effective urban logistics management and planning. The insights of this study have significant implications for both scientific research and practical applications in urban planners and policymakers.

The linear regression model developed in this research, using synthetic data from the BasGoed freight tour data, shows robust statistical validity. Explaining factors such as employment and surface area were identified as the most significant predictors of freight (trip) generation in all Logistic Segments. The validation of the model through statistical tests and comparisons with existing literature confirms its reliability and accuracy in predicting logistics movements. It was found by Sahu and Pani that employment is more suitable for representing FP, while surface area explains FA better. The results from this model confirm this observation. Although in both FP and FA are employment as surface area significant explaining factors. In FP, employment has a larger value of the standardized coefficient and thus a bigger magnitude in effect (more about that in section 5.3) whereas in FA, the surface area has a larger value of the standardized coefficient. This is in line with the observation from the literature since a larger standardized value indicated a stronger relationship between the predictor variable and the outcome variable. Pani et al. stated that employment is invariably considered the most preferred causal variable in FG/FTG models. 75 out of 76 models show that any kind of employment is a significant explaining factor to the FTG or FG. Only in the FP for NST/R goods type 7 (fertilizers) explaining factors in employment lack. This shows a very large preferred causal link with employment and confirms the observation from the literature.

The research provides several useful insights for policymakers. The significant impact of employment in various sectors on freight (trip) generation suggests that urban planners should consider employment when designing logistics-friendly urban environments. Additionally, the presence of a rail terminal significantly influences logistics attraction, indicating that enhancing these facilities could improve freight efficiency and reduce road congestion.

One of the evident findings is the inverse relationship between urban density and logistics trip attraction. Higher urban density appears to decrease logistics trip attraction, which could be explained by logistics bundling, transhipment to other modes which are not included in the model or the lack of business and industry in denser areas. It could also suggest that densely populated urban areas might have more efficient logistics networks that consolidate shipments, thereby reducing the number of trips required. Further research is necessary to explore this phenomenon.

The practical application of this model is conducted through a use case, demonstrating how changes in surface area and employment can predict logistics movements under different spatial planning scenarios. For instance, scenarios involving economic development that stimulate employment can be analyzed using the model to forecast the corresponding increase in freight movements. This can help planners optimize land use and infrastructure development to accommodate future logistics demands effectively. Moreover, a lot of questions like *How will these movements go through the network and what will their entry and exit points be?*, *Will the increase of movements fit on the current road infrastructure?*, *How can the movements to this area be regulated?*, *What will the optimal solution for supplying the area be?* and *Should another modality be considered for movements towards this area?* could arise. These can be addressed with the insights of this model.

The model of this study includes the following limitations:

- The use of the one-sided t-test could be a limiting factor. A one-sided t-test focuses on an increase or decrease of the explaining factor. In this case, it led to an increase since only positive t-values were allowed in the explaining factor selection. A two-sided t-test would allow a change in trips or weight (positive and negative explaining factors), and not only an increase or decrease. This would enhance the predictive power but would also affect the plausibility since it is unknown what the explanation for the factors that decrease the number of trips would be. When tested on a small scale, it showed that there were more explaining factors significant when using a two-sided t-test (so also negative ones) instead of a one-sided t-test. However, the R^2 did barely change and thus for this model only positive explaining factors are considered.
- As seen from the use case found in Appendix F, this model explaining factors does not cover all situations. Seaport areas, where a lot of freight is generated, are not suitable for this model. Also, the freight (trip) generation in big transhipment areas can not be explained correctly due to the vast amount of trips or weight generated by these areas. The model is based on all of the zones in the Netherlands and finds its best there, so the outliers like the seaports cannot be explained by it.
- Also, the model explanatory power of the model has its limitations in terms of accuracy. Since $0.67 \leq R^2 \leq 0.19$, not all data can be explained by the model. Therefore, some inconsistency in the results is expected.
- The positive trade-off between an independent and dependent variable, which occurs with endogeneity, is not covered in this model. For instance, the area of distribution centers is endogenous to trip generation. This means that the size of distribution centers is not only determined by external factors but also by the demand for transport services generated in that specific area. This means that an increase in trips could lead to an increase of distribution center surface area which will lead to an increase in trips and so forth. In this particular example, if there is high demand for transport services in a particular area, it may be interesting for distribution centres in that area to expand their area to meet that demand. This creates a positive trade-off between the surface area of distribution centers and trip generation in that area. The implication of that, in this model, is that it can lead to inaccuracies in forecasts and a lack of understanding of the actual dynamics between distribution centers and trip generation.

8 Conclusion

The conclusion of this research starts with the research goal: to gain more insight into the quantification of logistics by determining the explaining factors for logistics generation in the Netherlands and provide knowledge for spatial planners to more efficiently implement logistics into urban developments and know its impact. The first part of that goal is addressed in section 8.1, the second part is addressed in section 8.2 and finally this conclusion ends with answering the main research question in section 8.3.

8.1. Explaining Factors

To answer the goal of gaining more insight into the quantification of logistics by determining the explaining factors for logistics generation in the Netherlands, a study has been conducted where multiple datasets of the Netherlands were combined and analysed with the help of a linear regression analysis using OLS as a technique. This resulted in an empirically substantiated analysis which has lots of feasible applications (i.e. enhancing demand models by providing more accurate forecasts of demand) and results in a practical and usable table. The explaining factors from this model can predict the trip and weight production and attraction of logistics in terms of different Logistic Segments and different classes of NST/R.

The following explaining factors turned out to be significant - with at least a cumulative probability confidence level of 97.5% - for FTG and FG:

- Freight Trip Generation
  - The surface area of distribution centers
  - The employment in the agricultural, industrial, retail and services and other sector
  - The presence of rail or inland waterway terminals
  - The urban density level of 2, 3, 4, 5 and 6
  - The surface area of business, industrial and offices
- Freight Generation
  - The population
  - The surface area of distribution centers
  - The employment in the agricultural, industrial, retail and services and other sector
  - The presence of container, rail or inland waterway terminals
  - The urban density level of 2, 3, 4, 5 and 6

  - The surface area of business, industrial, retail and offices

Out of these factors, the surface area and employment had the biggest impact on freight (trip) generation. This is in line with the literature, employment in particular is found in many studies to be a significant predictor. Moreover, it is very plausible for the above predictors to explain freight: it is clear that business, industrial, offices and distribution centers establishments produce and attract freight, and most certainly it does with terminals. The increase in urban density, meaning the increase of possible employees and the increase of activity, is also very logical to generate freight (up to a certain level, 6 in particular). Therefore employment itself does generate freight too.

Furthermore, what has been noticed during the study:

- Age distribution does not affect the generation of freight. The explaining factor does not include the distribution of age but only focuses on the number of people in a zone: the population.
- Distribution centers in the LSs mainly cause freight weight to be attracted to a zone, the production of the weight is rather small. However, in terms of trips, both production and attraction are equally distributed (what goes in, also goes out).
- Rail terminals increase the attraction of trips and weight but do not cause (a lot of) production. This could mean that the transhipment of rail to truck is larger than truck to rail.
- An urban density level which is lower is more likely to attract freight than producing it.
- The attraction of both trips and weight decreases when the urban density level is increased from 5 to 6.

8.2. Impact Logistics

The impact on logistics by applying the prediction model estimated in this study can be of great improvement when considering a spatial planning scenario. Although not every spatial planning scenario can be used with this model, in scenarios like urban growth, economic development and land use the increase of movements in absolute numbers can create new implications for policymakers to address before the implementation phase of the scenario. Moreover, the results of this application can be input for other models, like traffic flow models. Next to logistics movements, the amount of weight created by spatial planning scenarios can be estimated. These weights can also lead to great new insights.

8.3. Main Research Question

Main research question: How can the logistics movements of urban environments in the Netherlands be predicted?

This study shows a model which predicts logistics in terms of freight trips and freight weight. The model results show highly significant explaining factors in the production and attraction of logistics, which have been seen in previous studies as well (Table 3.1). However, the model results from this study are based on a more extensive and detailed dataset with many more explaining factors. 12 out of the 26 explaining factors were found to be significant for Freight Trip Generation due to this study. Moreover, the different segments which are predicted with such precision also are a unique point of this study. In answering the main research question of how logistic movements can be predicted, the first part - on a scientific base - is by using the explaining factors shown and their corresponding coefficients. These factors show predictions in urban areas as in rural areas through the urban density level. The main explaining factors for logistic movement predictions are:

- The surface area of distribution centers
- The employment in the agricultural, industrial, retail and services and other sector
- The presence of rail or inland waterway terminals
- The urban density level of 2, 3, 4, 5 and 6
- The surface area of business, industrial and offices

Next to the significance of the explaining factors according to the model, the impact on logistics by applying the prediction model, and estimating absolute numbers in terms of movements, can be of great improvement when considering a spatial planning scenario. Although not every spatial planning scenario can be used with this model, in scenarios like urban growth, economic development and land use the increase of movements can create new implications for policymakers to address before the implementation phase of the scenario. Moreover, the results of this application can be input for other models, like traffic flow models.

9 Recommendations

This chapter presents recommendations based on the key findings of this study. The goal is to address the identified challenges and offer actionable steps for future work and practical applications in the quantification of FTG and FG. These recommendations are designed to benefit various stakeholders by providing feasible and impactful studies.

1. Validating using XML-data
   As shown in chapter 4, this research is based on synthetic data from the BasGoed freight tour module. However, what results would come from the model if the model was estimated on the original XML data which is used in the BasGoed model? This data is location-bounded and uses the original locations of freight production of attraction locations in the Netherlands. Would the results of the model then be the same? A recommendation would be to explore this in a new study.
2. The decrease with a higher urban density
   As seen in the results, an increase in urban density level (from 5 to 6) causes attraction to decrease. It would be a recommendation to explore this interesting observation. Would it be due to bundling logistics? Is there a transhipment happening between vehicles not addressed in the model?
3. Production/Attraction of households
   This model already predicts on a relatively low level; NRM zones. However, how would you model the production and attraction of households for example? This would give an extra dimension to the prediction of logistics when you can model the parcel deliveries in a given zone.
4. Attraction of construction logistics by office zoning
   As seen in the model results, construction logistics was attracted by the surface area of office zoning. Why is this happening? A recommendation for future research would be to explore this observation. Understanding these underlying factors could help policymakers to better manage and predict freight movements associated with construction activities.
5. Non-linearity
   In this study, all correlations are assumed to be linear: the amount of trips has a linear relation to the surface area of distribution centres, for example. However, it should be taken into consideration that some relations can be non-linear for example found by Sánchez-Díaz et al. for employment in industry sectors. Future research can investigate this non-linearity with this extensive dataset and look at what type of regression model fits the data even better, i.e. Poisson regression.

6. Explanatory variables for shipment size and vehicle type
   In light of the feasibility of the applications, the specific modelling in terms of vehicle types would enhance the way this model has an impact on urban environments. Therefore, a recommendation would be to explore the explanatory variables for vehicle types per segment to increase the practical usability for applications and policymakers. This can be done by the specific vehicle type which is incorporated in the BasGoed freight tour data. However, the shipment size, in combination with the weight transported with each trip, would even secure a more accurate prediction of vehicle trips and give more insights into the efficiency of the freight network. Vehicle type would then be the limiting factor in the amount of shipments transported per trip. Both aspects, vehicle type and shipment size, would be worthwhile for future work.
7. Liveability aspects
   While the previous recommendation is focused on explaining factors for vehicle type and shipment size, subsequently the explaining and predicting factors for vehicle types would have a large impact on the knowledge of a liveable environment. With the knowledge for explaining vehicle types, aspects like emissions, traffic conditions, congestion, accessibility and safety can also be taken into account since the amount and the type of vehicle influence these. Therefore, next to the explaining factors for vehicle type, a subsequent study could be conducted on how this will affect the urban environment in terms of the mentioned aspects. One study per aspect.
8. Apply this method to a dataset of vans or other missing vehicles
   If in the future the data availability for vans or other missing vehicles has increased, the same method used in this research could be applied to model the prediction of those. This is highly recommended and would have a great impact on knowledge of city logistic policies. The urge to know what happens in city centers is still substantially large and knowing the number of vans or other vehicles going into the city and having that knowledge while developing urban environments is needed.
9. To prevent endogeneity, a causality analysis could be conducted in future work. Different variables should then carefully be examined for causal relationships between variables and experiments or observational studies should be considered to modify endogeneity. Moreover, in the case of endogeneity, this work may need to be improved by using instrumental variables to correct the bias. Instrumental variables are external variables that are correlated with the independent variable of interest but not with the variable which has a positive trade-off with it.
10. Finally, this study gives great knowledge of the impact and quantification of logistics. However, the deliverable is a very broad and scientific report. Future work could be to find a way of implementing this knowledge, and perhaps knowledge from other recommended studies above, into the practitioners and policymakers. That is where this study can have a large impact and a lot of benefits.

References

‘Smart zone’ wordt mogelijke oplossing voor laad- en losprobleem in Utrechtse Twijnstraat. (2022, May). Retrieved May 11, 2024, from https://www.duic.nl/algemeen/smart-zone-wordt-mogelijke-oplossing-voor-laad-en-losprobleem-in-utrechtse-twijnstraat/

Alho, A. R., & Silva, J. d. A. e. (2014). Freight-Trip Generation Model: Predicting Urban Freight Weekly Parking Demand from Retail Establishment Characteristics [Publisher: SAGE Publications Inc]. *Transportation Research Record*, *2411*(1), 45–54. https://doi.org/10.3141/2411-06

Asuncion, J. S. (2014). The Geographic Adaptive Potential of Freight Transportation and Production System in the Context of Fuel and Emission Constraints [Publisher: University of Canterbury. Department of Mechanical Engineering]. Retrieved January 11, 2024, from http://hdl.handle.net/10092/9230

BasGoed. (n.d.). Retrieved December 4, 2023, from https://www.basgoed.nl/

BasGoed Functionele Documentatie. (2023, December).

Bastida, C., & Holguín-Veras, J. (2009). Freight Generation Models: Comparative Analysis of Regression Models and Multiple Classification Analysis [Publisher: SAGE Publications Inc]. *Transportation Research Record*, *2097*(1), 51–61. https://doi.org/10.3141/2097-07

Beagan, D., Fischer, M., & Kuppam, A. (2007). Quick Response Freight Manual II. Retrieved November 29, 2023, from https://www.semanticscholar.org/paper/Quick-Response-Freight-Manual-II-Beagan-Fischer/18f8b61e9d45dc17f08713b2135aa20b02668262

Bestemmingsplan Bedrijventerreinen: Regels. (n.d.). Retrieved April 25, 2024, from https://www.planviewer.nl/imro/files/NL.IMRO.0867.bpWWbdrijvnterrein-va01/r_NL.IMRO.0867.bpWWbdrijvnterrein-va01.html

Boer, A. M. (n.d.). Distribution Center facilities in South-Holland.

Campbell, S., Sanchez-Diaz, I., Lawson, C. T., Jaller, M., & Holguin-Veras, J. (2012). Comparison Between Industrial Classification Systems in Freight Trip Generation Modeling [Number: 12-3825]. Retrieved January 11, 2024, from https://trid.trb.org/View/1130510

Centraal Bureau voor de Statistiek. (n.d.). Microdata: Zelf onderzoek doen [Last Modified: 24-10-2023T09:29:31]. Retrieved December 4, 2023, from https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen

Cheah, L., Mepparambath, R. M., & Ricart Surribas, G. M. (2021). Freight trips generated at retail malls in dense urban areas. *Transportation Research Part A: Policy and Practice*, *145*, 118–131. https://doi.org/10.1016/j.tra.2021.01.015

Couzy, M. (2016, August). Gemeente start offensief tegen foutparkeren busjes. Retrieved June 13, 2023, from https://www.parool.nl/nieuws/gemeente-start-offensief-tegen-foutparkeren-busjes~b0d95775/

CROW & Topsector Logistiek. (2023, May). *Handreiking Duurzame gebiedsontwikkeling en logistiek* (tech. rep.). https://www-crow-nl.tudelft.idm.oclc.org/downloads/pdf/mobiliteit/handreiking-duurzame-gebiedsontwikkeling-en-logist.aspx?ext=.pdf

Data- en Kennishub Gezond Stedelijk Leven (DKH GSL). (n.d.). Ambitie. Retrieved June 13, 2023, from https://gezondstedelijklevenhub.nlover-ons/ambitie.html

den Boer, E., Kok, R., Ploos van Amstel, W., Quak, H. J., & Wagter, H. (2017). Outlook City Logistics 2017. Retrieved March 31, 2023, from https://repository.tno.nl/islandora/object/uuid%3Ac1e44ebd-833d-4515-9760-f4a47eddf53a

Dhulipala, S., & Patil, G. R. (2023). Regional freight generation and spatial interactions in developing regions using secondary data. *Transportation*, *50*(3), 773–810. https://doi.org/10.1007/s11116-021-10261-w

Dolati Neghabadi, P., Evrard Samuel, K., & Espinouse, M.-L. (2019). Systematic literature review on city logistics: Overview, classification and analysis [Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00207543.2018.1489153]. *International Journal of Production Research*, *57*(3), 865–887. https://doi.org/10.1080/00207543.2018.1489153

Gemeente Amsterdam. (2023, June). Agenda Amsterdam Autoluw. Retrieved June 22, 2023, from https://www.amsterdam.nl/verkeer-vervoer/agenda-amsterdam-autoluw/

Gemeente Delft. (n.d.). Autoluwe binnenstad. Retrieved June 22, 2023, from https://www.delft.nl/autoluwe-binnenstad

Gemeente Nieuwegein. (n.d.). De A12-zone. Retrieved May 13, 2024, from https://www.ikbennieuwegein.nl/projecten/de+a12-zone/default.aspx

Gemeente Utrecht. (n.d.-a). Beurskwartier: Bouw stadswijk. Retrieved January 25, 2024, from https://www.utrecht.nl/wonen-en-leven/bouwprojecten-en-stedelijke-ontwikkeling/bouwprojecten/beurskwartier-bouw-stadswijk/

Gemeente Utrecht. (n.d.-b). Merwede. Retrieved January 25, 2024, from https://www.utrecht.nl/wonen-en-leven/bouwprojecten-en-stedelijke-ontwikkeling/bouwprojecten/merwedekanaalzone/projecten-in-de-merwedekanaalzone/merwede/

Giuliano, G., Gordon, P., Pan, Q., Park, J., & Wang, L. (2010). Estimating Freight Flows for Metropolitan Area Highway Networks Using Secondary Data Sources. *Networks and Spatial Economics*, *10*(1), 73–91. https://doi.org/10.1007/s11067-007-9024-9

Gonzalez-Calderon, C. A., Holguín-Veras, J., Amaya, J., Sánchez-Díaz, I., & Sarmiento, I. (2021). Generalized noortman and van es' empty trips model. *Transportation Research Part A: Policy and Practice*, *145*, 260–268. https://doi.org/10.1016/j.tra.2021.01.005

Gonzalez-Feliu, J., Palacios-Argüello, L., & Suarez-Nuñez, C. (2020). Links between freight trip generation rates, accessibility and socio-demographic variables in urban zones [Publisher: Index Copernicus]. *Archives of Transport*, *53*(1), 7–20. https://doi.org/10.5604/01.3001.0014.1738

Grebe, S., Jong, G. d., Bok, M. d., Houwe, P. v., & Borremans, D. (2016). THE STRATEGIC FLEMISH FREIGHT MODEL AT THE INTERSECTION OF POLICY ISSUES AND THE AVAILABLE DATA.

Ha, D.-H., & Combes, F. (2016). Building a Model of Freight Generation with a Commodity Flow Survey. In U. Clausen, H. Friedrich, C. Thaller, & C. Geiger (Eds.), *Commercial Transport* (pp. 23–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-21266-1_2

Handleiding Webviewer BasGoed. (2022). https://www.basgoed.nl/assets/handleiding_webviewer_v2.0.pdf

Holguín-Veras, J., Jaller, M., Destro, L., Ban, X. (, Lawson, C., & Levinson, H. S. (2011). Freight Generation, Freight Trip Generation, and Perils of Using Constant Trip Rates [Publisher: SAGE Publications Inc]. *Transportation Research Record*, *2224*(1), 68–81. https://doi.org/10.3141/2224-09

Holguín-Veras, J., Jaller, M., Sánchez-Díaz, I., Campbell, S., & Lawson, C. T. (2014, January). 3 - Freight Generation and Freight Trip Generation Models. In L. Tavasszy & G. de Jong (Eds.), *Modelling Freight Transport* (pp. 43–63). Elsevier. https://doi.org/10.1016/B978-0-12-410400-6.00003-3

Holguín-Veras, J., & Patil, G. (2007). Integrated Origin-Destination Synthesis Model for Freight with Commodity-Based and Empty Trip Models. *Transportation Research Record*, *2008*, 60–66. https://doi.org/10.3141/2008-08

Iding, M. H. E., Meester, W. J., & Tavasszy, L. (2002). Freight trip generation by firms. Retrieved November 15, 2023, from https://www.econstor.eu/handle/10419/115840

Institute of Transportation Engineers. (2008). *Trip Generation Manual* (tech. rep. No. 8th Edition).

Jaller, M., Sanchez-Diaz, I., Holguín-Veras, J., & Lawson, C. T. (2014). Area Based Freight Trip Generation Models [Number: 14-4908]. Retrieved November 16, 2023, from https://trid.trb.org/view/1289664

Klimaatakkoord. (2019, June). https://www.klimaatakkoord.nl/binaries/klimaatakkoord/documenten/publicaties/2019/06/28/klimaatakkoord/klimaatakkoord.pdf

Kock, R. (2023, January). 'Er is niets beter voor stadslogistiek dan minder voertuigen'. Retrieved June 22, 2023, from https://www.bnr.nl/nieuws/mobiliteit/10502651/er-is-niets-beter-voor-stadslogistiek-dan-minder-voertuigen

Lawson, C. T., Holguín-Veras, J., Sánchez-Díaz, I., Jaller, M., Campbell, S., & Powers, E. L. (2012). Estimated Generation of Freight Trips Based on Land Use [Publisher: SAGE Publications Inc]. *Transportation Research Record*, *2269*(1), 65–72. https://doi.org/10.3141/2269-08

Logistieke hubs voor emissievrije stedelijke distributie. (2023, September). Retrieved September 25, 2023, from https://www.tno.nl/nl/newsroom/insights/2023/05/hubs-emissievrije-stedelijke-distributie/

Marcucci, E., Gatta, V., & Le Pira, M. (2023). *Handbook on City Logistics and Urban Freight: 0*. Edward Elgar Publishing. Retrieved December 4, 2023, from https://books.google.nl/books?hl=nl&lr=&id=8gPEEAAAQBAJ&oi=fnd&pg=PR1&dq=Handbook+on+City+Logistics+and+Urban+Freight&ots=hGXRGTIuOn&sig=zcygrEgVUNW1KdjGE9tmyGtjN2g

Matthijs Otten & Peter Scholten. (2020). *Outlook Hinterland and Continental Freight 2020* (tech. rep.). Retrieved May 20, 2024, from https://cedelft.eu/publications/outlook-hinterland-and-continental-freight-2020/

Merwede. (2022, February). Straks - Een stadswijk als geen ander. Retrieved June 22, 2023, from https://merwede.nl/straks/

Messelink, R. (2020, December). Veel letsel én doden in de binnenstad door roekeloos rijgedrag in bestelbusjes: 'Voer apart rijbewijs in'. Retrieved June 13, 2023, from https://eenvandaag.avrotros.nl/item/roekeloos-rijgedrag-in-bestelbusjes-zorgt-voor-veel-letsel-en-doden-in-de-binnenstad-voer-apart-rijbewijs-in/

Mohammed, R. A., Nadi, A., Tavasszy, L., & de Bok, M. (2023). Data Fusion Approach to Identify Distribution Chain Segments in Freight Shipment Databases. *Transportation Research Record: Journal of the Transportation Research Board*, *2677*, 036119812211470. https://doi.org/10.1177/03611981221147049

Mommens, K., Lier, T. v., & Macharis, C. (2017). Freight demand generation on commodity and loading unit level [Number: 1]. *European Journal of Transport and Infrastructure Research*, *17*(1). https://doi.org/10.18757/ejtir.2017.17.1.3179

Nederlandse Omroep Stichting. (2021, February). Grote bouwcoalitie presenteert actieplan voor 1 miljoen woningen in tien jaar. Retrieved July 4, 2023, from https://nos.nl/artikel/2369043-grote-bouwcoalitie-presenteert-actieplan-voor-1-miljoen-woningen-in-tien-jaar

Oliveira, L., Araújo, G., Bertoncini, B., Pedrosa, C., & Silva, F. (2022). Modelling Freight Trip Generation Based on Deliveries for Brazilian Municipalities. *Sustainability (Switzerland)*, *14*(16). https://doi.org/10.3390/su141610300

Otten, M., Meerwaldt, H., & den Boer, E. (2016, July). De omvang van stadslogistiek [Opdrachtgever: Topsector Logistiek.].

Overlast door bouw logistieke 'blokkendozen': 'Wegen niet ingesteld op zoveel vrachtwagens'. (2022, April). Retrieved May 11, 2024, from https://pointer.kro-ncrv.nl/overlast-door-bouw-logistieke-blokkendozen-wegen-niet-ingesteld-op-zoveel-vrachtwagens

Pani, A., Sahu, P., Patil, G., & Sarkar, A. (2018). Modelling urban freight generation: A case study of seven cities in Kerala, India. *Transport Policy*, *69*, 49–64. https://doi.org/10.1016/j.tranpol.2018.05.013

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830. Retrieved April 24, 2024, from http://jmlr.org/papers/v12/pedregosa11a.html

Ploos van Amstel, W. (2017, November). Verkeersveiligheid stadslogistiek: Tijd om aan de rem te trekken. Retrieved June 13, 2023, from https://www.logistiek.nl/159959/verkeersveiligheid-stadslogistiek-tijd-om-aan-de-rem-te-trekken

Rodoshi, N., Shahriyar, S., Al Noor, M., & Hoque, M. (2024). A REVIEW OF THE FREIGHT GENERATION MODELS AND DATA COLLECTION TECHNIQUES.

Sahu, P., & Pani, A. (2020). Freight generation and geographical effects: Modelling freight needs of establishments in developing economies and analyzing their geographical disparities. *Transportation*, *47*(6), 2873–2902. https://doi.org/10.1007/s11116-019-09995-5

Sanchez-Diaz, I. (2020). Assessing the magnitude of freight traffic generated by office deliveries. *Transportation Research Part A: Policy and Practice*, *142*, 279–289. https://doi.org/10.1016/j.tra.2020.11.003

Sánchez-Díaz, I. (2017). Modeling urban freight generation: A study of commercial establishments' freight needs. *Transportation Research Part A: Policy and Practice*, *102*, 3–17. https://doi.org/10.1016/j.tra.2016.06.035

Sánchez-Díaz, I., Holguín-Veras, J., & Wang, X. (2016). An exploratory analysis of spatial effects on freight trip attraction. *Transportation*, *43*(1), 177–196. https://doi.org/10.1007/s11116-014-9570-1

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*, *2010*.

Statistiek, C. B. v. d. (n.d.-a). Bedrijventerrein [Last Modified: 23-02-2022T17:52:37]. Retrieved April 25, 2024, from https://www.cbs.nl/nl-nl/onze-diensten/methoden/begrippen/bedrijventerrein

Statistiek, C. B. v. d. (n.d.-b). Werkgelegenheid [Last Modified: 23-02-2022T17:56:23]. Retrieved April 22, 2024, from https://www.cbs.nl/nl-nl/onze-diensten/methoden/begrippen/werkgelegenheid

Statistiek, C. B. v. d. (2021, March). Statistische gegevens per vierkant en postcode 2020-2019-2018 [Last Modified: 2021-03-05T14:30:00+01:00]. Retrieved January 28, 2024, from https://www.cbs.nl/nl-nl/longread/diversen/2021/statistische-gegevens-per-vierkant-en-postcode-2020-2019-2018

United Nations Department of Economic and Social Affairs. (2018, May). 68% of the world population projected to live in urban areas by 2050, says UN. Retrieved June 1, 2023, from https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html

Veras, J., Jaller, M., Sánchez-Díaz, I., Wojtowicz, J., Campbell, S., Levinson, H., Lawson, C., Powers, E., & Tavasszy, L. (2012, January). *NCHRP Report 739 / NCFRP Report 19: Freight Trip Generation and Land Use*.

Zaken, M. v. A. (2013, February). Stadsvernieuwing stadshavens Rotterdam - Ruimtelijke ordening en gebiedsontwikkeling - Rijksoverheid.nl [Last Modified: 2023-11-02T14:16 Publisher: Ministerie van Algemene Zaken]. Retrieved April 24, 2024, from https://www.rijksoverheid.nl/onderwerpen/ruimtelijke-ordening-en-gebiedsontwikkeling/gebiedsontwikkeling-per-regio/stadsvernieuwing-stadshavens-rotterdam

Zhang, R., & Yao, Y. (2022). Exploring the Impacts of the Built Environment on Urban Freight Generation: A Case Study of Light Freight Vehicles in Shanghai, 2593–2603. https://doi.org/10.1061/9780784484265.244

Zhou, Y., Xie, R., Tang, S., & Chen, Y. (2018). The Calculation Model of Freight Quantity Generation from a Survey of Beijing Restaurants, 499–507. https://doi.org/10.1061/9780784481523.050

A

Scientific Paper

Quantifying freight trip and freight generation from spatial developments in the Netherlands

Wim Gommans Delft University of Technology
Lóri Tavasszy Delft University of Technology
Michiel de Bok Delft University of Technology
Maarten Kroesen Delft University of Technology
Jan Anne Annema Delft University of Technology
Thymo Vlot Movares Utrecht
Jessica van Rijn Movares Utrecht

Abstract

This research aims to enhance the understanding and quantification of logistics in the Netherlands by identifying factors influencing freight activities in urban areas. Using a linear regression model developed from BasGoed freight tour data, CBS location data, and Dutch zonal plan data, the study integrates 26 variables to predict freight trip generation and freight generation. The main explaining factors include the surface area of distribution centers, employment in various sectors, the presence of rail or inland waterway terminals, urban density, and the surface area of business spaces. The model, validated through statistical tests and literature comparisons, provides reliable predictions, with employment influencing freight production and surface area affecting freight attraction.

A case study in the Utrecht A12 zone demonstrated the practical relevance of the model, offering insights for policymakers to design logistics-friendly urban environments. Limitations include simplified statistical procedures and reduced accuracy in seaport and transhipment areas. Despite these constraints, the model delivers valuable predictions for various spatial scenarios.

Future research should explore model performance using CBS XML-microdata, consider non-linear relationships, and examine variables for shipment size and vehicle type. Addressing causality and endogeneity with instrumental variables is also recommended. This study contributes significantly to both scientific and societal domains, providing a robust model for freight trip generation and freight generation, and guiding municipalities in effective spatial planning and decision-making for sustainable urban development.

Keywords: Freight trip generation, Freight generation, Logistics quantification, Regression model, Spatial planning

1 Background

This research seeks to expand the understanding of logistics quantification in the Netherlands by identifying key factors influencing Freight Trip Generation (FTG) and Freight Generation (FG). Using a regression model, containing various data sources, the study aims to quantify freight activities in urban areas, offering valuable insights for policymakers to optimize logistics integration into spatial development projects.

Inefficient spatial planning leads to increased congestion and delays in logistics, causing higher costs and environmental impacts. Poor alignment of distribution centers with transportation networks disrupts goods distribution and neighbourhood environments. The rapid expansion of logistics, driven by e-commerce and urbanization, presents challenges for cities, including finding space for logistic functions and managing freight movements. Efforts to promote sustainability, such as zero-emission zones and consolidation centers, require a detailed study on logistics quantification. Without accurate data on logistics movements, municipalities struggle to plan effectively, leading to congestion, safety issues, and inefficiencies. Improved spatial planning tools and index numbers for logistics are needed to optimize freight operations, reduce environmental impacts, and ensure sustainable urban development.

The scientific relevance of this study lies in its development of a statistically significant model that incorporates recent data and identifies new explanatory factors for FTG and FG. While previous research has emphasized the importance of employment and surface area in predicting freight movements, this study aims to build upon these findings by integrating additional variables such as urban density. Furthermore, it utilizes a high-quality data source that comprehensively covers the entire Netherlands.

From a societal perspective, due to the absence of suitable statistics and tools in the field of logistics, there is an urge for a guide that can assist municipalities in identifying and understanding the specific logistics requirements of a given area and acquiring these figures easily. With this guide, presented as index numbers, municipalities can quantify what kind of logistics is generated in a particular area. By providing this valuable information, the guide can contribute to educated decision-making and effective planning regarding spatial use within city boundaries as well as outside.

The primary goal of this research is to enhance the understanding of logistics generation in urban environments in the Netherlands. Specifically, seeking to identify the explaining factors for freight (trip) generation in urban areas. By doing so, it aims to offer valuable insights for spatial planners to more efficiently implement logistics into urban developments, considering both current and future demands. The objectives are:

1. To analyze existing FTG and FG models and their explaining factors
2. To develop a statistically significant model based on recent data
3. To apply the developed model to a relevant use case and assess its practical utility

The outline of this paper is as follows: section 2 provides a literature review on existing FTG and FG models. section 3 discusses the research design, including gaps, objectives, questions, methodology and data used in this study. Moreover, the model development is described. section 4 shows the results of this study. section 5 applies the model to the A12 zone use case. Finally, section 6 presents the discussion and section 7 the conclusion and recommendations.

2 Literature Review

2.1 Freight (trip) generation models

Freight Trip Generation and Freight Generation models are essential tools in urban logistics planning. These models estimate the number of freight trips and the volume of goods transported within urban areas, providing critical insights for traffic management and infrastructure development. Existing literature on FTG and FG models reveals several methodologies, primarily focusing on regression analysis to identify key explaining factors. Table 1 gives an overview of this existing literature with freight models based on the study area, the method used, results, whether it is on an aggregate or disaggregate level and finally the dependent[1] and independent[2] variables. Concerning the data collection and size of the models, a more recent study by Rodoshi et al. highlighted the data collection techniques used in FTG/FG models. The conclusion of this study was that truck trip-based studies have large sample sizes

[1] The dependent variable is the outcome variable that is being predicted or explained by the independent variables in a regression model. It represents the variable that is being observed and is influenced by the independent variables in the model. Here the dependent variables can be the amount of freight trips or the amount of freight weight.

[2] The independent variable is the variable that is presumed to have an effect on the dependent variable.

that range from 5,276 (Holguín-Veras and Patil 2007) to 135,564 (Gonzalez-Calderon et al. 2021).

A general analysis of this overview shows that the Netherlands is barely represented in the studies conducted. A fair share is conducted in the United States of America, some across Europe and some in parts of Asia. A large part of the studies have a regression analysis as a method for the model, have vehicle trips as the dependent variable and show those trips on a disaggregate level. Moreover, the amount of independent variables taken into consideration in these studies is quite considerate and ranges from share of transport costs to commodity type, and from employment to storage and parking spaces.

2.1.1 Methods

From previous work, four methods for FG/FTG models can be determined: linear regression my means of Ordinary Least Squares (OLS), Multiple Classification Analysis (MCA), trip rate and input-output. The regression method, being mostly based on a statistically robust technique, is seen most in previous work (Table 1). It is easy to use and can uncover patterns, gain insights into data and make predictions for the future. Although by use of a MCA, parameters can freely change across intervals of the independent variable (Lawson et al. 2012), comparative analysis revealed that MCA performed slightly better than regression models for estimating Freight Trip Attraction (FTA), although there was only a very small difference to the overall error (Veras et al. 2012). Trip rate is a method and can be derived from trip- and tour data, trips can then be divided by a factor (i.e. employees) which shows the rates per employee. Finally, input-output is based on economic interdependencies between different zones. Trips or volumes are then predicted based on the economic value, often for aggregate areas.

[3] Aggregate/Disaggregate.
[4] Multiple Classification Analysis.

2.1.2 Explaining factors

Explaining factors are the independent variables which affect the dependent variables (vehicle trips or volume/weight) in a model. From previous work, it can be derived that employment (17 times) and commodity type (7 times) are in most cases the explaining factors for FG/FTG models. Area, number of employees and population are also derived to be significant in some models. Pani et al. suggested in his study to use of multiple-variable models for quantifying freight activities from establishments subjected to data availability. Single-variable FG models may be used when the data availability is limited to one variable.

For employment, Jaller et al. results showed that the strongest correlation between employment and area is seen in commercial areas, office and retail in particular. The employment-based variables perform better than the area-based in these situations. Moreover, Pani et al. explained that employment is invariably considered the most preferred causal variable in FG/FTG models. It is often considered along with other variables such as business area, commodity type, industry segment, gross floor area, building area or parking area, or both while modelling the freight models. Also what is interesting regarding employment as an explaining factor, is that Sánchez-Díaz, Holguín-Veras, and Wang described that FTA is better modelled using non-linear models for all industry sectors. Specifically, the FTA of business establishments is concave with employment, flattening as employment increases. That matched the result of Oliveira et al., where the results of the study showed that the number of employees has a more significant influence in small cities and a lower influence in medium-sized municipalities. Finally, Sahu and Pani stated that the predictive ability of business size variables suggests that employment is more suitable for representing FP, while area explains FA better.

2.1.3 Freight (Trip) Generation

The literature also points out a few arguments regarding whether a FTG or FG model should be used. Holguín-

Study	Study area/country	Method; results; A/D[3]	Dependent variable	Independent variable
Iding, Meester, and Tavasszy 2002	Germany	Regression; $R^2 \leq 0.88$; D	Vehicle trips	No of employees & area firm, Site area, Employment
Beagan, Fischer, and Kuppam 2007	Wisconsin, Florida, Indiana, Tennessee	Regression; $R^2 \leq 0.98$; A	Weight, Vehicle trips	Employment type, No of households, Population
Institute of Transportation Engineers 2008	USA	Trip rate; D	Vehicle trips	Land use
Bastida and Holguín-Veras 2009	New York, Manhatten, Brooklyn	Regression, MCA[4]; D	Vehicle trips	Commodity type, Industry sector, Employment
Giuliano et al. 2010	Los Angeles	Input - output; A	Weight	Inter and intra-regional commodity flow, Employment
Holguín-Veras et al. 2011	New York, USA	Regression; $R^2 \leq 0.93$; D	Vehicle trips	Employment, Commodity type
Campbell et al. 2012	New York, USA	MCA; D	Vehicle trips	Employment, Commodity type
Lawson et al. 2012	New York, USA	Regression; $R^2 \leq 0.49$; D	Vehicle trips	Employment, Land Use, Commodity type
Alho and Silva 2014	Lisbon, Portugal	Regression; $R^2 \leq 0.67$; D	Vehicle trips	Employment, Frontage width, Warehouse area, Sales area, Commodity type
Asuncion 2014	Canterbury, New Zealand	Regression; $R^2 \leq 0.92$; D	Vehicle trips	Employment, Retail trading area, Storage and parking spaces, Product variation number
Holguín-Veras et al. 2014	New York, USA	Regression; $R^2 \leq 0.60$; D	Weight, Vehicle trips	Employment, Commodity type
Jaller et al. 2014	New York, New Jersey	Regression; $R^2 \leq 0.91$; D	Vehicle trips	Employment & area, Employment, Area, Commodity type
Ha and Combes 2016	France	Regression; $R^2 \leq 0.68$; D	Weight	Employment, No of clients, No of carriers, Share of transport cost in product value
Mommens, Lier, and Macharis 2017	Belgium	Regression; D	Vehicle trips	Employment, Area per activity, Population
Sánchez-Díaz 2017	Gothenburg, Sweden	Regression; $R^2 \leq 0.75$; D	Weight, Vehicle trips	Employment, Area, Commercial sector
Pani et al. 2018	Kerala, India	Regression; $R^2 \leq 0.64$	Weight	No employees & gross floor area, No of years in business
Gonzalez-Feliu, Palacios-Argüello, and Suarez-Nuñez 2020	Lyon, France	Regression; $R^2 \leq 0.88$; D	Vehicle trips	Population, Employment, Distance between zones
Sanchez-Diaz 2020	Stockholm, Sweden	Regression; A	Vehicle trips	No of employees, Area
Cheah, Mepparambath, and Ricart Surribas 2021	Singapore	Regression; D	Vehicle trips	Employment, Establishment type
Oliveira et al. 2022	Brazil	Regression	Vehicle trips	No of employees, Area
Zhang and Yao 2022	Shanghai	Regression; $R^2 \leq 0.55$; D	Vehicle trips	Population, Employment, Population density, Land use
Dhulipala and Patil 2023	India	Regression; D	Vehicle trips	Employment

Table 1: Overview of past FG/FTG studies (adapted from Pani et al. 2018).

Veras et al. describes that the correlation of establishment size variables (e.g., employment) with FG is relatively stronger than FTG due to its direct physical correlation with the scale of commodity production and attraction. FTG, on the other hand, is influenced by shipment size and logistic decisions. Veras et al. depicts FG as an expression of economic activity performed at a business establishment. Here the input materials are processed and transformed generating an output that, in most cases, is transported elsewhere for further processing, storage, distribution, or consumption. FTG, on the other hand, is the result of the logistic decisions concerning how best to transport the FG in terms of shipment size, frequency of deliveries, and the vehicle or mode used. The ability of the distributor to change shipment size to minimize total logistic costs, as it allows carriers to increase the cargo transported (the FG) without proportionally increasing the corresponding FTG is of great importance. As a result, FTG cannot be universally assumed to be proportional to business size because large establishments could receive larger amounts of cargo without the accompanying increases in FTG. This has big implications for FTG modelling, as normally it is implicitly assumed that FTG and business size variables (e.g., square footage, employment) are in proportion. The case studies in the work of Veras et al. confirm that proportionality between FTG and business size only happens in a minority of industry segments.

2.2 Conclusion

Despite the comprehensive methodologies, current FTG and FG models exhibit several limitations. Firstly, there is a lack of representation of the Netherlands in the existing studies, with most research conducted in the United States, Europe, and Asia. Secondly, few models provide an extensive analysis per commodity type, limiting their applicability to diverse logistics scenarios. Additionally, the highest sample size of 135,564 trucks, though significant, still leaves room for improvement in data granularity and coverage. Moreover, this review concludes that regression analysis using OLS is the most suitable method for developing a model to quantify logistics in urban areas.

3 Methodology

3.1 Research Design

The research is divided into three phases: desk research, model development, and application. The desk research phase involves a comprehensive literature review and identification of research gaps. This phase sets the foundation for the study by providing a thorough understanding of existing knowledge and highlighting areas for further investigation. The model development phase focuses on creating a regression model using the BasGoed freight tour data and two additional sources. This phase involves data preprocessing, variable selection, and regression analysis. The final phase is the application of the model to a relevant use case, testing its validity and practical utility.

3.1.1 Desk Research

The desk research phase involves an extensive review of existing literature on freight (trip) generation models. This includes identifying key explaining factors, understanding the methodologies used in previous studies, and assessing the limitations of existing models. The literature review helps in identifying gaps and areas for further investigation, guiding the development of the model.

3.1.2 Model Development

Variable selection is guided by the literature review and the availability of data, focusing on factors that have been identified as significant predictors of freight generation. The regression analysis involves applying Ordinary Least Squares techniques to estimate the relationships between the selected variables and freight movements. This step includes testing for statistical significance and ensuring the robustness of the model.

3.1.3 Application

The application phase tests the model on a relevant use case, specifically the A12 zone. This phase involves applying the model to estimate Freight Trip Generation based on the identified variables, assessing its accuracy and practical utility. The results of this phase provide insights into the spatial distribution of logistics activities and their implications for infrastructure development.

3.2 Data Collection

This research utilizes three primary data sources: BasGoed freight tour data, CBS location data, and Dutch zoning plan data. These datasets provide comprehensive information on freight movements, socioeconomic indicators, and land use characteristics, respectively. The BasGoed data includes detailed records of freight trips, offering insights into the origins, destinations, and characteristics of each trip. The CBS location data provides demographic and economic indicators at a granular level, essential for linking freight movements to broader urban characteristics. The Dutch zoning plan data offers information on land use designations, helping to understand the spatial distribution of logistics activities. An overview of the sources is shown in Figure 1.

3.2.1 BasGoed Freight Tour Data

The BasGoed dataset is a key resource for this research, containing over 4 million trips recorded in 2018. It provides detailed information on each trip, including the type of commodity transported, vehicle type, trip origin and destination, and trip purpose. This dataset is crucial for understanding the patterns of freight movements and identifying key explanatory variables. The comprehensive nature of the BasGoed data allows for a detailed analysis of logistics activities, making it a valuable resource for developing robust models.

3.2.2 CBS Location Data

The CBS location data includes information on employment, population, and other socio-economic factors. This data is available at a granular level, allowing for a detailed analysis of the relationship between socioeconomic characteristics and freight generation. By linking this data with the BasGoed dataset, the research can identify patterns and trends in logistics movements. The CBS data is essential for understanding the broader context of freight generation and trip generation, providing insights into how socioeconomic factors influence logistics activities.

3.2.3 Dutch Zoning Plan Data

Zoning plans designate land areas for specific uses, such as residential, retail, or industrial. The Dutch zoning plan data helps in understanding the spatial distribution of logistics activities and their relationship with urban planning decisions. This data provides insights into how different land uses contribute to freight generation, offering valuable information for developing the model. By incorporating zoning plan data, the research can assess the impact of land use decisions on logistics movements, providing a comprehensive view of urban logistics.

3.3 Model Development

The model development phase involves creating a multiple regression-based FTG/FG model using the selected three datasets. The model aims to predict freight movements based on various explaining factors, including employment, surface area, and urban density. A total number of 26 variables is used in this regression model. The accuracy of the model is verified through statistical tests, including R-squared and t-test, to ensure its predictive capabilities. Moreover, it incorporated a stratified test and train dataset to ensure a generalized outcome.

(a) BasGoed freight tour data

(b) CBS location data

(c) Dutch zoning plan data

Figure 1: Impression of the three different data sources

4 Results

The regression analysis reveals several significant explaining factors for freight trip generation and freight volume generation. The surface area of distribution centers, employment in various sectors, and the presence of terminals emerge as the most influential factors. Urban density levels also play a role in affecting the spatial distribution of logistics movements.

The model identifies the surface area of distribution centers, the employment in the agricultural, industrial, retail and services and other sector, the presence of rail or inland waterway terminals, the urban density level and the surface area of business, industrial and offices as key factors influencing FTG.

For FG, the model highlights the population, the surface area of distribution centers, the employment in the agricultural, industrial, retail and services and other sector, the presence of container, rail or inland waterway terminals, the urban density level and the surface area of business, industrial, retail and offices as significant factors.

5 Application

The application process begins with the identification of feasible spatial scenarios that the models can address. These scenarios include urban growth, economic development, and changes in land use, based on the significant explaining factors for FTG. By simulating these scenarios, we can observe how the models predict changes in freight trip movements and identify potential challenges and opportunities for urban logistics planning. The first step in this phase involves defining the use case, which includes specifying the area under study, its parameters, and the expected outcomes.

For this study, we focus on a specific use case involving the A12 zone. To ensure suitable and affordable housing in the Utrecht region, the construction of new housing is very important. The A12 zone is the area on both sides of the A12 motorway between junction Oudenrijn and Lunetten. This area has been designated as a site with many opportunities for new housing, businesses, and nature and recreational areas after 2030 (Gemeente Nieuwegein n.d.). In total, 15250 new houses are planned to be built and 9300 new jobs are realised. An overview of the corresponding NRM zones to the use case is shown in Figure 2.

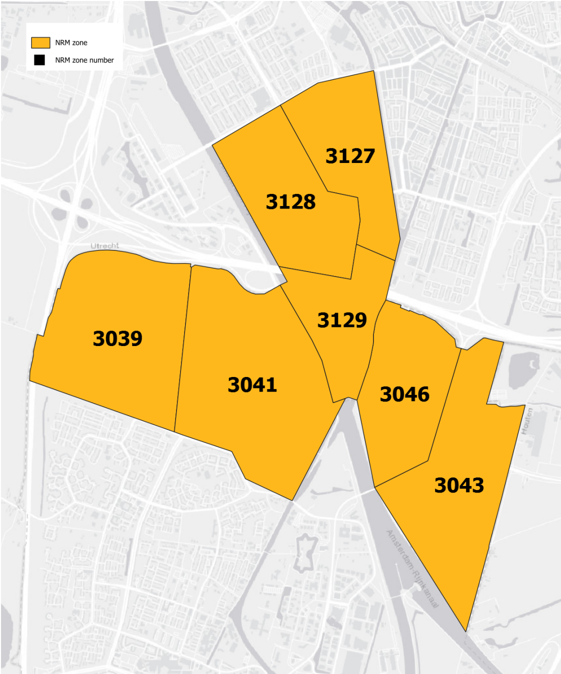


Figure 2: The corresponding NRM zones of the A12 zone area

Once the use case is defined, the next step involves simulating the impact of the spatial scenario on freight trip movements. This is conducted by adding the increase of employment to the corresponding NRM zones and with a linear factor also increasing the surface area of that NRM zone. Both the increase in surface area and employment are finally added to the designated variable (i.e. surface area business) according to the total weight of the employment or surface area in that zone. The expected increase of the parameters is applied to the current parameters. Those are consequently applied to the explaining coefficients which results in a future freight trip scenario. The impact is calculated using the additive growth method, meaning that the predicted impact is added to the base scenario (base dataset).

The application of the explaining coefficients resulted in a predicted freight trip production increase of 414 trips/workweek and a predicted attraction increase of 408 trips/workweek of the study area, see Table 2. The total predicted future production/attraction trips are 3586 trips/workweek and 3556 trips/workweek, respectively. Considering the initial production/attraction trips of 3172 and 3148, respectively, this is a significant increase of $\approx 13\%$ trips in the study area.

In summary, the application of the FTG model to the A12 zone provided valuable insights into the relationship between a spatial planning scenario and freight movements. The quantified results underscore the utility of the model in forecasting and planning for future logistics needs.

6 Discussion

Using a robust linear regression model based on synthetic data from the BasGoed freight tour data, the study identified employment and surface area as the most significant predictors of Freight Generation. The statistical validity of the model was confirmed through rigorous testing and alignment with existing literature, noticeably supporting findings by Sahu and Pani and Pani et al. Employment emerged as having a high impact, significantly affecting Freight Trip Generation across various models, while the surface area was more indicative of Freight Attraction. The study also highlighted the importance of rail terminals in enhancing logistics efficiency and reducing road congestion. Additionally, an inverse relationship between urban density and Freight Trip Attraction was observed, suggesting more efficient logistics networks in densely populated areas. Practical applications of the model demonstrated its utility in forecasting logistics movements under different spatial planning scenarios, aiding policymakers in optimizing infrastructure development. However, limitations such as the use of one-sided t-tests, the model's applicability to seaport and transhipment areas, and issues with endogeneity were noted, indicating areas for further refinement and research.

7 Conclusion

The goal of this research was to gain insights into the quantification of logistics in the Netherlands by determining the key factors influencing logistics generation

Production (trips/workweek)								
NRM zone	**3039**	**3041**	**3043**	**3046**	**3127**	**3128**	**3129**	**Total**
Observed trips	229	113	151	1184	618	517	360	3172
Model outcome current	183	160	173	1172	449	363	258	2759
Model outcome future	211	277	181	1244	463	518	280	3174
Predicted freight trip impact	28	117	8	73	13	154	21	414
Predicted future freight trips	257	230	159	1257	631	671	381	3586
Attraction (trips/workweek)								
NRM zone	**3039**	**3041**	**3043**	**3046**	**3127**	**3128**	**3129**	**Total**
Observed trips	233	128	166	1109	616	506	390	3148
Model outcome current	201	191	197	1069	521	355	274	2809
Model outcome future	225	335	205	1133	537	488	294	3217
Predicted freight trip impact	24	144	8	64	16	132	20	408
Predicted future freight trips	257	272	174	1173	632	638	410	3556

Table 2: Results of the A12 zone use case

and to provide knowledge for spatial planners to efficiently integrate logistics into urban developments. Our study employed an extensive analysis combining multiple datasets from the Netherlands and using Ordinary Least Squares regression techniques. The significant explaining factors identified for Freight Trip Generation and Freight Generation include the surface area of distribution centers, employment in various sectors, the presence of rail or inland waterway terminals, urban density levels, and the surface area of business, industrial, and office spaces. These factors, particularly surface area and employment, demonstrated the most substantial impact on freight (trip) generation. The findings indicate that the model can predict logistics movements with high accuracy, providing valuable input for spatial planning scenarios related to urban growth, economic development, and land use. This model also offers implications for enhancing traffic flow models and other planning tools, thereby aiding policymakers in making informed decisions to optimize logistics integration and soften potential challenges in urban environments.

This study prompts several recommendations for future research. First, validating the model with the original CBS XML microdata used in BasGoed could reveal differences in results due to location-specific factors. Observations such as the decrease in attraction with increased urban density and the unexpected link between construction logistics and office zoning show the need for further investigation. Additionally, modelling production and attraction at the household level, incorporating variables for shipment size and vehicle type, and understanding their impact on liveability factors like emissions and congestion are crucial. Non-linear relationships, such as those identified by Sánchez-Díaz, Holguín-Veras, and Wang, should be considered to refine regression models. Extending the methodology to include vans and other missing vehicles would enhance city logistics policies. Finally, translating these findings into actionable insights for practitioners and policymakers is essential for maximizing practical impact.

List of Abbreviations

FA Freight Attraction. 3, 8

FG Freight Generation. 1–9

FP Freight Production. 3

FTA Freight Trip Attraction. 3, 8

FTG Freight Trip Generation. 1–9

MCA Multiple Classification Analysis. 3

OLS Ordinary Least Squares. 3, 5, 9

References

Alho, André Romano and João de Abreu e Silva (Jan. 2014). "Freight-Trip Generation Model: Predicting Urban Freight Weekly Parking Demand from Retail Establishment Characteristics". en. In: *Transportation Research Record* 2411 (1). Publisher: SAGE Publications Inc, pp. 45–54.

Asuncion, Janice Sy (2014). "The Geographic Adaptive Potential of Freight Transportation and Production System in the Context of Fuel and Emission Constraints". en. In: Publisher: University of Canterbury. Department of Mechanical Engineering.

Bastida, Carlos and José Holguín-Veras (Jan. 2009). "Freight Generation Models: Comparative Analysis of Regression Models and Multiple Classification Analysis". en. In: *Transportation Research Record* 2097 (1). Publisher: SAGE Publications Inc, pp. 51–61.

Beagan, D., M. Fischer, and A. Kuppam (Sept. 2007). "Quick Response Freight Manual II". In.

Campbell, Shama et al. (2012). "Comparison Between Industrial Classification Systems in Freight Trip Generation Modeling". In: Number: 12-3825.

Cheah, Lynette, Rakhi Manohar Mepparambath, and Gabriella Marie Ricart Surribas (Mar. 2021). "Freight trips generated at retail malls in dense urban areas". In: *Transportation Research Part A: Policy and Practice* 145, pp. 118–131.

Dhulipala, Sowjanya and Gopal R. Patil (June 2023). "Regional freight generation and spatial interactions in developing regions using secondary data". en. In: *Transportation* 50 (3), pp. 773–810.

Gemeente Nieuwegein (n.d.). *De A12-zone*. nl.

Giuliano, Genevieve et al. (Mar. 2010). "Estimating Freight Flows for Metropolitan Area Highway Networks Using Secondary Data Sources". en. In: *Networks and Spatial Economics* 10 (1), pp. 73–91.

Gonzalez-Calderon, Carlos A. et al. (Mar. 2021). "Generalized noortman and van es empty trips model". In: *Transportation Research Part A: Policy and Practice* 145, pp. 260–268.

Gonzalez-Feliu, Jesus, Laura Palacios-Argüello, and Carlos Suarez-Nuñez (Apr. 2020). "Links between freight trip generation rates, accessibility and socio-demographic variables in urban zones". en. In: *Archives of Transport* 53 (1). Publisher: Index Copernicus, pp. 7–20.

Ha, Duy-Hung and François Combes (2016). "Building a Model of Freight Generation with a Commodity Flow Survey". en. In: *Commercial Transport*. Ed. by Uwe Clausen et al. Lecture Notes in Logistics. Cham: Springer International Publishing, pp. 23–37.

Holguín-Veras, José and Gopal Patil (Jan. 2007). "Integrated Origin-Destination Synthesis Model for Freight with Commodity-Based and Empty Trip Models". In: *Transportation Research Record* 2008, pp. 60–66.

Holguín-Veras, José et al. (Jan. 2011). "Freight Generation, Freight Trip Generation, and Perils of Using Constant Trip Rates". en. In: *Transportation Research Record* 2224 (1). Publisher: SAGE Publications Inc, pp. 68–81.

Holguín-Veras, José et al. (Jan. 2014). "3 - Freight Generation and Freight Trip Generation Models". In: *Modelling Freight Transport*. Ed. by Lóránt Tavasszy and Gerard de Jong. Oxford: Elsevier, pp. 43–63.

Iding, Mirjam H. E., Wilhelm J. Meester, and Lóri Tavasszy (2002). "Freight trip generation by firms". eng. In: Louvain-la-Neuve: European Regional Science Association (ERSA).

Institute of Transportation Engineers (2008). *Trip Generation Manual*. Tech. rep. 8th Edition.

Jaller, Miguel et al. (2014). "Area Based Freight Trip Generation Models". In: Number: 14-4908.

Lawson, Catherine T. et al. (Jan. 2012). "Estimated Generation of Freight Trips Based on Land Use". en. In: *Transportation Research Record* 2269 (1). Publisher: SAGE Publications Inc, pp. 65–72.

Mommens, Koen, Tom van Lier, and Cathy Macharis (Jan. 2017). "Freight demand generation on commodity and loading unit level". en. In: *European*

Journal of Transport and Infrastructure Research 17 (1). Number: 1.

Oliveira, L.K.D. et al. (2022). “Modelling Freight Trip Generation Based on Deliveries for Brazilian Municipalities”. English. In: *Sustainability (Switzerland)* 14 (16).

Pani, A. et al. (2018). “Modelling urban freight generation: A case study of seven cities in Kerala, India”. In: *Transport Policy* 69, pp. 49–64.

Rodoshi, Noushin et al. (Feb. 2024). “A REVIEW OF THE FREIGHT GENERATION MODELS AND DATA COLLECTION TECHNIQUES”. In.

Sahu, P.K. and A. Pani (2020). “Freight generation and geographical effects: modelling freight needs of establishments in developing economies and analyzing their geographical disparities”. In: *Transportation* 47 (6), pp. 2873–2902.

Sanchez-Diaz, Ivan (Dec. 2020). “Assessing the magnitude of freight traffic generated by office deliveries”. In: *Transportation Research Part A: Policy and Practice* 142, pp. 279–289.

Sánchez-Díaz, Iván (Aug. 2017). “Modeling urban freight generation: A study of commercial establishments freight needs”. In: *Transportation Research Part A: Policy and Practice*. SI: Freight Behavior Research 102, pp. 3–17.

Sánchez-Díaz, Iván, José Holguín-Veras, and Xiaokun Wang (Jan. 2016). “An exploratory analysis of spatial effects on freight trip attraction”. en. In: *Transportation* 43 (1), pp. 177–196.

Veras, José et al. (Jan. 2012). *NCHRP Report 739 / NCFRP Report 19: Freight Trip Generation and Land Use*.

Zhang, R. and Y. Yao (2022). “Exploring the Impacts of the Built Environment on Urban Freight Generation: A Case Study of Light Freight Vehicles in Shanghai”. In: pp. 2593–2603.

B

Trip Rates NST/R

Description	Prod. trips/Employee/Workweek	Attr. trips/Employee/Workweek
Agricultural products and live animals	0.019	0.018
Foodstuffs and animal fodder	0.032	0.032
Solid mineral fuels	0.005	0.005
Petroleum products	0.002	0.002
Ores and metal waste	0.002	0.002
Metal products	0.006	0.006
Crude and manufactured minerals, building materials	0.014	0.015
Fertilizers	0.001	0.001
Chemicals	0.022	0.021
Machinery, transport equipment, manufactured articles and miscellaneous articles	0.194	0.191
Empty trips	0.006	0.006

Table B.1: Weekly trip rate production/attraction per NST/R goods type

C

Additional Data Exploration

Label	Description
NST/R 0	Agricultural products and live animals
NST/R 1	Foodstuffs and animal fodder
NST/R 2	Solid mineral fuels
NST/R 3	Petroleum products
NST/R 4	Ores and metal waste
NST/R 5	Metal products
NST/R 6	Crude and manufactured minerals, building materials
NST/R 7	Fertilizers
NST/R 8	Chemicals
NST/R 9	Machinery, transport equipment, manufactured articles and miscellaneous articles
NST/R -1	Empty trips

Table C.1: NST/R goods type label description

Variable	Max.	Min.	Avg.	SD
Trips/Workweek				
FTP NST/R 0	1714	0	25.39	57.18
FTP NST/R 1	3400	0	43.21	91.56
FTP NST/R 2	334	0	6.36	9.08
FTP NST/R 3	1238	0	2.09	16.78
FTP NST/R 4	363	0	3.07	8.39
FTP NST/R 5	989	0	7.94	22.14
FTP NST/R 6	1110	0	19.38	29.14
FTP NST/R 7	503	0	0.83	7.99
FTP NST/R 8	4447	0	29.89	83.40
FTP NST/R 9	22184	0	260.30	485.29
FTP NST/R -1	8917	0	8.24	133.43
FTA NST/R 0	1551	0	24.23	48.25
FTA NST/R 1	3056	0	42.59	84.28
FTA NST/R 2	327	0	6.42	8.92
FTA NST/R 3	993	0	2.07	13.70
FTA NST/R 4	353	0	3.04	8.04
FTA NST/R 5	871	0	7.78	17.22
FTA NST/R 6	1113	0	19.60	27.90
FTA NST/R 7	356	0	0.80	5.93
FTA NST/R 8	3792	0	28.80	70.82
FTA NST/R 9	20540	0	256.95	440.65
FTA NST/R -1	8939	0	8.17	124.73
Tonnes/Workweek				
FP NST/R 0	26159.95	0	265.77	716.68
FP NST/R 1	52703.85	0	515.96	1240.83
FP NST/R 2	5666.80	0	78.30	170.45
FP NST/R 3	19910.21	0	21.96	267.49
FP NST/R 4	2742.38	0	21.82	73.72
FP NST/R 5	10804.79	0	76.98	246.81
FP NST/R 6	16651.38	0	230.39	498.69
FP NST/R 7	8745.62	0	10.27	135.14
FP NST/R 8	46130.27	0	271.33	891.32
FP NST/R 9	247814.93	0	2721.26	5543.19
FP NST/R -1	29252.93	0	27.03	437.74
FA NST/R 0	21228.58	0	257.44	539.02
FA NST/R 1	49005.68	0	511.08	1031.55
FA NST/R 2	6439.59	0	81.59	120.34
FA NST/R 3	4057.37	0	21.99	76.05
FA NST/R 4	3088.17	0	21.59	71.50
FA NST/R 5	9584.87	0	76.14	183.19
FA NST/R 6	19788.32	0	240.03	358.84
FA NST/R 7	6149.19	0	10.14	98.22
FA NST/R 8	34977.09	0	263.45	676.23
FA NST/R 9	206676.92	0	2706.47	4293.93
FA NST/R -1	29325.10	0	26.80	409.17

Table C.2: Dependent variable list and characteristics

Variable	Additional description
Surface area distribution centers	The amount of surface area (in km^2) of distribution centers in the zone.
Employment agricultural	The number of employment in the agricultural sector in the zone.
Employment industrial	The number of employment in the industrial sector in the zone.
Employment retail	The number of employment in the retail sector in the zone.
Employment services	The number of employment in the services sector in the zone.
Employment government	The number of employment in the governmental sector in the zone.
Employment others	The number of employment in any other sector than above in the zone.
Urban density	The degree of urban density in the zone, see Table 5.4 for an explanation per level.
Surface area business	The amount of surface area (in km^2) of business area in the zone. For business zoning, it means that it cannot simply be turned into a residential zoning. For light activity, however, a residential zoning can be used as a business zoning.
Surface area industrial	The amount of surface area (in km^2) of industrial area in the zone. It is a land in use for industry, commerce and business services. Business park includes: factory site; port site; auction site; exhibition site; livestock market (indoor or outdoor); wholesale complex; site with banks and insurance companies etc; associated storage area and parking; garage (incl. car park); garage of bus company; office building; associated parking areas (Statistiek, n.d.-a).
Surface area city center	The amount of surface area (in km^2) of a city center in the zone.
Surface area retail	The amount of surface area (in km^2) of retail in the zone.
Surface area mixed	The amount of surface area (in km^2) of mixed land-use in the zone.
Surface area horeca	The amount of surface area (in km^2) of horeca in the zone.
Surface area office	The amount of surface area (in km^2) of offices in the zone. Moreover, a building for the commercial provision of services where the public is not or only to a minor extent directly addressed or assisted ("Bestemmingsplan Bedrijventerreinen: Regels", n.d.).
Surface area residential	The amount of surface area (in km^2) of residential in the zone.
Surface area urban development	The amount of surface area (in km^2) of urban development in the zone.

Table C.3: Additional information for surface areas

D
Estimation Results NST/R

Table D.1: Estimation results for explaining Freight Trip Production per NST/R

Trips/Workweek/Zone N = 5152		Intercept		Population	Surface area distribution centers	Employment: Agricultural	Employment: Industrial	Employment: Services	Employment: Others	Terminal: Rail	Terminal: Inland waterway	Urban density: 3	Urban density: 4	Urban density: 5	Urban density: 6	Surface area: Business	Surface area: Industrial	Surface area: Office	R²	Adj. R²	Predicted R²
	Unit	-		x1000	km²	x1000				-	-	-	-	-	-	km²					
NST/R 0	**Coefficient**	**-4.89**			**172.95**	**312.16**	**3.05**		**1.71**		**7.12**	**5.25**	**7.76**	**11.32**	**11.27**	**94.06**	**8.07**		**0.68**	0.68	0.53
	St. coefficient		25.61		4.18	28.76	1.55		1.47		3.17	1.61	3.10	4.98	4.52	34.64	2.78				
Agricultural products and live animals	*T-value*	*-4.59*	*55.98*		*7.90*	*60.65*	*2.85*		*2.87*		*6.80*	*3.08*	*5.44*	*8.32*	*7.50*	*72.75*	*5.28*				
	VIF				*1.40*	*1.12*	*1.77*		*1.73*		*1.32*	*1.09*	*1.14*	*1.20*	*1.37*	*1.15*	*1.41*				
NST/R 1	**Coefficient**	**10.73**			**143.83**	**28.18**	**76.97**				**6.92**			**5.19**		**111.80**	**7.05**		**0.54**	0.54	0.50
	St. coefficient		43.37		3.47	2.60	39.19				3.09			2.28		41.17	2.43				
Foodstuffs and animal fodder	*T-value*	*8.90*	*51.89*		*3.65*	*3.07*	*40.76*				*3.63*			*2.69*		*47.52*	*2.54*				
	VIF				*1.37*	*1.08*	*1.61*				*1.19*			*1.10*		*1.14*	*1.40*				
NST/R 2	**Coefficient**	**2.40**		**0.09**			**7.67**	**0.69**			**4.60**					**4.79**	**1.04**		**0.44**	0.44	0.33
	St. coefficient		6.40	0.27			3.90	0.51			2.05					1.76	0.36				
Solid mineral fuels	*T-value*	*17.75*	*74.09*	*3.02*			*39.21*	*5.56*			*23.40*					*19.97*	*3.74*				
	VIF			*1.32*			*1.64*	*1.30*			*1.25*					*1.11*	*1.30*				
NST/R 3	**Coefficient**	**0.38**					**3.18**								**0.99**	**7.41**			**0.22**	0.22	0.09
	St. coefficient		1.92				1.62								0.40	2.73					
Petroleum products	*T-value*	*3.55*	*21.55*				*17.86*								*4.43*	*30.12*					
	VIF						*1.14*								*1.05*	*1.09*					
NST/R 4	**Coefficient**	**0.06**			**108.84**		**4.60**			**4.98**	**0.68**		**0.41**	**0.74**	**0.74**	**8.42**	**0.92**		**0.66**	0.66	0.46
	St. coefficient		3.08		2.63		2.34			0.39	0.30		0.16	0.33	0.30	3.10	0.32				
Ores and metal waste	*T-value*	*0.57*	*51.74*		*38.65*		*33.80*			*6.14*	*5.00*		*2.40*	*4.71*	*4.36*	*48.28*	*4.48*				
	VIF				*1.37*		*1.70*			*1.13*	*1.27*		*1.09*	*1.12*	*1.13*	*1.22*	*1.48*				
NST/R 5	**Coefficient**	**1.68**		**0.21**			**17.88**									**17.19**			**0.40**	0.40	0.39
	St. coefficient		7.85	0.58			9.10									6.33					
Metal products	*T-value*	*5.69*	*38.42*	*2.86*			*43.82*									*30.48*					
	VIF			*1.11*			*1.19*									*1.09*					
NST/R 6	**Coefficient**	**4.94**					**26.84**	**1.64**			**11.69**	**2.28**	**3.26**	**3.61**	**3.22**	**22.50**	**2.69**		**0.53**	0.53	0.40
	St. coefficient		19.49				13.66	1.21			5.21	0.70	1.30	1.59	1.29	8.29	0.93				
Crude and manufactured minerals, building materials	*T-value*	*8.68*	*75.76*				*45.70*	*4.35*			*19.86*	*2.38*	*4.11*	*4.84*	*3.92*	*31.41*	*3.26*				
	VIF						*1.70*	*1.37*			*1.31*	*1.05*	*1.10*	*1.15*	*1.29*	*1.11*	*1.30*				
NST/R 7	**Coefficient**	**-0.13**					**0.49**									**9.91**	**0.88**		**0.24**	0.24	0.30
	St. coefficient		0.83				0.25									3.65	0.30				
Fertilizers	*T-value*	*-1.28*	*8.96*				*2.41*									*38.80*	*2.98*				
	VIF						*1.38*									*1.09*	*1.28*				
NST/R 8	**Coefficient**	**5.45**			**145.37**		**61.35**			**38.20**			**4.89**	**4.23**		**78.93**	**5.71**		**0.53**	0.53	0.31
	St. coefficient		29.78		3.51		31.23			3.00			1.95	1.86		29.07	1.97				
Chemicals	*T-value*	*5.95*	*46.49*		*4.80*		*42.41*			*4.39*			*2.90*	*2.75*		*42.39*	*2.60*				
	VIF				*1.37*		*1.59*			*1.12*			*1.06*	*1.08*		*1.20*	*1.46*				
NST/R 9	**Coefficient**	**41.60**			**1866.68**	**412.95**	**448.41**			**185.90**	**44.21**	**29.26**	**41.63**	**57.67**	**46.55**	**451.42**	**29.40**	**599.10**	**0.67**	0.66	0.44
	St. coefficient		260.74		45.10	38.05	228.28			14.61	19.72	8.98	16.61	25.35	18.66	166.24	10.13	7.42			
Machinery, transport equipment, manufactured articles and miscellaneous articles	*T-value*	*5.27*	*76.80*		*11.61*	*10.82*	*57.51*			*4.02*	*5.69*	*2.32*	*3.94*	*5.74*	*4.31*	*45.37*	*2.50*	*2.15*			
	VIF				*1.37*	*1.12*	*1.72*			*1.13*	*1.32*	*1.09*	*1.12*	*1.14*	*1.18*	*1.23*	*1.49*	*1.10*			
NST/R -1	**Coefficient**	**-13.05**													**11.72**	**233.05**	**15.70**		**0.35**	0.35	0.43
	St. coefficient		8.75												4.70	85.82	5.41				
Empty trips	*T-value*	*-6.86*	*5.39*												*2.89*	*52.86*	*3.33*				
	VIF														*1.01*	*1.01*	*1.01*				

Table D.2: Estimation results for explaining Freight Trip Attraction per NST/R

Trips/Workweek/Zone N = 5152		Intercept		Population	Surface area distribution centers	Employment: Agricultural	Employment: Industrial	Employment: Services	Employment: Others	Terminal: Rail	Terminal: Inland waterway	Urban density: 2	Urban density: 3	Urban density: 4	Urban density: 5	Urban density: 6	Surface area: Business	Surface area: Industrial	Surface area: Office	R²	Adj. R²	Predicted R²
	Unit	-		*x1000*	*km²*	*x1000*				-	-	-	-	-	-	-	*km²*					
NST/R 0	**Coefficient**	**-2.36**			**169.12**	**223.65**	**5.78**		**1.51**	**19.42**	**6.66**	**3.99**	**6.14**	**7.97**	**11.79**	**11.69**	**72.53**	**5.64**		**0.61**	*0.61*	*0.45*
Agricultural products and live animals	St. coefficient		24.38		4.09	20.61	2.94		1.29	1.53	2.97	1.05	1.88	3.18	5.18	4.69	26.71	1.94				
	T-value	*-2.08*	*58.96*		*8.54*	*48.05*	*5.96*		*2.80*	*3.45*	*7.04*	*2.17*	*3.72*	*5.60*	*8.59*	*7.85*	*59.77*	*3.94*				
	VIF				*1.40*	*1.15*	*1.78*		*1.73*	*1.13*	*1.35*	*1.07*	*1.10*	*1.15*	*1.21*	*1.38*	*1.23*	*1.49*				
NST/R 1	**Coefficient**	**9.95**			**262.43**		**58.66**		**1.70**	**43.29**	**5.83**			**5.87**	**8.70**	**5.82**	**130.58**			**0.65**	*0.65*	*0.50*
Foodstuffs and animal fodder	St. coefficient		42.83		6.34		29.86		1.46	3.40	2.60			2.34	3.83	2.33	48.09					
	T-value	*7.84*	*63.96*		*8.54*		*39.45*		*1.95*	*4.91*	*3.81*			*3.07*	*4.88*	*2.93*	*67.60*					
	VIF				*1.29*		*1.59*		*1.71*	*1.07*	*1.26*			*1.12*	*1.19*	*1.35*	*1.18*					
NST/R 2	**Coefficient**	**2.58**		**0.09**			**7.27**	**0.71**			**4.57**						**4.61**	**1.07**		**0.43**	*0.43*	*0.32*
Solid mineral fuels	St. coefficient		6.46	0.26			3.70	0.52			2.04						1.70	0.37				
	T-value	*19.11*	*74.87*	*2.89*			*37.23*	*5.68*			*23.29*						*19.25*	*3.87*				
	VIF			*1.32*			*1.64*	*1.30*			*1.25*						*1.11*	*1.30*				
NST/R 3	**Coefficient**	**0.46**					**2.61**									**0.99**	**8.54**			**0.32**	*0.32*	*0.11*
Petroleum products	St. coefficient		1.95				1.33									0.40	3.15					
	T-value	*5.10*	*26.19*				*17.53*									*5.34*	*41.56*					
	VIF						*1.14*									*1.05*	*1.09*					
NST/R 4	**Coefficient**	**0.20**			**113.50**		**3.29**		**0.16**	**4.98**	**0.71**			**0.50**	**0.76**	**0.72**	**8.48**	**1.02**		**0.65**	*0.65*	*0.44*
Ores and metal waste	St. coefficient		3.05		2.74		1.68		0.13	0.39	0.32			0.20	0.34	0.29	3.12	0.35				
	T-value	*1.82*	*52.27*		*40.63*		*24.11*		*2.07*	*6.26*	*5.33*			*2.97*	*4.90*	*4.14*	*49.60*	*5.05*				
	VIF				*1.40*		*1.78*		*1.71*	*1.13*	*1.27*			*1.12*	*1.19*	*1.36*	*1.22*	*1.48*				
NST/R 5	**Coefficient**	**2.63**		**0.19**			**11.69**								**1.41**		**18.21**			**0.49**	*0.49*	*0.36*
Metal products	St. coefficient		7.72	0.54			5.95								0.62		6.71					
	T-value	*12.46*	*55.99*	*3.89*			*42.31*								*4.47*		*47.78*					
	VIF			*1.27*			*1.22*								*1.23*		*1.09*					
NST/R 6	**Coefficient**	**5.34**			**39.25**		**23.48**	**1.28**	**0.83**		**11.36**		**2.32**	**3.25**	**3.79**	**3.20**	**24.83**	**2.20**		**0.56**	*0.56*	*0.39*
Crude and manufactured minerals, building materials	St. coefficient		19.70		0.95		11.95	0.94	0.71		5.07		0.71	1.30	1.67	1.28	9.14	0.76				
	T-value	*10.33*	*84.69*		*3.53*		*42.80*	*3.42*	*2.50*		*21.34*		*2.68*	*4.51*	*5.57*	*4.23*	*37.76*	*2.83*				
	VIF				*1.40*		*1.81*	*1.66*	*2.09*		*1.31*		*1.06*	*1.13*	*1.20*	*1.42*	*1.14*	*1.41*				
NST/R 7	**Coefficient**	**-0.52**							**0.33**	**4.81**					**0.37**	**0.50**	**10.97**	**1.05**		**0.51**	*0.51*	*0.43*
Fertilizers	St. coefficient		0.82						0.28	0.38					0.16	0.20	4.04	0.36				
	T-value	*-6.19*	*14.15*						*4.55*	*6.15*					*2.66*	*3.16*	*67.19*	*6.00*				
	VIF								*1.49*	*1.12*					*1.12*	*1.27*	*1.10*	*1.11*				
NST/R 8	**Coefficient**	**7.71**			**232.23**		**43.51**			**47.39**				**4.87**	**4.87**		**83.87**	**5.46**		**0.62**	*0.62*	*0.34*
Chemicals	St. coefficient		28.79		5.61		22.15			3.72				1.94	2.14		30.89	1.88				
	T-value	*10.91*	*58.27*		*9.94*		*39.00*			*7.06*				*3.74*	*4.11*		*58.41*	*3.22*				
	VIF				*1.37*		*1.59*			*1.12*				*1.06*	*1.08*		*1.20*	*1.46*				
NST/R 9	**Coefficient**	**60.46**			**2284.95**	**283.77**	**351.69**			**223.72**	**43.70**		**29.62**	**42.19**	**63.11**	**50.64**	**469.98**	**27.39**	**518.06**	**0.67**	*0.67*	*0.44*
Machinery, transport equipment, manufactured articles and miscellaneous articles	St. coefficient		257.65		55.20	26.15	179.03			17.58	19.49		9.09	16.84	27.75	20.31	173.07	9.43	6.42			
	T-value	*8.57*	*84.92*		*15.90*	*8.32*	*50.47*			*5.41*	*6.29*		*2.62*	*4.47*	*7.03*	*5.24*	*52.85*	*2.61*	*2.08*			
	VIF				*1.37*	*1.12*	*1.72*			*1.13*	*1.32*		*1.09*	*1.12*	*1.14*	*1.18*	*1.23*	*1.49*	*1.10*			
NST/R -1	**Coefficient**	**-19.65**							**5.04**	**98.79**					**6.77**	**12.48**	**254.24**	**11.16**		**0.51**	*0.51*	*0.41*
Empty trips	St. coefficient		8.59						4.32	7.76					2.97	5.00	93.63	3.85				
	T-value	*-10.30*	*6.57*						*3.11*	*5.57*					*2.15*	*3.47*	*68.70*	*2.81*				
	VIF								*1.49*	*1.12*					*1.12*	*1.27*	*1.10*	*1.11*				

Table D.3: Estimation results for explaining Freight Production per NST/R

Tonnes/Workweek/Zone *N = 5152*		Intercept		Population	Surface area distribution centers	Employment: Agricultural	Employment: Industrial	Employment: Retail	Employment: Services	Employment: Others	Terminal: Rail	Terminal: Inland waterway	Urban density: 3	Urban density: 4	Urban density: 5	Urban density: 6	Surface area: Business	Surface area: Industrial	Surface area: Mixed	Surface area: Office	R²	Adj. R²	Predicted R²
	Unit	-		x1000	km²	x1000					-	-	-	-	-	-	km²						
NST/R 0	**Coefficient**	**-91.14**			**664.94**	**3672.42**				**24.36**		**79.96**	**64.47**	**93.57**	**128.52**	**125.54**	**1315.64**	**105.47**			**0.68**	*0.68*	*0.50*
Agricultural products and live animals	St. coefficient		268.66		16.07	338.41				20.86		35.66	19.78	37.34	56.50	50.33	484.50	36.33					
	T-value	*-6.62*	*45.40*		*2.39*	*55.21*				*3.23*		*5.92*	*2.93*	*5.09*	*7.34*	*6.48*	*79.35*	*5.60*					
	VIF				*1.35*	*1.12*				*1.66*		*1.31*	*1.09*	*1.13*	*1.18*	*1.36*	*1.11*	*1.27*					
NST/R 1	**Coefficient**	**95.84**				**304.16**	**1103.28**	**111.06**				**71.90**					**1371.46**	**78.42**			**0.48**	*0.48*	*0.49*
Foodstuffs and animal fodder	St. coefficient		516.45			28.03	561.65	27.33				32.07					505.05	27.01					
	T-value	*5.83*	*42.63*			*2.30*	*41.22*	*2.23*				*2.61*					*40.71*	*2.03*					
	VIF					*1.08*	*1.53*	*1.14*				*1.21*					*1.12*	*1.30*					
NST/R 2	**Coefficient**	**-5.81**		**1.66**			**168.92**	**15.09**	**12.79**			**101.21**					**64.40**	**11.92**	**16.65**		**0.52**	*0.52*	*0.38*
Solid mineral fuels	St. coefficient		78.67	4.70			85.99	3.71	9.39			45.13					23.71	4.11	3.65				
	T-value	*-2.43*	*51.53*	*2.69*			*48.84*	*2.05*	*5.60*			*29.09*					*15.18*	*2.43*	*2.38*				
	VIF			*1.77*			*1.65*	*1.76*	*1.40*			*1.25*					*1.11*	*1.30*	*1.02*				
NST/R 3	**Coefficient**	**-2.89**					**50.40**									**14.45**	**91.30**				**0.18**	*0.18*	*0.07*
Petroleum products	St. coefficient		19.18				25.66									5.79	33.62						
	T-value	*-1.71*	*13.79*				*18.15*									*4.16*	*23.78*						
	VIF						*1.14*									*1.05*	*1.09*						
NST/R 4	**Coefficient**	**-4.43**			**562.01**		**43.49**				**40.82**	**2.99**			**5.08**	**7.96**	**104.73**	**7.34**		**133.07**	**0.62**	*0.62*	*0.52*
Ores and metal waste	St. coefficient		21.91		13.58		22.14				3.21	1.33			2.23	3.19	38.57	2.53		1.65			
	T-value	*-4.71*	*36.31*		*19.72*		*31.54*				*4.96*	*2.17*			*3.49*	*4.94*	*59.38*	*3.52*		*2.69*			
	VIF				*1.37*		*1.69*				*1.13*	*1.23*			*1.12*	*1.16*	*1.22*	*1.48*		*1.09*			
NST/R 5	**Coefficient**	**9.56**		**1.95**			**199.63**										**172.43**				**0.36**	*0.36*	*0.38*
Metal products	St. coefficient		75.94	5.54			101.63										63.50						
	T-value	*2.75*	*31.67*	*2.31*			*41.68*										*26.03*						
	VIF			*1.11*			*1.19*										*1.09*						
NST/R 6	**Coefficient**	**-10.25**					**531.48**	**59.41**	**32.57**			**251.34**				**28.79**	**300.36**				**0.57**	*0.57*	*0.44*
Crude and manufactured minerals, building materials	St. coefficient		231.37				270.56	14.62	23.91			112.08				11.54	110.61						
	T-value	*-1.73*	*53.55*				*59.22*	*3.15*	*4.91*			*25.62*				*2.55*	*25.08*						
	VIF						*1.34*	*1.36*	*1.48*			*1.21*				*1.28*	*1.10*						
NST/R 7	**Coefficient**	**-3.94**															**165.24**	**14.27**			**0.22**	*0.22*	*0.30*
Fertilizers	St. coefficient		10.10														60.85	4.92					
	T-value	*-2.34*	*6.28*														*37.85*	*3.06*					
	VIF																*1.00*	*1.00*					
NST/R 8	**Coefficient**	**36.05**					**688.44**				**333.88**						**856.87**				**0.49**	*0.49*	*0.31*
Chemicals	St. coefficient		269.49				350.47				26.23						315.55						
	T-value	*4.41*	*36.65*				*46.91*				*3.45*						*41.02*						
	VIF						*1.09*				*1.07*						*1.15*						
NST/R 9	**Coefficient**	**478.95**				**4587.70**	**5952.58**		**189.62**		**1731.38**	**564.97**			**265.13**		**4791.22**				**0.66**	*0.66*	*0.45*
Machinery, transport equipment, manufactured articles and miscellaneous articles	St. coefficient		2722.88			422.75	3030.31		139.21		136.03	251.95			116.56		1764.42						
	T-value	*8.24*	*69.15*			*10.62*	*72.41*		*3.41*		*3.34*	*6.30*			*2.92*		*42.60*						
	VIF					*1.07*	*1.38*		*1.19*		*1.07*	*1.20*			*1.11*		*1.17*						
NST/R -1	**Coefficient**	**-42.80**														**38.45**	**764.55**	**51.52**			**0.35**	*0.35*	*0.43*
Empty trips	St. coefficient		28.69													15.42	281.55	17.75					
	T-value	*-6.86*	*5.39*													*2.89*	*52.86*	*3.33*					
	VIF															*1.01*	*1.01*	*1.01*					

Table D.4: Estimation results for explaining Freight Attraction per NST/R

Tonnes/Workweek/Zone N = 5152		Intercept		Population	Surface area distribution centers	Employment: Agricultural	Employment: Industrial	Employment: Others	Terminal: Container	Terminal: Rail	Terminal: Inland waterway	Urban density: 2	Urban density: 3	Urban density: 4	Urban density: 5	Urban density: 6	Surface area: Business	Surface area: Industrial	R²	Adj. R²	Predicted R²
	Unit	-		x1000	km²	x1000			-	-	-	-	-	-	-	-	km²				
NST/R 0	Coefficient	-7.92			875.04	1770.13	27.41	21.11		169.74	49.76	47.94	68.09	85.78	129.38	138.55	945.61	74.51	0.58	0.58	0.41
Agricultural products and live animals	St. coefficient		258.97		21.14	163.11	13.95	18.08		13.34	22.19	12.59	20.89	34.23	56.88	55.55	348.23	25.67			
	T-value	-0.59	53.13		3.75	32.26	2.40	3.32		2.56	4.46	2.21	3.50	5.11	8.00	7.90	66.11	4.42			
	VIF				1.40	1.15	1.78	1.73		1.13	1.35	1.07	1.10	1.15	1.21	1.38	1.23	1.49			
NST/R 1	Coefficient	143.21			2062.77		432.54	30.42		490.45	46.79			90.93	126.71	99.91	1900.68	98.79	0.63	0.63	0.48
Foodstuffs and animal fodder	St. coefficient		516.21		49.84		220.20	26.05		38.53	20.87			36.29	55.71	40.06	699.94	34.03			
	T-value	8.74	59.85		5.00		21.45	2.71		4.17	2.37			3.69	5.50	3.89	75.23	3.31			
	VIF				1.40		1.78	1.71		1.13	1.27			1.12	1.19	1.36	1.22	1.48			
NST/R 2	Coefficient	61.86			645.58	47.19	10.14				11.00			6.60	10.13		70.60		0.16	0.16	0.14
Solid mineral fuels	St. coefficient		81.61		15.60	4.35	5.16				4.91			2.63	4.45		26.00				
	T-value	32.12	68.79		11.99	3.62	3.97				4.06			2.11	3.52		21.25				
	VIF				1.26	1.08	1.47				1.21			1.09	1.11		1.13				
NST/R 3	Coefficient	5.81					7.79			21.68				5.14	7.68	10.55	113.27		0.48	0.48	0.28
Petroleum products	St. coefficient		21.79				3.97			1.70				2.05	3.38	4.23	41.71				
	T-value	5.28	34.87				6.18			2.64				2.90	4.68	5.95	63.53				
	VIF						1.28			1.07				1.06	1.09	1.06	1.15				
NST/R 4	Coefficient	0.71			663.66		6.93	2.09	14.08	50.00				6.09	8.20	7.76	106.91	16.45	0.59	0.59	0.42
Ores and metal waste	St. coefficient		21.81		16.03		3.53	1.79	1.31	3.93				2.43	3.60	3.11	39.37	5.67			
	T-value	0.69	37.64		23.68		5.13	2.76	2.01	6.27				3.68	5.32	4.49	61.16	8.08			
	VIF				1.43		1.76	1.70	1.23	1.14				1.09	1.17	1.31	1.26	1.52			
NST/R 5	Coefficient	24.88		1.35			75.32			53.33				9.61	23.96	13.72	231.71		0.46	0.46	0.33
Metal products	St. coefficient		75.82	3.85			38.34			4.19				3.84	10.54	5.50	85.33				
	T-value	8.76	49.86	2.39			24.57			2.67				2.22	5.88	3.03	53.38				
	VIF			1.88			1.29			1.07				1.19	1.38	1.47	1.15				
NST/R 6	Coefficient	152.29		3.27	1895.16	75.83	49.20	8.66		157.47	26.80			29.68	38.36		347.40		0.34	0.34	0.23
Crude and manufactured minerals, building materials	St. coefficient		240.70	9.28	45.79	6.99	25.05	7.42		12.37	11.95			11.84	16.87		127.93				
	T-value	26.04	76.51	2.71	12.89	2.19	7.04	2.06		3.80	3.73			3.57	5.01		38.32				
	VIF			1.88	1.34	1.13	1.59	1.79		1.07	1.26			1.15	1.27		1.19				
NST/R 7	Coefficient	-11.20						4.80		89.42					5.63	9.15	182.79	12.37	0.49	0.49	0.44
Fertilizers	St. coefficient		10.43					4.11		7.03					2.48	3.67	67.32	4.26			
	T-value	-7.66	10.40					3.86		6.58					2.33	3.32	64.47	4.06			
	VIF							1.49		1.12					1.12	1.27	1.10	1.11			
NST/R 8	Coefficient	75.29			1223.24		226.40			505.77				63.00	66.05	39.51	1036.93	80.14	0.61	0.61	0.35
Chemicals	St. coefficient		263.77		29.55		115.25			39.74				25.14	29.04	15.84	381.86	27.61			
	T-value	8.57	53.22		5.22		20.06			7.49				4.47	5.05	2.79	71.64	4.69			
	VIF				1.37		1.68			1.13				1.06	1.09	1.07	1.20	1.47			
NST/R 9	Coefficient	833.57			17695.0	2606.39	2356.93			2236.74	321.83	328.43	448.41	632.00	881.46	722.48	5584.92	357.48	0.61	0.61	0.40
Machinery, transport equipment, manufactured articles and miscellaneous articles	St. coefficient		2715.15		427.52	240.17	1199.85			175.73	143.52	86.26	137.58	252.21	387.52	289.68	2056.70	123.15			
	T-value	9.45	84.18		11.59	7.19	31.90			5.09	4.36	2.29	3.49	5.71	8.30	6.45	59.04	3.21			
	VIF				1.37	1.15	1.71			1.13	1.35	1.07	1.09	1.12	1.13	1.14	1.23	1.49			
NST/R -1	Coefficient	-64.48						16.55		324.10					22.20	40.93	834.06	36.62	0.51	0.51	0.41
Empty trips	St. coefficient		28.19					14.17		25.46					9.76	16.41	307.15	12.62			
	T-value	-10.30	6.57					3.11		5.57					2.15	3.47	68.70	2.81			
	VIF							1.49		1.12					1.12	1.27	1.10	1.11			

E

Practical Results NST/R

Table E.1: Coefficients for explanatory variables of Freight Trip Generation per NST/R

	Trips/Workweek/Zone *N = 5152*	Intercept	Population	Surface area distribution centers	Employment: Agricultural	Employment: Industrial	Employment: Services	Employment: Others	Terminal: Rail	Terminal: Inland waterway	Urban density: 2	Urban density: 3	Urban density: 4	Urban density: 5	Urban density: 6	Surface area: Business	Surface area: Industrial	Surface area: Office
	Unit	-	*x1000*	*km²*	*x1000*				-	-	-	-	-	-	-	*km²*		
Production	**Agricultural products and live animals**	-5		173	312	3		2		7		5	8	11	11	94	8	
	Foodstuffs and animal fodder	11		144	28	77				7				5		112	7	
	Solid mineral fuels	2	0			8	1			5						5	1	
	Petroleum products	0				3									1	7		
	Ores and metal waste	0		109		5			5	1			0	1	1	8	1	
	Metal products	2	0			18										17		
	Crude and manufactured minerals, building materials	5				27	2			12		2	3	4	3	22	3	
	Fertilizers	0				0										10	1	
	Chemicals	5		145		61			38				5	4		79	6	
	Machinery, transport equipment, manufactured articles and miscellaneous articles	42		1867	413	448			186	44		29	42	58	47	451	29	599
	Empty trips	-13													12	233	16	
Attraction	**Agricultural products and live animals**	-2		169	224	6		2	19	7	4	6	8	12	12	73	6	
	Foodstuffs and animal fodder	10		262		59		2	43	6			6	9	6	131		
	Solid mineral fuels	3	0			7	1			5						5	1	
	Petroleum products	0				3									1	9		
	Ores and metal waste	0		114		3		0	5	1			0	1	1	8	1	
	Metal products	3	0			12								1		18		
	Crude and manufactured minerals, building materials	5		39		23	1	1		11		2	3	4	3	25	2	
	Fertilizers	-1						0	5					0	1	11	1	
	Chemicals	8		232		44			47				5	5		84	5	
	Machinery, transport equipment, manufactured articles and miscellaneous articles	60		2285	284	352			224	44		30	42	63	51	470	27	518
	Empty trips	-20						5	99					7	12	254	11	

Table E.2: Coefficients for explanatory variables of Freight Generation per NST/R

	Tonnes/Workweek/Zone *N = 5152*	Intercept	Surface area distribution centers	Employment: Agricultural	Employment: Industrial	Employment: Retail	Employment: Services	Employment: Others	Terminal: Container	Terminal: Rail	Terminal: Inland waterway	Urban density: 2	Urban density: 3	Urban density: 4	Urban density: 5	Urban density: 6	Surface area: Business	Surface area: Industrial	Surface area: Mixed	Surface area: Office
	Unit	-	*km²*	*x1000*					-	-	-	-	-	-	-	-	*km²*			
Production	**Agricultural products and live animals**	-91	665	3672				24			80		64	94	129	126	1316	105		
	Foodstuffs and animal fodder	96		304	1103	111					72						1371	78		
	Solid mineral fuels	-6			169	15	13				101						64	12	17	
	Petroleum products	-3			50											14	91			
	Ores and metal waste	-4	562		43					41	3				5	8	105	7		133
	Metal products	10			200												172			
	Crude and manufactured minerals, building materials	-10			531	59	33				251					29	300			
	Fertilizers	-4															165	14		
	Chemicals	36			688					334							857			
	Machinery, transport equipment, manufactured articles and miscellaneous articles	479		4588	5953		190			1731	565				265		4791			
	Empty trips	-43														38	765	52		
Attraction	**Agricultural products and live animals**	-8	875	1770	27			21		170	50	48	68	86	129	139	946	75		
	Foodstuffs and animal fodder	143	2063		433			30		490	47			91	127	100	1901	99		
	Solid mineral fuels	62	646	47	10						11			7	10		71			
	Petroleum products	6			8					22				5	8	11	113			
	Ores and metal waste	1	664		7			2	14	50				6	8	8	107	16		
	Metal products	25			75					53				10	24	14	232			
	Crude and manufactured minerals, building materials	152	1895	76	49			9		157	27			30	38		347			
	Fertilizers	-11						5		89					6	9	183	12		
	Chemicals	75	1223		226					506				63	66	40	1037	80		
	Machinery, transport equipment, manufactured articles and miscellaneous articles	834	17695	2606	2357					2237	322	328	448	632	881	722	5585	357		
	Empty trips	-64						17		324					22	41	834	37		

F

Use Case Stadshavens Rotterdam

This chapter discusses the application of the predictive model in another use case: Stadshavens Rotterdam. The use case discussed in this chapter showed to be less relevant than the one in the main report due to the having large amount of trip generation caused by a seaport.

F.1. Background

The Stadshavens Rotterdam project aims to transform and renew the Stadshavens Rotterdam area into an area combining innovative and modern working and living environments.

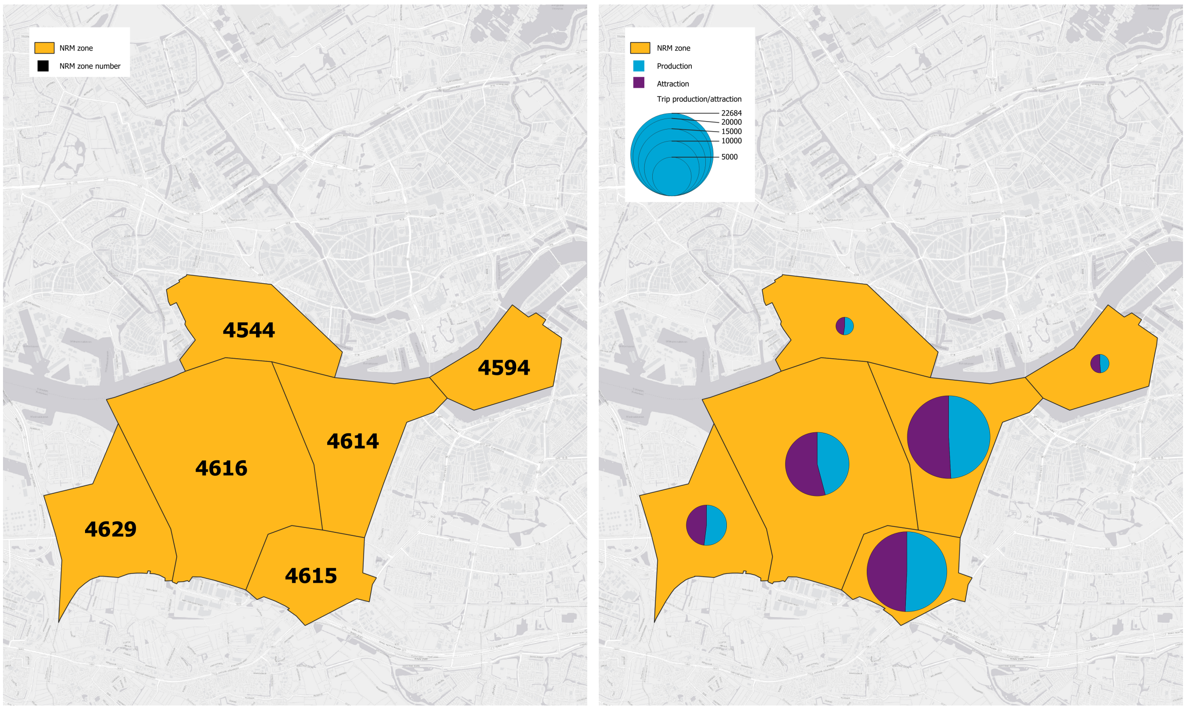
Current activities in Stadshavens will be broadened (diversified) and the business location climate will improve. This will strengthen the economic structure of the port. But also knowledge institutes and innovative companies are intended to be added to the area. Moreover, building housing in Stadshavens will increase the supply of inner-city housing to improve the living and working environment.

The project aims to add (Zaken, 2013):

- 10,000 new houses
- Education for at least 1,000 students
- 13,000 new jobs

F.2. Area Parameters

From the BasGoed freight tour data, which is used to predict the freight trips, the initial observed situation is determined regarding the trip production and attraction of the individual zones. These are distributed over the different NRM zones according to Figure F.1b. Furthermore, the corresponding NRM zones, covering the study area, are presented in Figure F.1a.



(a) NRM zone numbers of study area (b) Distribution of the trip generation

Figure F.1: Study area visualized using the model dataset

De total FTP and FTA of the different zones are the combination of different trips per Logistic Segment. How the distribution of segments over the total amount of production and attraction per zone is defined is shown in Table F.1. The table shows that LS 1 is overruling the other segments with more than 50% of all the trips. This segment contains general cargo, which also explains the size of trips containing this segment considering the study area has a seaport. Moreover, the total production and attraction are both the same size, meaning all the incoming trips also leave the area.

Production (trips/workweek)								
LS 0	LS 1	LS 2	LS 3	LS 4	LS 5	LS 6	LS 7	Total
3894	16493	4156	1243	1747	2213	265	1961	31972
Attraction (trips/workweek)								
LS 0	LS 1	LS 2	LS 3	LS 4	LS 5	LS 6	LS 7	Total
4143	17054	4316	1286	1751	2168	254	1974	32946

Table F.1: Current FTG of Stadshavens Rotterdam per LS

Considering the explaining factors found in the results, the explaining factors for the prediction of logistic movement are the surface area of distribution centers, employment in the agricultural, industrial, retail, services and others sector, the presence of a rail or inland waterway terminals, the urban density level of 2, 3, 4, 5 and 6 and the surface area of business, industrial and offices in the area. From the base dataset, the BasGoed freight tour data, CBS and zonal data, the parameters of the corresponding NRM zones for these variables are presented. This is shown in Table F.2.

NRM zone	4544	4594	4614	4615	4616	4629	Total
Surface area distribution centers (km^2)	0	0.013	0.018	0.042	0.037	0.020	0.130
Employment agricultural ($x1000$)	0.002	0.001	0	0	0	0	0.003
Employment industrial ($x1000$)	0.423	0.304	2.112	1.765	0.810	0.733	6.147
Employment retail ($x1000$)	0.451	0.208	0	0.006	0.022	0	0.687
Employment services ($x1000$)	0.798	5.400	0.539	0.716	0.452	0.056	7.961
Employment other ($x1000$)	1.297	4.105	2.883	3.270	2.704	0.253	14.512
Terminal rail	0	0	0	1	0	0	1
Terminal inland waterway	1	1	1	0	1	1	5
Urban density level	6	6	6	6	6	5	1x5 | 5x6
Surface area business (km^2)	0.657	0.036	0.849	0.768	3.212	0.859	6.381
Surface area industrial (km^2)	0.042	0	0	0	0	0	0.042
Surface area office (km^2)	0.005	0	0.051	0.015	0.006	0	0.077

Table F.2: Current situation Stadshavens Rotterdam

As seen in the table, the surface area of business is relatively big in the study area. To visualize this, an overview of the surface areas is depicted in Figure F.2, where the port area in zone 4616 is visible.

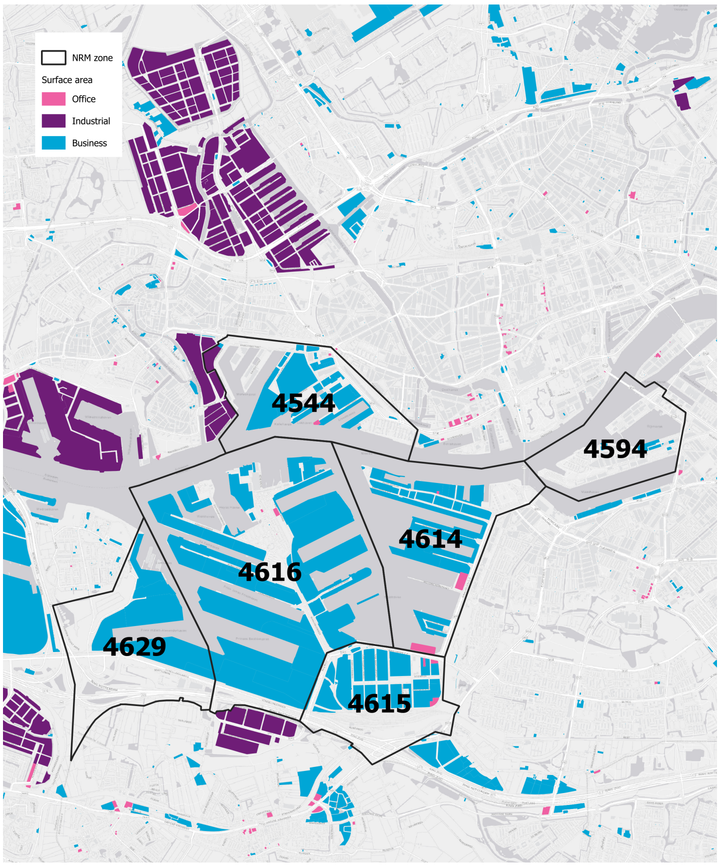


Figure F.2: Surface area distribution in study area

F.3. Increase Estimation

The project specifications, being 10,000 new houses, education for at least 1,000 students, and 13,000 new jobs, are transformed into a coefficient for the explaining factors. For each project goal, the implications for the area are described and the results are shown in Table F.3.

- 10,000 new houses
  The implication for 10,000 new homes in logistic movements will be that the urban density level of all 6 zones will increase to level 6 for all zones. This, however, will decrease the amount of movements according to the explaining factors. This is added to the table as 1-time urban density level 6 and -1-time urban density level 5.
- Education for at least 1,000 students
  The implication for at least 1,000 new students is assumed to not affect the logistic movements in the area. The implications of the increase in students are supposed to be found in the new jobs.
- 13,000 new jobs
  The implication for 13,000 new jobs in the study area is calculated using both the surface area and the employment. Since distribution centers, business, industrial and offices all have a relation with employment, the amount of employees per km^2 is calculated using Equation F.1. This factor is then used to calculate the increase in surface area, see Equation F.2. These two values (increase of new jobs and increase of surface area) are then added to the designated variable according to the total weight of the employment or surface area, respectively. The results are shown in Table F.3.

$$\text{Employees per surface area} = \frac{\text{Total employees}}{\text{Total surface area}} = \frac{29,310}{6.63} = 4,420.40 \text{ employees}/km^2 \tag{F.1}$$

$$\text{Increase in surface area} = \frac{13,000 \text{ new jobs}}{4,420.40 \text{ employees}/km^2} = 2.94km^2 \tag{F.2}$$

	Current situation	Increase
Surface area distribution centers (km^2)	0.130	0.058
Employment agricultural ($x1000$)	0.003	0.001
Employment industrial ($x1000$)	6.147	2.726
Employment retail ($x1000$)	0.687	0.305
Employment services ($x1000$)	7.961	3.531
Employment other ($x1000$)	14.512	6.437
Terminal rail	1	0
Terminal inland waterway	5	0
Urban density level	1x5 | 5x6	6x6
Surface area business (km^2)	6.381	2.830
Surface area industrial (km^2)	0.042	0.019
Surface area office (km^2)	0.077	0.034

Table F.3: Current and expected situation of the study area

F.4. Model Application

After the initial parameters and the expected parameters have been defined, the model and its explaining factors are applied to the study area.

As a reference, the FTG is calculated using the input parameters from Table F.2 and the explaining factors found in Table 6.6. This results in a total production of 12264 trips/workweek and a total attraction of 12375 trips/workweek of the study area (also see Table F.4). The prediction from the model shows an error of roughly 20000 trips/workweek in both production and attraction in comparison to the results from the initial dataset (Table F.1) which is quite significant. An explanation for this would be that other factors influence the production and attraction in this study area or particular port activities generate a load of trips which are not covered by the model which follows a trend considering all the NRM zones.

F.5. Results

However, this application was intended to show the increase in logistic movements and not validate the current model. Therefore the expected increase from Table F.3 is applied to the explaining factors. This resulted in a total production increase of $4,923.67$ trips/workweek and a total attraction increase of $4,773.63$ trips/workweek of the study area (also see Table F.4). This is a significant increase of roughly 1/6th of the current movements in the area.

Table F.4: Results of predicting Freight Trip Generation for the study area

	Trips/Workweek/Zone N = 5152			Employment					Terminal		Urban density					Surface area			
		Intercept	Surface area distribution centers	Agricultural	Industrial	Retail	Services	Others	Rail	Inland waterway	2	3	4	5	6	Business	Industrial	Office	
	Unit	-	km²	x1000					-	-	-	-	-	-	-	km²			
Production	Food (general cargo)	12	230		105	7				9			6	8		110			
	Miscellaneous (general cargo)	18	431	419	261					43			30	45	50	593	62		
	Conditioned transport	-1	175	329	40					11		7	11	15	15	141	10		
	Facility logistics	3	98		35		1		23					3	3	37			
	Construction logistics	12	125		91							5	6	8	6	40		107	
	Waste	2	924		48		2		59	8			4	6	6	62	11		
	Parcel (consolidated flows between sorting centres)	0	104		6		0		6	1				0	1	6	1		
	Hazardous materials	6			58								5	4		65			
	Total explaining factors	53	2087	748	644	7	3	0	88	72	0	12	61	90	81	1054	84	107	
	Input values current situation	-	0	0	6	1	8	15	1	5				1	5	6	0	0	**Total FTP current situation**
	FTP current situation	**317**	**271**	**2**	**3960**	**5**	**25**	**0**	**88**	**362**	**0**	**0**	**0**	**90**	**405**	**6727**	**4**	**8**	**12264**
	Input values increase	-	0	0	3	0	4	6						-1	1	3	0	0	**Total increase FTP**
	Increase FTP	**0**	**121**	**1**	**1756**	**2**	**11**	**0**	**0**	**0**	**0**	**0**	**0**	**-90**	**81**	**2983**	**2**	**4**	**4871**
Attraction	Food (general cargo)	14	365		81				37	8			8	12	7	125			
	Miscellaneous (general cargo)	20	957	288	196			12	248	37		20	36	54	52	608	44		
	Conditioned transport	3	228	227	33			2	35	10		7	10	15	13	132	5		
	Facility logistics	2	139		25		1		27	1	2	3	3	6	5	41			
	Construction logistics	14	217		74							5	6	10	8	48		106	
	Waste	4	915		37		2		59	9			4	7	7	66	13		
	Parcel (consolidated flows between sorting centres)	1	105		5		0		6	1			0	1	1	6	1		
	Hazardous materials	8			44				24				5	4		66			
	Total explaining factors	66	2926	515	496	0	3	15	438	65	2	36	74	107	92	1092	64	106	
	Input values current situation	-	0	0	6	1	8	15	1	5	0	0	0	1	5	6	0	0	**Total FTA current situation**
	FTA current situation	**397**	**380**	**2**	**3046**	**0**	**23**	**213**	**438**	**326**	**0**	**0**	**0**	**107**	**462**	**6970**	**3**	**8**	**12375**
	Input values increase	-	0	0	3	0	4	6	0	0	0	0	0	-1	1	3	0	0	**Total increase FTA**
	Increase FTA	**0**	**170**	**1**	**1351**	**0**	**10**	**95**	**0**	**0**	**0**	**0**	**0**	**-107**	**92**	**3091**	**1**	**4**	**4707**