GLOBAL AIR CARGO FLOWS ESTIMATION BASED ON O/D TRADE DATA

L.J.J.A. MEIJS

TU Delft

SEABURY CONSULTING
now part of accenture

GLOBAL AIR CARGO FLOWS ESTIMATION BASED ON O/D TRADE DATA

L.J.J.A. MEIJS

to obtain the degree of Master of Science in Aerospace Engineering
at the Delft University of Technology
to be defended on Thursday September 7^{th}, 2017 at 2:30 PM.

Student number:	4244028	
Thesis committee:	Prof. dr. Richard Curran,	Faculty of Aerospace Engineering
	Dr. ir. Bruno F. Lopes dos Santos,	Faculty of Aerospace Engineering
	Prof. dr. ir. Lóránt Tavasszy,	Faculty of Technology, Policy, and Management
	Dr. Daniele Ragni	Faculty of Aerospace Engineering
	Daniel dos Reis Miranda,	Seabury Consulting

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

PREFACE

After a successful research at Seabury Consulting about wind implications on the air cargo capacity of an aircraft in the Summer of 2015, the opportunity arose of doing a graduation research project in the cargo field. The original idea Seabury had, was to solve a puzzle which looked like a big Sudoku in the first place, but has appeared to be an multi dimensional Rubik's Cube. At least Seabury was the one with the right data to get started, so the idea had to be converted into a proposal with a strong framework to get to an academic research project. From the moment Lori got involved and brought some fundamental methodologies to approach the problem, the kick-off was a fact.

The constant awareness of both developing an academic supported model, but also keeping in mind that this was the first step of a potential commercial product, was very challenging and kept me motivated. The fact that I needed to learn Java in order to create a model which estimated air cargo flows took some time, but was rewarding in the end. After a year of programming, calibrating, validating and analyzing, Java is still not my best friend, but the model and I can get along pretty well.

I like to thank Daniel for his daily support and his industry knowledge which helped me to put the enormous amount of data into the right perspective. Further, I like to thank Bruno for the guidance during this research and the clear and critical feedback during the meetings, and Lori for all the help and the enthusiasm he brought into this project.

After exactly ten years of studying in which I planned and executed my learning curve pretty steep but gradually, I'm very proud to deliver this Master research. The developed Air Cargo Flow Model has, from my point of view, some high potential from both an academic as a commercial perspective.

Lars Meijs
August 2017

SUMMARY

INTRODUCTION

Where a relatively small part (less than 1%) by volume of world trade is represented by air transport, the orientation towards high value products results in a share of almost 35% of international trade in value. But where data of the amount of trade between countries is highly available, the actual routing of those traded goods is not known at the moment. This means that up until today it is unknown for airlines, airports, forwarders and other important stakeholders how trade between countries have been flown.

This research includes the development of a model to estimate yearly air cargo flows over the global air transport network by using country-to-country air trade data. The novelty in this model lies in the fact that this is the first model in academic literature that estimates air cargo flows in such a detailed airport-to-airport level. Both on the demand side of the model as on the supply side very granular data is used. O/D air trade data on country-to-country level is the main demand input, where route choices are based on scheduled flight services reported by the airline, serve as supply input.

MODEL FRAMEWORK

The model includes a total of three building blocks which form the framework. For each O/D country pair a choice set of routes by a shortest path algorithm from airport to airport, based on actual scheduled airline services. The possibility of intermediate transshipments and trucking connections are included. By using a path size logit model together with a generalized cost function, the route choice probabilities are calculated and applied to the O/D trade data for the distribution of the flows. The airport impedance parameter A_p and the scaling parameter μ are introduced for calibration purposes. The scaling parameter μ determines the impact that the cost differentials have on the percentage of trade assigned to each alternative. The parameter A_p includes all relevant, measurable and hidden service characteristics of airports, such as fuel costs, airport charges, handling costs, congestion costs, etc.

model data & calibration

The model uses actual air trade data between two countries on the demand side, and scheduled airline services on the supply side. Other network data used in the model are: county nodes which captures the weighted average location of all air freight moving to, or from the country; Road Feeder Services to allow the model to also use airports from neighbouring airports to import/export trade. The main network attributes of time and costs for transhipment and transportation are presented in the model parameters which were obtained from academic literature and industry analysis.

Calibration of the model includes the adaption of the scale parameter μ and the airport impedance parameter A_p. The main objective of this calibration is to minimize the absolute difference between the observed throughput reported by ACI and the modeled throughput estimated by the model. Validation of the model will be done by using BTS data and the Seabury Cargo Capacity Database.

us - brazil

The US-BR case study had the purpose to prove the capabilities of the model to reproduce air cargo flows from airport to airport and yearly airport throughput. First, a detour factor an capacity share analysis was performed from which we could conclude that potential interference of O/D trade is minimized within the trade market between the US regions and Brazil. Therefore it could be used as a proper case study.

Secondly, the case study was set-up, which includes six US and six Brazilian airports and a service network to distribute the trade between those countries. For calibration purposes, assumptions had to be made to determine the observed throughput data for the twelve airports. After calibration, validation analysis showed that the model is able to reproduce the airport throughput very accurate, with a R-squared of 95%. Also the air cargo flows between the airports are estimated rather accurately.

Furthermore, verification of the model has proven that the model is robust when parameters are changed. From this we could conclude that the model is able to estimate air cargo flows with an acceptable accuracy and can be used to apply on a global scale.

FULL WORLD

The Full World model is a model which estimates global air cargo flows based on O/D trade data. First, a set of airports, countries and flight services was created by defining three selection criteria. Those selection criteria created a set of 318 airports, representing 148 countries and a network consisting out of +7,500 distinct flight routes. With this it captured 96% of network capacity, 99% of airport throughput and 98% of air trade and was able to estimate the air cargo flows on a global scale within one minute. After converging of the model due to calibration, it resulted in a fit which explains 88% of the amount of variation between airport throughput volumes. With this high R-squared value, we could conclude that the model is able to reproduce airport throughput rather accurately.

By analyzing the correlation graph which presented the modeled throughput compared with the observed throughput, we identified four group of airports, namely, top 20 airports, integrator airports, 'island' airports and overestimated airports. We also discussed the airports which are located near the assumed cargo center of gravity.

Most of the conclusions drawn from this analysis are due to two main factors, namely, the use of the shortest path algorithm and the method of calibration. By using the shortest path algorithm, only one route between an airport pair will be added to the choice set. This narrows down the amount of possibilities to distribute the flows and will capture transshipment and island airports to a lesser extend. The number of direct routes and available airports for the distribution of trade is a determining factor in the current model to reproduce airport throughput when a shortest path algorithm is used. The calibration method of minimizing the sum of absolute gaps between observed and modeled airport throughput forces the model to focus on the more larger airports, for the most efficient approach. This results in a general overestimation of smaller airports, because NSGA-II also minimizes the difference between observed and modeled transshipments. Because in the current model only for the top 14 the (estimated) observed transshipment is inserted, the calibration will be forced towards a value of zero for the rest of the airports, which may not be the case.

The development of the Air Cargo Flow Model including the research from this report is the accomplishment of the main objective of this project. The work presented give airlines and other stakeholders in the aviation industry sufficient insight into the now largely unknown air traffic flows of the competitive air cargo market. This helps them to understand the behavior of competitor airlines, balance the capacity and for route development in order to increase the profitability. On top of that, it will also help them with market analysis, pricing strategies and new route analysis to improve their network.

Further research, which now becomes available by the introduction of this version of an air cargo flow model, will allow us to study the current and future state of air cargo flows on a global scale. Recommendations in order to improve and extend the model are mainly focused on two topics. Firstly, including the k-shortest path algorithm to make the allocation of flows more realistic and to increase the reliability of the calibration process. Secondly, adding capacity constrains on routes or trade lanes to improve the quality of the model.

Lastly, further research into the airport impedance parameters is recommended. Especially when the model is extended with the k-shortest path and a capacity constrain, the quality of the Ap-values will increase. There can be investigated if and how they are related to each other and how the correlation is between airport characteristics known from public data and the assigned impedance value by the model.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

The air cargo industry plays an import role in world trade, doubling the volume flow every 10 years since 1970 (Chang and Yeh, 2007). According to Boeing and Airbus, this trend is expected to continue over the next 20 years with an annual average growth rate between 4% and 6% (Boeing, 2014; Airbus, 2015). Although a relatively small part (less than 1%) by volume of world trade is represented by air transport, the orientation towards high value products results in a share of almost 35% of international trade in value (IATA, 2015). On top of that, between 1995 and 2004 air cargo transport has grown about 50% faster than passenger transport (Wong et al., 2009). As a result, many passenger carriers place strategic emphasis on the growth of cargo business along their passenger route networks (Zhang et al., 2004).

The ongoing competition between airlines force them to manage their air cargo operations efficiently and focus on strategic decisions to stay flexible for unexpected changes. Those challenges resulted in an increasing amount of research on air cargo operations since 1990 (Feng et al., 2015). One of those challenging problems is the imbalance of cargo capacity. This occurs, for example, when trade goods are being produced in country A and consumed in country B. The air cargo would be in this case "unidirectional" in terms of demand. Flying the aircraft back to country A with an empty hold would result in high operating costs and a weak competitive position.

Unfortunately, the actual routing of trade flows are not available at a global scale. This means that up until today it is unknown for airlines, airports, forwarders and other important stakeholders in the aviation industry how trade between countries have been flown. Every single stakeholder has the isolated view of its operations but there is still a lack of understanding on the global flows. Among other reasons, this can be explained by the complexity of air cargo modelling compared to passengers or vehicles in urban networks for example. On top of that, most real-world problems are lacking in having the right data, both for input to run the models as to validate their results (Jourquin and Beuthe, 1996; Jourquin, 2005; Maia and Couto, 2013). Other research do manage to model air freight but are missing level of granularity on the demand and the supply side (Heinitz and Meincke, 2013; ITF, 2015; Newton, 2009).

In the world were data-driven decision making emerged as a key approach to solving problems arising, it is hard to believe that air cargo stakeholders do not have sufficient insight into the competitive cargo traffic flows to make the right strategic decisions. Available data sets in air transport published as aggregated high-level figures from countries, airports and airlines, together can serve as valuable pieces to solve the puzzle of estimating those air cargo traffic flows.

1.1 RESEARCH OBJECTIVE & RESEARCH QUESTIONS

Based on the aforementioned research problems stated, the main objective of this research is formulated as follows:

> *To contribute to the development of a model which estimates the air cargo flow per flight leg, flown by a certain airline and equipment, on a worldwide basis by providing a best-fit model which generates low-level estimates by combining and allocating high-level figures, on yearly basis.*

The main research question is formulated as follows:

> *How can the air cargo flow per flight leg on a worldwide basis be estimated from aggregated traffic data?*

This main research question is sub-divided into a set of central questions. By answering these questions the problem will be clarified and a model can be developed. The questions are formulated as follows:

1. **Which data regarding air cargo transport is available in order to estimate air cargo flows?**
   Because we noticed during the literature study that in order to take some first steps in developing an air cargo flow model which estimates at a certain detail, one needs accurate input data. This is true for both the supply side as for the demand side. The right data has to be found to meet these requirements and to serve the main research question.

2. **Which solution technique will be used to estimate the air cargo flow on a flight leg basis?**
   The estimation of air cargo flows on a global scale with granular input data is never reported in current academic literature. Therefore techniques filtered out from the literature review will be used to develop a new model.

3. **How and which data should be used in order to test the model on its quality?** The output itself of the to be developed model alone will not give enough information about the quality of the model. Therefore calibration is needed to perform verification and validation in a later stadium of the project. Case studies will be developed and be analyzed.

4. **Which steps have to be taken to develop the model further?** Because this research will be the first step in the development of an air cargo flow model, possible improvements and extensions can be recommended for further research.

1.2 RESEARCH SCOPE

The model which will be presented in this research will be a simplified representation of reality. Several details are omitted to achieve this simplification. To define the purpose of the model including its boundaries, a certain scope is created. The scope of this research is formulated by some model characteristics and is based on Sun and Wang, (2015).

1. **Scale** The research considers international air freight transport on the global level. The definition of air freight included traded commodities that use this particular mode of transport, thus leaving air mail outside of the scope.

2. **Objective** The model assumes that flows are allocated to an air transportation network from a profit maximization approach from the forwarder.

3. **Modes** Besides the air freight transportation mode, also Road Feeder Services are included in the model.

4. **Commodity** The model will only consider air trade as a whole and will not consider a split into commodities. It will be measured in yearly metric tonnes.

5. **Network Resources** The model will not take network capacity into consideration. Links and nodes are assumed to have unlimited capacity.

6. **Data Assumption** The model will be assumed as a deterministic model with fixed value of certain data (i.e. distance, demand, cost).

7. **Perspective** The perspective of the forwarders will be used. This because they have the knowledge of the main trade route alternatives between the countries by using the different transport modes.

1.3 RESEARCH RELEVANCE

The research in this thesis has the purpose to have both scientific and a real-world contribution. The following subsections describe these contributions respectively.

1.3.1 *Scientific contribution*

The main contribution of this thesis scientifically is in the field of modelling freight transport, specifically in air freight. The developed model in this thesis fills the gaps identified in the research area of modelling global air freight transport.

The novelty of this work is in the fact that it uses actual air trade data between two countries on the demand side, and scheduled airline services on the supply side. Due this combination and to the fact that the estimated flows are calibrated against observed traffic data, the model is unique it its field of air freight transport modelling. The granular demand and supply data used, allowed us to create an accurate model, while it is still applicable on a global scale.

1.3.2 *Industry contribution*

The models and results produced by this research are intended to support market analysis processes of different aviation stakeholders which deal with problems related to air freight transport. We have identified several of those stakeholders that may utilize the results of this research and the way they can benefit from this research.

Airlines

Each airline is very much aware of their operation performance in terms of the amount of FTKs they fly and the O/D markets they offer on their flights. However, this is an isolated view because airlines do not have the insight on how the competition is performing on the same route. Neither that they have an understanding of how many tonnes of air traded goods flow on specific routes. This information could be very helpful when making decisions regarding opening new routes or even with complete network changes.

Airports

Also airports are dealing with an isolated view on air cargo performance. They know to which regions they connect to in terms of airline services, but lack in having the information about what or how the competition is doing. An airport is highly interested in being a well connected node within a network which contains interesting trade markets in order to increase their performance. To attract new markets to connect to, it will be very valuable to have insight into trade routes and competing nodes nearby.

Forwarders

A forwarder, who arranges the shipping from point A to point B, wants to have the most complete overview of route options to deliver the shipment as cheap and fast as possible. Forwarders constantly face challenges because of a fast rate of change in those options. Mode choices can change because of competition between rail, ocean and air transport shipping but also competition between forwarders creates uncertainty. Unfortunately a global insight in optimal routes to transport goods from one city to the other is not available at the moment.

All of these challenges, which are dealt with on a management and/or strategic level within the stakeholders company, will be coped with in the model presented in this research or in future model extensions.

1.4 RESEARCH STRUCTURE

Chapter 2 includes the literature study of this research project, focused on the topic of estimating traffic flows. All relevant foundings from academic literature provide a solid basis to developed a new air cargo traffic estimation model.

Chapter 3 presents the framework of the model. First, the requirements are stated, then the assumptions of the model are defined coming from the requirements. Secondly, the building blocks which form the framework, are explained.

Chapter 4 explains the data and parameters used to run the model. Also the calibration methodology is described.

Chapter 5 includes a case study where trade is distributed from the U.S. regions towards Brazil. The purpose of this is to illustrate the capabilities of the model to reproduce air cargo flows from airport to airport and yearly airport throughput. On top of that, before applying the model on a global scale, some validation and verification checks are performed to ensure the ability and robustness of the model.

Chapter 6 includes the Full World model where trade is distributed on a global scale. This chapter contributes to literature of global air freight transport modelling.

Chapter 7 presents the conclusions along with several recommendations for future improvements of the model.

2

FREIGHT TRANSPORT MODELLING: STATE-OF-THE-ART

This chapter includes the literature study of this research project, focused on the topic of estimating traffic flows. All relevant foundings from academic literature provide a solid basis to developed a new air cargo traffic estimation model.

To get insight into possible techniques which can be used to estimate flows on a global scale, the main objective of this literature review is to give an overview of the relevant in different fields of research where aggregated data is used to estimate its dis-aggregated figures. Firstly the commonly used four-stage model will be explained with the focus on traffic assignment in freight transport. Secondly we will focus more on estimation models which are able to perform on a global scale since this the aim of our research scope.

2.1 TRANSPORT MODELLING

Since 1950 the basic concepts of estimating route and link flows in transport modelling have found their evolution in research. It has been widely used in all kind of transportation networks, from urban traffic to freight distribution networks.

The traffic estimation process to predict route and link flows, traditionally follows four sequential stages: *trip generation, trip distribution, modal split* and *traffic assignment*. This four-stage model has been recognized as a standard in transport modelling. Some studies in freight transport even added a fifth stage just before the assigning the traffic, to identify and incorporate elements that influence and shape the route choice (Tavasszy, 2006). The output for each stage in the model is used as an input for the stage that follows, and the link and route flows from the last stage are used as feedback for the framework previous stages (Figure 1).

Within the trip generation stage, the amount of trips between an origin (trip production) and destination (trip attraction) at each zone in a network are determined. Next comes the step of trip distribution, were the generated trips are connected to form an origin-destination (O/D) matrix to see how the trips are distributed. In the third step, the demand for each O/D-pair is partitioned into different travel modes. The traffic assignment is the fourth and last step of the traditional modelling methodology. Flows from the O/D matrix will be assigned onto routes and links of the transport network.

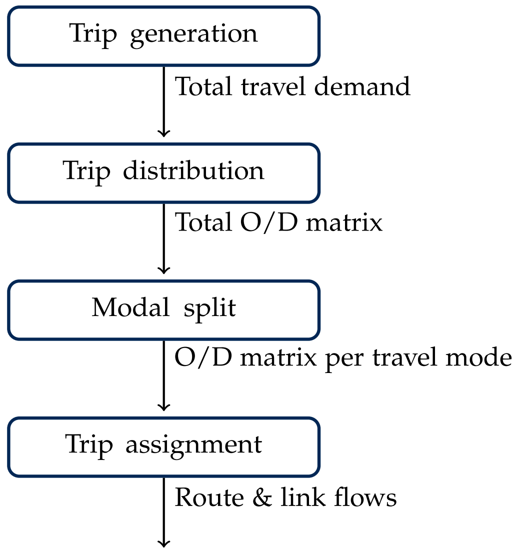

Figure 1: The traditional four-stage model (Ortuzar and Willumsen, 2011)

In our case we have aggregated data of trade between the true origin and final destination country, which will serve as an O/D matrix. Because the available trade figures representing *air* trade, the modal split stage is not applicable to our problem.

Therefore, in our search to the traffic flows in the air cargo network, we will focus in this section on the literature review on the different solution techniques for traffic assignment in the domain of transport modelling. This final step to estimate flows within a network is a well researched part in transport modelling since its output gives the most detailed and valuable information for strategic and tactical decisions. The main aim of this chapter is to give an overview of the different solution techniques which are used in previous research regarding the traffic assignment problem. Because the available aggregated figures do not vary in space and time, dynamic assignment techniques will not be discussed.

2.1.1 *Traffic Assignment*

As shortly discussed above, the major aim of the traffic assignment process is to estimate the volume of traffic on the links of the network. The basic inputs required for assignment are (Patriksson, 1994):

- an O/D matrix expressing estimated demand.
- a network, consisting of nodes, links and routes.
- principles or route selection rules thought to be relevant to the problem in question.

The output of the procedure depends on the developed model but it always includes the following:

- estimation of the link flows.
- performance of the link in terms of costs.

With those output a broad scope of a specific transportation system can be evaluated which forms the basis of strategic operational decisions or assessments in the various characteristics of the network.

2.1.2 *History of Traffic Assignment*

The fundamentals of traffic assignment evolved shortly after the end of World War II, when vehicle movements between zones became available for use in determining traffic loads on proposed new routes. At that time it was assumed that the flows on the links were independent of travel time and costs[1], which resulted in the allocation of *all* vehicles on a origin-destination route, or *no* vehicles on that origin-destination pair (Campbell, 1950). This technique is commonly known as *shortest path* or *All-or-nothing* (AoN) assignment.

A few years later it was noticed that the assumption used for the AoN assignment was very unnatural and it would be more realistic if the assignment technique would recognize that travel times and costs will increase with the flows on the links. Wardrop was the first who implemented the assumption that the travelers are minimizing their individual travel cost, known as the *user equilibrium* principle, or that they would choose their routes to minimize the travel time in the total transportation system, the so called *system optimum*. These two principles are still widely used in academic literature and are know as the two Wardrop principles (Wardrop, 1952).

In 1956 two important events took place for the advancement in traffic assignment. Firstly, Beckmann used nonlinear optimization and showed that the two Wardrop principles corresponds to the solution of convex nonlinear optimization problems subject to flow conservation constraints (Beckmann and Winston, 1956). Secondly, Frank and Wolfe published a paper for the solution of convex, quadratic optmization problems by using an iterative algorithm (Frank and Wolfe, 1956). It is still seen as one of the most major improvements on the standard iterative method and is as of present day serves as a standard code in transportation planning packages to solve traffic assignment problems.

Now that new traffic assignment techniques included that the link flows were dependent of time and costs, it was still a simplified model

1 the term 'costs' is mostly represented in literature by travel distances or fixed travel times

because the costs are deterministic, which means that they are fixed. This limits the user to make cognitive decisions and therefore it can be assumed that it will always take the best path in a network, which is not behaviorally sound when more alternatives are possible.This critic which can be found in literature introduces the motivation for stochastic models.

In stochastic models, user costs are considered as random variables which emphasizes the variability in drivers' perceptions of costs. The technique developed by Burrell in the sixties has been widely used for many years (Burrell, 1968). He used an uniform distribution of costs for each link in the network for a more realistic distribution than the AoN technique. Also Dial described a method based on a loading algorithm which splits trips between all possible exit nodes to create efficient paths (Dial, 1977).

When in a stochastic model the links are restrained in terms of capacity, the models developed seek an equilibrium condition where each user chooses the route with the minimum 'perceived' travel cost. This is therefore called: Stochastic User Equilibrium and was firstly introduced by Daganzo (Daganzo and Sheffi, 1977).

From this short description of historic evolvement in traffic assignment, two main factors can be filtered out, namely: the individual perception of the traveler and the possible congestion effects which can make some routes less attractive. These two factors define the classes where the considerable number and variety of traffic assignment models in present literature are in. The classic assignment techniques can therefore be classified as given in Table 1.

CAPACITY CONSTRAINT?	DETERMINISTIC	STOCHASTIC
No	All-or-Nothing	Pure stochastic Dial's, Burrell's
Yes	Wardrop's equilibrium	Stochastic user equilibrium

Table 1: Classification of Traffic Assignment techniques (Ortuzar and Willumsen, 2011) - adapted

In the following parts of this section we will further discuss the four big groups presented in the table. Because the literature available on the traffic assignment in air cargo networks is very limited, this literature study will broaden its scope by firstly reviewing basic techniques used in traffic modelling and after that exploring evolvement of those techniques in freight transport. In the last section of this chapter the different techniques will be discussed to see if can be of additional value for our research.

Traffic Assignment in Freight Modelling

In deterministic traffic assignment, it is considered that each user has the same perception of costs. No randomness is included so that the route choice decisions are uniquely based on the lowest generalized costs. These generalized costs can represent a lot of factors, like travel time, travel distance, fuel costs or congestion. The aim of the traffic assignment problem is to minimize these costs.

The most simple and basic traffic assignment technique is the *'All-or-Nothing'* (AoN) assignment, which already was shortly discussed in the previous section. When the available links in the network are not restrained by capacity, all users consider the same attributes to choose their route and also perceive and weight them in the same way, the AoN technique can assign the network. In this method the trips from any origin to any destination are loaded onto a single, minimum cost, link between them as given in the general notation for this type of traffic assignment shown in formula 1

$$x(i) = \sum x(j) \tag{1}$$

where x is the volume assigned to a link between nodes i and j when this link represents the route with the minimum costs (e.g. minimum travel time, shortest distance between nodes). The assignment is complete if the total volume has been assigned to a unique link.

Campbell started with the AoN assigent method and applied it in an urban transportation study (Campbell, 1950). After that, this technique is not used very often in freight models. Most of the literature is focussed on assigment of trucks, or jointly with passenger traffic, since freight traffic usally is only a small part of total traffic. However, a few models are worth mentioning regarding the AoN traffic assignment for freight modelling.

Jourquin used the AoN assignment when introducing a model for freight moving, loading and unloading, to analyze a multimodal freight transportation policy in Europe. The proposed optimization algorithm aimed to minimize the total transportation costs which was one of the techniques used to develop the NO/DUS software (Jourquin and Beuthe, 1996).

More recently, an article is published where the developed model assigned the intermodal cargo using the AoN technique (Maia and Couto, 2013). All traffic used the least costly path and using different transportation modes as a part of a freight traffic assignment model for multimodal (road and rail) networks.

Also Cascetta (Cascetta and Vitillo, 2013) developed a freight demand model for Europe which uses fixed costs for the road, rail and sea modes. By using fixed costs for the traffic assignment problem it implies the all-or-nothing technique.

The *shortest route problem* is a more advanced version of the AoN technique and is based on an algorithm developed in operations research. Through a network with fixed travel times (or costs or distance) on the links, it will find the route with the lowest travel time(Dijkstra, 1959). Southworth showed that this technique can still be used to generate routes to construct a multimodal network for international freight (Southworth, 2000). It covered five million origin- to destination freight shipments reported as part of a 1997 United States commodity flow survey.

If stochastic effects are ignored and one concentrates on capacity restraint to assign trips on a network, other methods will apply. As described before, models which take into account capacity restraint (or congestion) recognize that travel times and costs will increase with the flows on the links. Models in this category try to reach a point of equilibrium when no user is able to reduce their transport costs by choosing a different route. This was firstly outlined by Wardrop's first principle (Wardrop, 1952):

> *Under congested conditions drivers will choose routes until no one can reduce their costs by switching to another path*

This is also known as the *user equilibrium* (UE). The general notation for the UE traffic assignment can be given as shown in formula 2

$$\text{Minimize Z} = \sum_a \int_0^x t_a(x_a)dx, \tag{2}$$

$$\begin{aligned}
\text{subject to} \sum_k f_k^{rs} &= q_{rs} \quad , \forall\ r, s \\
x_a &= \sum_r \sum_s \sum_k \delta_{a,k}^{rs} f_k^{rs} \quad , \forall\ a \\
f_k^{rs} &\geq 0 \quad , \forall\ k, r, s \\
x_a &\geq 0 \quad , a\ \epsilon\ A
\end{aligned}$$

where k is the path, x_a equilibrium flows in link a, t_a travel time on link a, f_r^{sk} the flow on path k connecting O-D pair $r-s$, q_{rs} the trip rate between r and s and $\delta_{a,k}^{rs}$ is a definitional constraint and is given by:

$$\delta_{a,k}^{rs} = \begin{cases} \text{1 if a link a belongs to path k,} \\ \text{0 otherwise} \end{cases}$$

Wardrop also proposed an alternative way which is usually referred to as Wardrop's second principle:

> *Drivers will choose routes in such a way that the total (average) costs are minimized*

This is also known as the *system optimum* (SO).The general notation for the SO traffic assignment can be given as shown in formula 3

$$\text{Minimize Z} = \sum_a t_a(x_a)dx, \tag{3}$$

$$\begin{aligned} \text{subject to} \sum_k f_k^{rs} &= q_{rs} \quad ,\forall\ r,s \\ x_a &= \sum_r \sum_s \sum_k \delta_{a,k}^{rs} f_k^{rs} \quad ,\forall\ a \\ f_k^{rs} &\geq 0 \quad ,\forall\ k,r,s \\ x_a &\geq 0 \quad ,a\ \epsilon\ A \end{aligned}$$

Most of the traffic assignment techniques in this category are evolved from those equilibrium principles. Beckmann demonstrated that the UE can be formulated mathematically as a convex-optimization problem subjected to flow conservation constraints (Beckmann and Winston, 1956). Continuing in the research to optimization programming techniques, in the same year the most common algorithm for Traffic Assignment problems was developed, namely: the Frank and Wolf (FW) algorithm (Frank and Wolfe, 1956). This technique is based on nonlinear optimization and in each iteration of the algorithm, the linearized objective function is solved by assigning all traffic to minimum cost routes. The main advantage of the FW algorithm is the relatively small memory requirement because only total link flows need to be stored to evaluate the objective function. The disadvantage is the slow convergence rate when the optimal solution is approached. Therefore the papers of LeBlanc et al. became very popular because first he used the FW algorithm to solve a large scale network equilibrium assignment and in addition he significantly improved the convergence rate (LeBlanc and Pierskalla, 1975)(LeBlanc et al., 1985).

At the end of the nineties, there also came some first development in the deterministic, congested freight traffic assignment models. STAN, an network assignment method for the strategic analysis and planning of national rail freight transport systems, was developed in Canada (Crainic and Leal, 1990)(Guelat and Crainic, 1990), making use of an equilibrium assignment technique. This research, based on the Wardrop equilibrium, proofed that the technique is suited for networks with capacity limits. Cost penalties were given when traffic was close to the capacity.

In stochastic traffic assignment, it is considered that each user has perception of costs and they try to minimize a certain composite measure (distance, time, costs). To better simulate the behavior of an user, stochastic models introduce a certain randomness in the route choice to highlight the fact that drivers normally do not have the perfect information about network conditions.

As discussed in previous section, in a deterministic case without congestion the traffic assignment uses the *all-or-nothing* or *shortest path* technique. In stochastic models the routes will be distributed according to probability calculations to determine which route is chosen. In this way, contrarily to the *all-or-nothing* assignment, it ensures that not all the traffic will be allocated to the route with the least generalized costs. In literature this field of traffic assignment research is known as the *stochastic network loading problem*. In general, two methods can be distinguished in this field, namely: *simulation-based* methods and *proportion-based* methods (Ortuzar and Willumsen, 2011).

Burell used a Monte Carlo simulation to introduce variability in perceived costs and to create a distribution model(Burrell, 1968). The advantage of this simulation approach is that is reduces the number of second-best routes to be considered, which is one of the main problems compared to the deterministic case. For the proportion-based methods, the algorithm of Dial (Dial, 1977) has been seen as the basic approach. His algorithm is based on a multinomial logistic curve and calculates the percentage of trips realized from an origin to a destination through a specific route.

Both basic approaches were evaluated by Daganzo, which concluded that Dial's algorithm assigns too much traffic to sets of routes which overlap heavily and is therefore not adequate to use(Daganzo and Sheffi, 1977). The shortcoming of Burrell's method is that it samples link travel times only once for each origin which results in varying results from execution to execution. Despite the serious deficiencies associated to those models, there are still a few papers in literature who use the proposed basics to model freight transport (Jourquin, 2005) (Tsamboulas, 2007).

From the table presented earlier in this section we have now discussed three of the four classification groups. In the pure stochastic assignment technique routes were spread between origin and destination because of perceived routes costs. In the deterministic UE technique an equilibrium was reached because of capacity constraints. In realistic networks it is expected that a combination of both techniques will take place.

Where Daganzo critically evaluated the traffic assignment techniques when there is no (or light) congestion, he also introduces the *Stochastic User Equilibrium* (SUE)(Daganzo and Sheffi, 1977). It is defined as the state in which no user can improve his/her perceived travel time by unilaterally changing routes. With this approach the model assumes that drivers perception of costs on any given route are not identical and that trips between each O/D pair are divided among the routes with the most cheapest route attracting most trips.

The SUE problem is well covered in literature when applying on urban networks (Sheffi, 1985)(Thomas, 1991)(Patriksson, 1994)(Bell, 1995). However, SUE used in freight traffic assignment problems is, according to our best knowledge, very limited. (Maia and Couto, 2013) is the only one who describes a freight traffic assignment model which includes both stochasticity and capacity constraints. A stochastic multi-flow technique is combined with the AoN assignment to distribute general and intermodal cargo into the network. The capacity constraints are implemented in the model by gradually introduce traffic in the network. The mix of techniques applied in this study resulted in a SUE model which is simple an fast to run as well as it produced satisfactory results when used on a fictitious network.

2.2 GLOBAL FREIGHT MODELLING

In the research scope in the previous chapter we stated that we want to create a model which considers international air freight transport on the global level. Therefore, in this section of the literature review, we focus on freight models which are able to estimate flows on a global scale.

Worldnet is one of the few models in literature which tried to map air cargo flows on a global scale, with its research within the sixth framework programme called Worldwide Cargo Flows (Newton, 2009). Worldnet employs a top down approach and uses multimodal assignment and uses Eurostat COMEXT trade data, UN COMTRADE trade data and Eurostat transport data as main inputs. The main outputs are an extended freight origin-destination database for the year 2005, extended road and rail networks, and new maritime and air-cargo networks. Despite the effort of the Worldnet modellers, there is a clear lack of data both on the demand as on the supply side, which the authors also highlight themselves.

Heinitz and Meincke, 2013 uses a multi step, twofold capacity-restrained itinerary-based demand assignment algorithm to predict global airfreight flows. Main inputs for this model are period timetables on a sample week basis, including all scheduled air cargo services on the supply side and on the demand side it used an estimated weekly demand figure based on IATA quantity structures and published airport and airline statistics. This model is not useful to apply on a yearly level because it tries to bridge the gap between a macroeconomic approach to predictions of cargo volumes on the one hand and network-oriented analyses on the other.

Another model which is able to perform on a global scale is the ITF international freight model, which projects international freight transport activity and related CO_2 emissions on alternative world scenarios (ITF, 2015). It uses a shortest path algorithm to estimate trade tonne-km between each production/consumption centroid for each transport mode. With underlying trade projection which includes a regional aggregation of 26 zones the input data is not very granular. A trade value mode share model is used to split the O/D trade into different transport modes and a weight model converts the value of the traded goods into weight.

The final freight model which is able to perform flow estimations on a global scale is the World Container Model (Tavasszy et al., 2011). The model considers more than 400 major ports, 237 countries, and more than 800 shipping lines. It is built on multinomial logit theory for explaining the mode and route choice behaviour of the shippers across alternative routes for each origin and destination. As the Worldnet model, the origin and destination demand matrix data is obtained based on international trade statistics (COMTRADE) and two European statistics database (EUROSTAT). The model includes import, export and transshipment flows of containers at ports, as well as hinterland flows and is able to reproduce maritime flows rather accurately. The methodology and features included in this model seem very promising to serve as fundamental building blocks for the current research into the estimation of air cargo flows from O/D trade data.

3 MODEL FRAMEWORK

This chapter explains the fundamental aspects of building the model. First, the requirements of the model are stated, then assumptions are defined coming from those. Lastly the model blocks which form the structure are presented.

3.1 REQUIREMENTS

Before creating a structure of model, it is important to firstly define the model requirements. We stated the following five model requirements:

- **Global** - The air trade market is in such a way interconnected that if this problem can be approached from a global level it will be of the most value. Therefore the model should in the end be able to estimate worldwide flows where it captures close to 100% of the air trade market.
- **Feasibility** - In order to make the model practical and usable for (future) market analysis purposes, computational feasibility is one of the model requirements. There model should be easy to run within a few minutes.
- **Data input** - The data input, needed to run the model and to produce results, should be available on a global level.
- **Calibration** - Calibration of the model should be possible by using observed data which is reported on a global scale and is related to the air transport network and the trade which flows through that.
- **Output** - The model should estimate air cargo flows from airport to airport which will be the main output.

All five requirements will be taken into consideration when the model is specified and developed in the next phases of this research.

3.2 ASSUMPTIONS

In order to cope with the requirements stated in the previous section, the complexity of modelling air cargo is narrowed down by some assumptions. The following main assumptions are used in our model:

- The Value of Time (VOT) is the same for all traded goods and is an average monetary value (USD/ton/day).

- The model will not take cargo capacity into consideration. Links and nodes are assumed to have unlimited capacity.
- Air Trade is transported with Wide Body (WB) freight, passenger, and Narrow Body (NB) aircraft
- International Freight Handled reported by airports includes the airport throughput which accounts for the air trade flowed through the airport.

The context around those assumptions will become more clear along the report.

3.3 MODEL STRUCTURE

The presented model is an air freight traffic assignment model, made for the estimation of air cargo flows on a yearly basis. The framework of the model is based on the World Container Model (WCM), which is an existing strategic network choice model for global container flows (Tavasszy et al., 2011). The model considers more than 400 major ports, 237 countries, and more than 800 shipping lines.

Figure 2 presents the general framework of the Air Cargo Flow Model. The blue highlighted trapezium shaped blocks represent the two main inputs of the model, namely demand and supply. On the demand side *O/D Air Trade Data* is used, which means that input demand data only refers to air transport demand. A mode split is therefore not needed in this model. On the supply side, the input consists of a physical network and a service network which together form the *Transport Network* of the model. The service network includes scheduled flight routes reported by airlines. Apart from this, the model also includes the possibility to use Road Feeder Services which makes it multimodal. These main data inputs are described in more detail in Chapter 4.

For each of the O/D country pairs from the Air Trade Data a *choice set* is created with possible routes. After this a *route choice* model determines the trade share of each route. In the end, *flow assignment* assures that air cargo flows are assigned between the airports and aggregated on airport level. *Calibration* of the model will be based on the total absolute difference between observed and estimated throughput on an airport level. After this, the choice set can be recreated for each O/D pair. When the model is clearly converging, the model is calibrated and ready to be used.

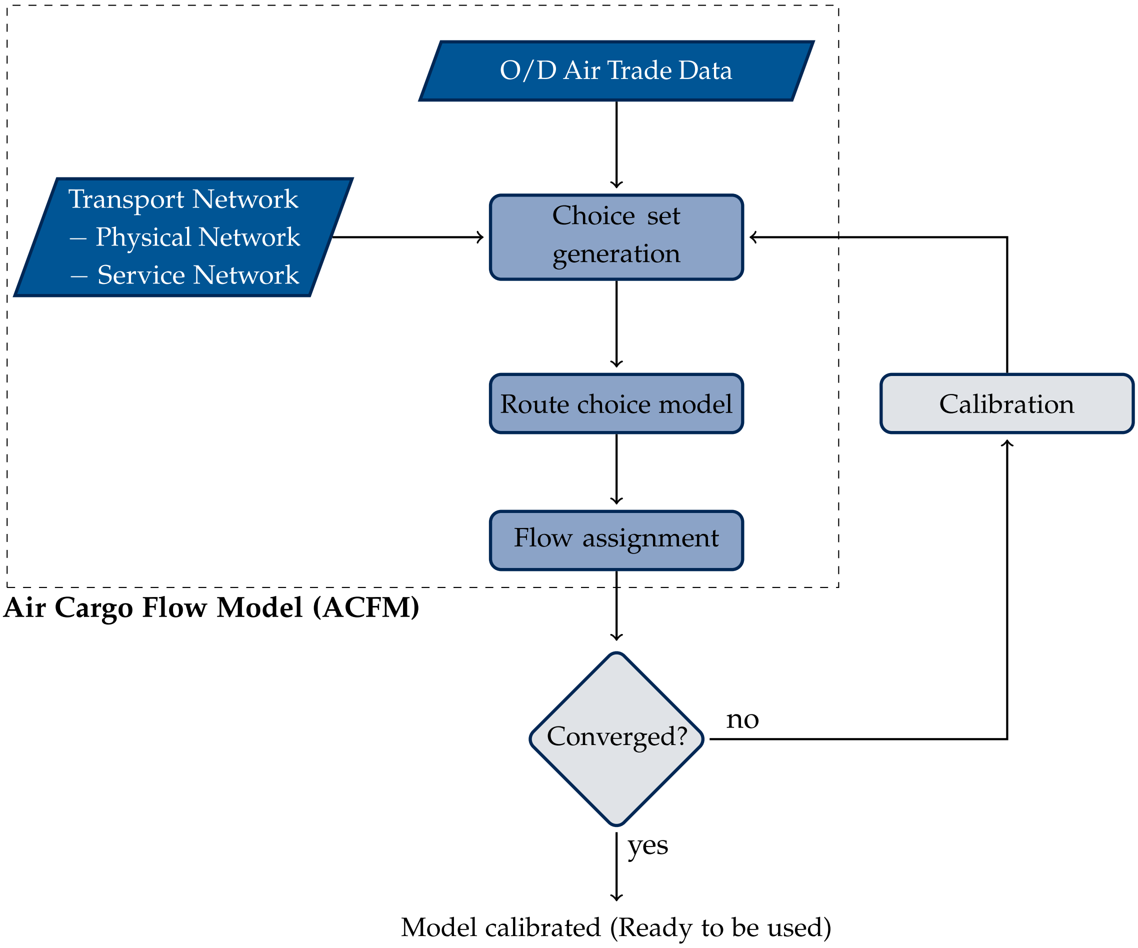

Figure 2: General framework of the model

3.4 MODEL BLOCKS

In this section, the building blocks which form the model framework are explained in more detail. These components of the framework are using the data inputs which will be explained in the next chapter.

3.4.1 *Multimodal*

In order to combine the physical network, including air transport and RFS, with the service network, a super network approach is used to allow a simultaneous choice of mode of transport and route, including transshipment points (Sheffi, 1985). The transshipment links represent the possibility of modal change between the different modes and contain the informations regarding the costs and times required. The multimodal super network is illustrated in Figure 3.

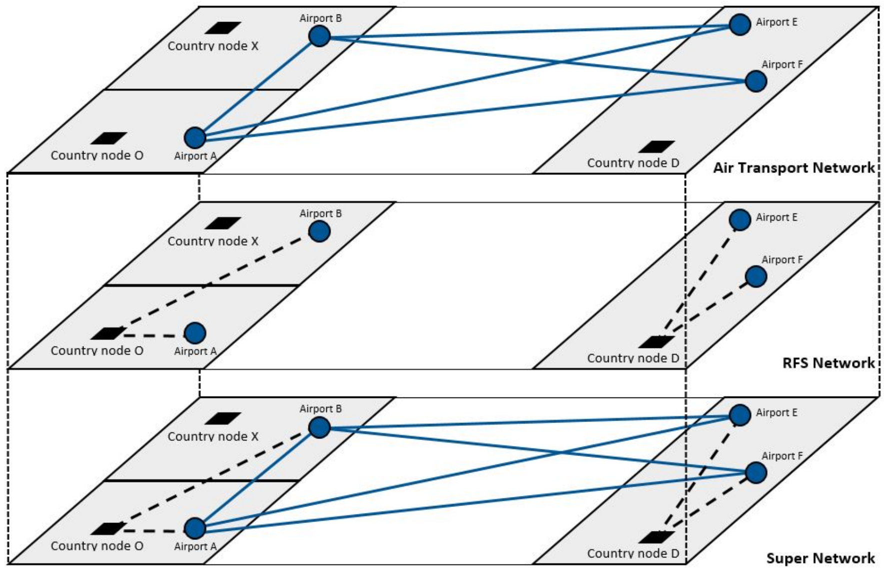

Figure 3: Sketch of the combined air transport and RFS network

3.4.2 *Choice set generation*

The choice set of the routes are generated by using a shortest path algorithm based on generalized costs for each available origin airport towards each available destination airport within an O/D country pair. This means that the choice set of each O/D pair is equal to the number of available airports at origin country multiplied by the number of available airports at destination country.

A small example of a choice set generation is illustrated in Figure 4, where the O/D country pair is origin country O and destination country D. The available airports of country O are airport A, but also airport B in neighbouring Country X may be used and can be reached with trucking possibilities. From, in this specific case, 2 airports the model creates a route towards the 2 available destination airports in country D, airport E and F. A route consists of a sequence of different air transport service-links with the possibility to transship or transfer at an airport. For each airport pair only the route which is known as the shortest path in terms of generalized costs, is added to the choice set. The total choice set of routes possible to use for the trade between Country O and Country D is presented in Table 2. In this example it is assumed that the direct flight option is the cheapest in terms of generalized costs, and therefore chosen as the shortest path.

3.4.3 *Route choice model*

To determine the routes taken by traded goods between two countries, we assume that route choices are made by profit maximizing forwarders who have knowledge of the main routing alternatives over

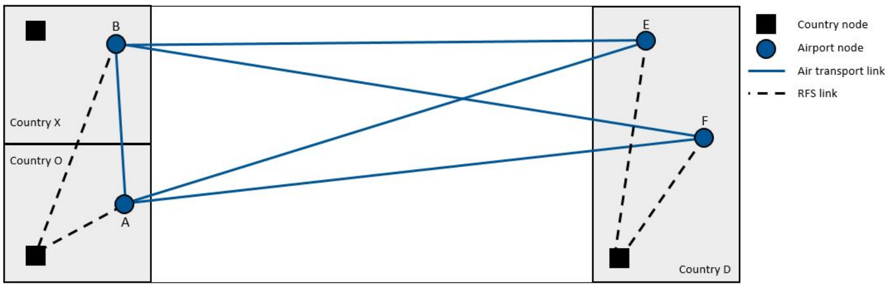

Figure 4: Example of routes between an O/D country pair

O/D PAIR AVAILABLE AIRPORTS	SERVICE LINK TRADE IMPORT	SHORTEST PATH
A - E	A - E	✓
	A - B - E	
A - F	A - F	✓
	A - B - F	
B - E	B - E	✓
	B - A - E	
B - F	B - E	✓
	B - A - F	

Table 2: Total choice set of O/D country pair

air. The basis for the route choice model is a logit route choice model where the probabilities depend on route specific generalized costs (Ben-Akiva and Bierlaire, 1999). The logit model does not take into account the path overlap between alternative routes. An overlapping service link makes it less attractive and therefore it should be taken into account while choosing the best route. This is done by adding the path size overlap variable (Hoogendoorn-Lanser et al., 2005). The route probability is given by:

$$P_r = \frac{e^{-\mu(C_r + lnS_r)}}{\sum_{h\epsilon CS} e^{-\mu(C_h + lnS_r)}} \tag{4}$$

where P_r, the choice probability of route r; C_h, generalized costs of route N; CS, the choice set; μ, logit scale parameter. This last parameter will be used for calibration, which will be discussed in Chapter 4. The path size overlap variable is defined as:

$$S_r = \sum_{a\epsilon\Gamma} \frac{(z_a)}{(z_r)} \frac{1}{N_{ah}} \tag{5}$$

where S_r, degree of path overlap; a, link in route r; Γ_r, set of links in route r; z_a, length of link a; z_r length of route r, N_{ah}, number of times link a is found in alternative routes generated within the choice set. The generalized cost function per route r is given by the following equation:

$$C_r = \sum_{p\epsilon r} A_p + \sum_{l\epsilon r} c_l + \alpha \cdot (\sum_{p\epsilon r} T_p + \sum_{l\epsilon r} t_l) \tag{6}$$

where C_r, costs of route r; p, airports used by the route; l, links used by the route, A_p, airport impedance parameter for airport p; c_l, total cost of transportation over link l; α, value of transport time (USD/day/ton); T_p, time spent during transshipment at airport p; t_l, time spent during transportation over link l.

Two values are unknowns and need to be estimated, namely, the airport impedance parameter A_p and the scaling parameter μ. The scaling parameter μ determines the impact that the cost differentials have on the percentage of trade assigned to each alternative. The parameter A_p includes all relevant, measurable and hidden service characteristics of airports, such as fuel costs, airport charges, handling costs, congestion costs, etc. As part of the route costs it is stated in USD/ton. The value of time (α) denotes the average preference for the shippers for either a faster (and thus more expensive) or a slower (and thus cheaper) transport option. The α for freight was inferred from an earlier study on value of freight (Jong et al., 2014), which is equal to $4,075$ USD/ton/day.

3.4.4 *Flow Assignment*

After the different possible routes are generated, the logit route choice probabilities are calculated for the alternatives within the choice set. To distribute the flows over the links in the network, the probabilities are multiplied with the O/D trade reported. Flows will be aggregated on an airport level, which determines the calculated throughput for the airports in the model.

3.5 SUMMARY

This chapter has described the requirements of the model and the assumptions which are made. It also presented the model framework including several model blocks which together are responsible for producing an estimated value of air cargo flows. Each step is described with the support of examples, figures and formulas. In the next chapter the input data of the model is described which feed the model blocks. Also the data and method to calibrate the model is discussed.

4 MODEL DATA & CALIBRATION

This chapter describes all the data which is used as input for the model and/or for supporting analysis. It also contains the explanation of the calibration methodology and the data used for that.

4.1 AIR TRADE DATA

The main input data for the model on the demand side were obtained from one main source, namely the *Seabury Global Trade Database* (Seabury Consulting, 2014b). The trade database captures historical global air trade between more than 200 countries. The primary data sources for the global trade database are custom offices. Each month, raw data is directly sourced from 45 individual countries' customs/statistics offices representing around 97% of worldwide international trade by weight. The remaining data is sourced indirectly by UN Comtrade data (22 countries) which is available on a yearly basis and enables to include the coverage of countries which are not directly reporting (see Appendix A). Hereby the U.S. and China are reporting their trade on a more granular level so they are further split up into different regions, 6 and 4 respectively (Figure 5).

For each trade lane, which is between true origin and destination country, data is available for 2,000 commodities such as laptops, cell phones or footwear. Apart from historical analysis, a five year forecasting module enables to forecast air trade per trade land for 70 key industries. The true origin and destination country together form an O/D pair which will be used in the model. Data is reported both in *Air Value* (in USD) and in *Air Weight* (in metric tonnes) on a monthly basis from 2006 onwards. Yearly air weight from the year 2014 was used as an input for this model. The year of 2014 is chosen because this is the latest year internal Seabury data and external data could be combined to be able to create a model which could be created on a global scale. As stated in the scope, within current research we do not consider commodities but take air trade as a whole.

4.2 NETWORK DATA

On the supply side of the model an transport network is generated. This network is the combination of input data about country nodes, airports, the air transport network and the trucking network.

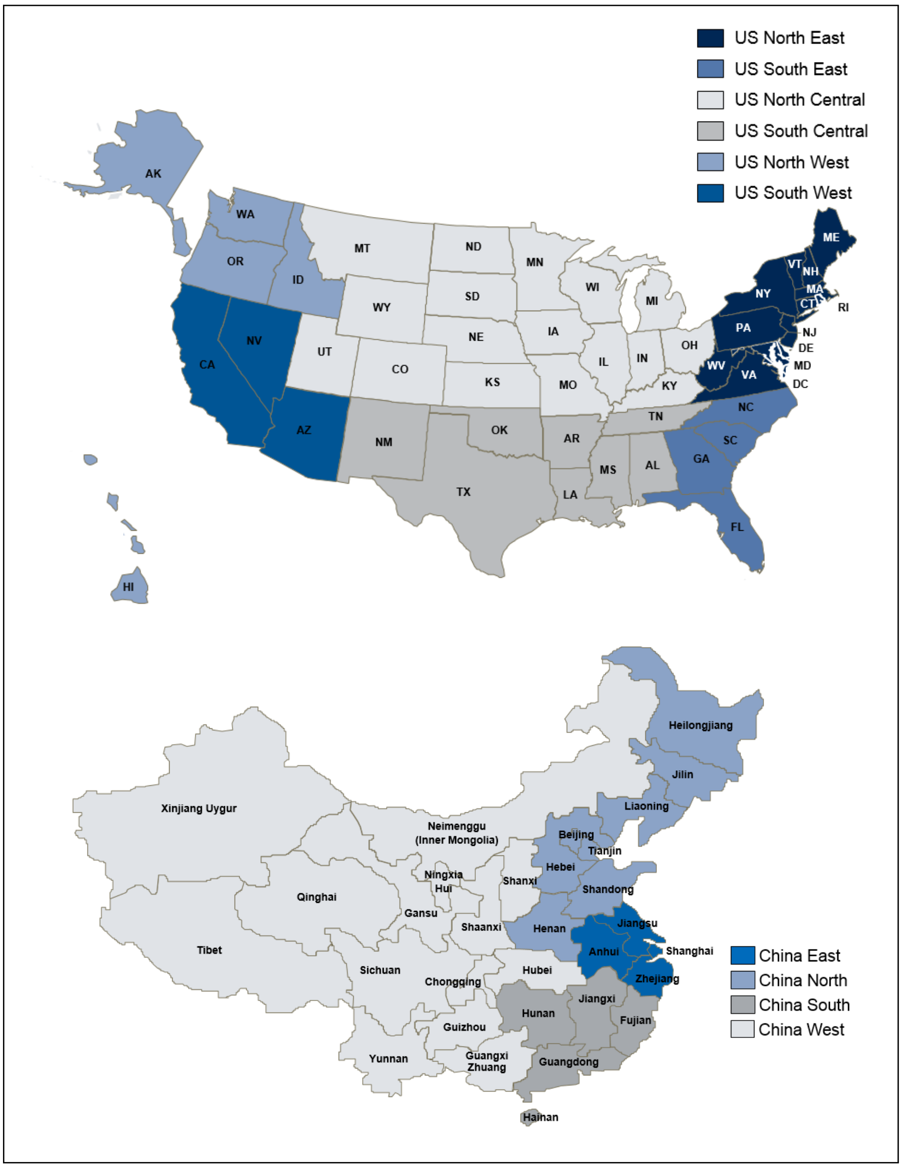

Figure 5: U.S. and China split up into regions from reporting trade data

4.2.1 *Country nodes*

In order to connect the O/D Air Trade Data from country to country with the Transport Network, country data has to be loaded into the model. Hereby a number of simplifications has been applied. First, the multimodal super network is connected to every country of origin or destination by a single centroid. Hereby we assume that this location captures the weighted average location of all air freight moving to, or from the country, the so called cargo center of gravity (CG). In most cases this centroid is the capital city of the respective country. This assumption is true in most cases with some exceptions. Within Germany, for example, the biggest cargo airports in terms of throughput are Frankfurt (FRA), Leipzig (LEJ) and Cologne (CGN). From this one can be assumed that the CG would be near FRA instead near the capital of Berlin (Figure 6). A total overview of the exception cases are presented in Table 3.

For the regions of the U.S. and China, the location of the centroid is based on the same assumption as used for the countries. This resulted that those U.S. and China region centroids laying at the largest airport of each region in terms of yearly airport throughput. Only for US North West, Seattle (SEA) is chosen instead of the larger airport of Anchorage (ANC), which is mainly used for transpacific stop-overs (Figure 7).

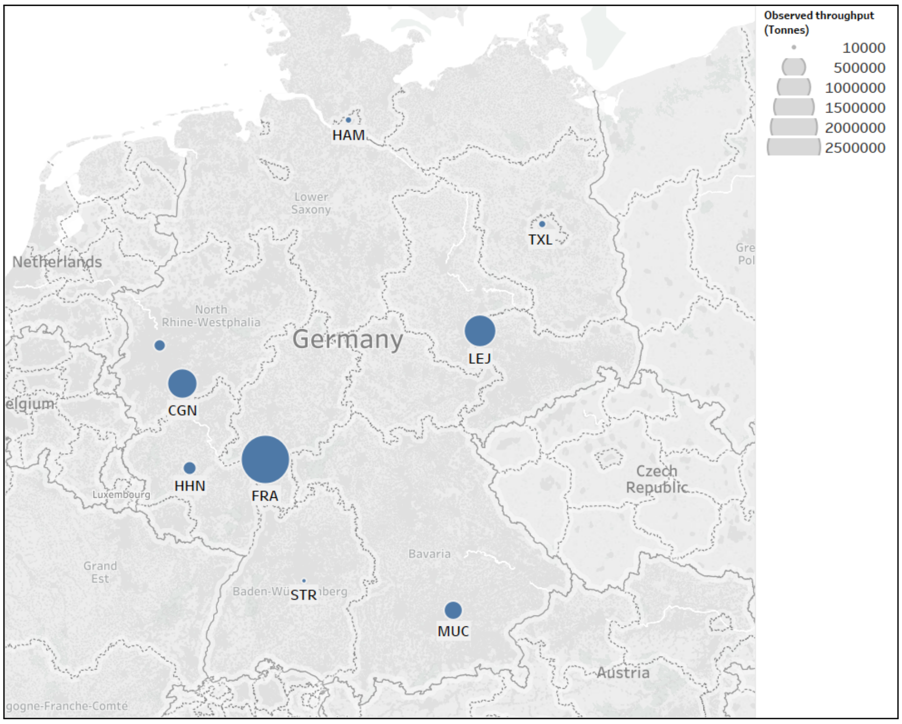

Figure 6: Main cargo airports of Germany

4.2.2 *Airports and Air Transport Network*

The supply side of the model consists of the transport network. This includes the physical network and the network of air transport and trucking possibilities. The airports incorporated into the model are based on airports associated with Seabury airport data. For the air transport network, the model is using scheduled flight services provided by Innovata (**Innovata**). An extract sample is presented in Table 4. The data includes full route codes with multiple stops, performed by a certain carrier, flown by a specific body type (NB = Narrow Body, WB = Wide Body) and configuration (Pax = passenger, Frt = freighter).

For the model, WB passenger and freighter services, and NB freighter services are used, because we can assume that those aircraft are the

COUNTRY	CAPITAL CITY	AIRPORTS WHICH HAVE BIGGEST INFLUENCE ON CG	CG
Germany	Berlin	FRA, LEJ, CGN	±20km North of Frankfurt
Italy	Rome	MXP, FCO, BGY	±70km South-East of Milaan
Turkey	Ankara	IST	Istanbul
Brazil	Brasilia	GRU, VCP	São Paolo
Canada	Ottawa	YYZ	horizontal center of Canada
Vietnam	Hanoi	HAN, SGN	vertical center of Vietnam
Kazakhstan	Astana	ALA	Almaty
Iraq	Bagdad	EBL	Erbil
Morocco	Rabat	CMN	Cassablanca
Tanzania	Dodoma	DAR	Dar es Salaam

Table 3: Overview of the exception cases where the country capital is not the cargo CG

main providers of transporting international trade worldwide. Integrator and charter services are not included in this scheduled flight service dataset.

YEAR	CARRIER	FLIGHTNO.	FULL ROUTE CODE	BODY TYPE	CONFIG
2014	KL	701	AMS - EZE - SCL	WB	Pax
2014	DL	281	ATL - SEA - HKG	WB	Pax
2014	EK	4993	DXB - FRA - ATL	WB	Frt
...	...	...	...	...	...

Table 4: Extract sample of flight route data (Innovata, 2014)

4.2.3 *Road Feeder Services*

As an addition to the transport network, also continental transport in the form of Road Feeder Services (RFS) is considered in the model. Because there can be assumed that the trade of goods will not only use the airports within the origin and destination country, alternative route options will be captured by adding the possibility of using airports of neighbouring countries. Those airports can be used with the creation of a simplified continental trucking network of links was

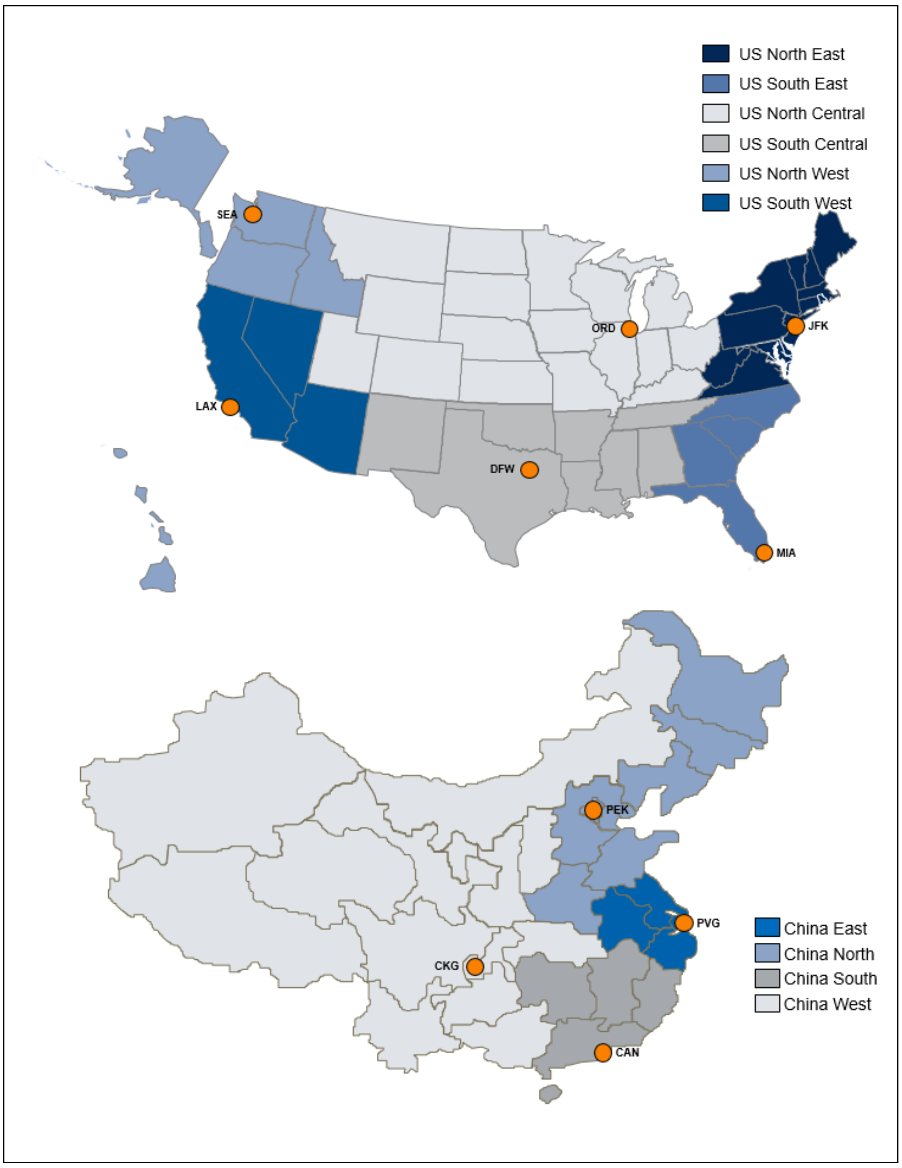

Figure 7: Centroids of the U.S. and China regions

created. Airports are included if they are within a radius of 1,000 km from one of the airports from the origin or destination country (and within the airport set of the model) of the specific O/D pair for which the choice set is generated. An example can be seen in Figure 8, were the available airports are presented for Germany. This means that the choice set of routes will be expanded with the shortest path from these available airports towards the available airports for the destination country.

4.3 MODEL PARAMETERS

The main network attributes are times and costs for transhipment and transportation. Over the whole network average level of trans-

Figure 8: Available airports for Germany which can be reached with trucking

portation costs and speeds are used. The speeds and tariffs used in the model are presented in Table 5. As discussed in Chapter 3, the multimodal transport network only considers air transport and trucking. Rail transportation is left out of the scope of this model. The Air Transport speed is an average obtained from the individual flight duration's reported by Innovata for the scheduled flight services. The Air Transport Cost is a computed average from published IATA Air Freight Yield figures (IATA, 2014). Both the trucking speed and costs are taken from an earlier study on the movement of containers, which besides a maritime network also included trucking options (Tavasszy et al., 2011). Because trucking facilities and quality of the network can be different per continent, trucking costs in Oceania, Asia and Africa are multiplied by 2, 3 and 4, respectively. Costs in Europe and America will not be multiplied. Furthermore, transshipment time of 18 hours is used in the model. This includes the loading and unloading time of the cargo, the customs clearance and warehousing (Ohasi, 2005).

	AIR TRANSPORT	ROAD TRANSPORT/ TRUCKING
Speed (km/hr)	750	42
Tariff (USD/km/tonnes)	0.175	0.057

Table 5: Description of the characteristics for the services used.

4.4 CALIBRATION

To increase the fitting of the model, the model will be calibrated against observed airport throughput data. The main objective of this exercise is to minimize the absolute difference between the modeled airport throughput and the observed, or reported, airport throughput.

4.4.1 *Calibration data*

Airports Council International (ACI) is the global trade representative of the world's airport authorities and yearly reports airport traffic statistics in the *ACI World Airport Traffic Reports* (ACI, 2014). Amongst other data, this reports include the amount of international handled freight per airport. Here we assume that when trade is transported via air and therefore taking off and landing at airports are included in this number. This international handled freight, or observed airport throughput, is used for the calibration of the model. Unfortunately, the amount of transshipped freight is not available in the ACI reports. Therefore transshipment data is taken from public data sources.

4.4.2 *Model fitting*

As described in Chapter 3, two values are introduced for model fitting purposes, namely (1) the scale parameter μ and (2) the airport impedance parameter A_p. The main objective of this calibration is to minimize the absolute difference between the observed throughput reported by ACI and the modeled throughput estimated by the model.

μ calibration

First, the scale parameter μ is adapted and the absolute throughput difference is calculated by running the model. In this stage, the A_p-value of all airports is equal to zero. The μ value where the difference is the lowest, gives an indication of the range were the model fitting will be the most promising.

A_p calibration

Secondly, the airport impedance parameter A_p is modified for each individual airport. To minimize both the difference in throughput as in transshipment, a multi-objective optimization method NSGA-II is used, which is a state of the art algorithm for solving multi-objective optimization problems and is rooted in evolutionary theory (Reed et al., 2007). Multiple solution points are used to explore the solution space and find the best solutions for the two objectives. NSGA-II uses crossover and mutation operators that mimic evolutionary processes when searching the solution space. Together with a solution selection routine, the algorithm explores and exploits the solution space efficiently (Crepinsek et al., 2013). The computation process is carried out using a simulation-based optimization framework called SIMEON (Halim and Seck, 2011).

To reduce the search space, boundaries will be set for the to be assigned A_p-values. After the values are assigned to the airports and the model is generating the choice set, calculates the route probability and assigns the flows. Because the routes are generated with a shortest path algorithm based on generalized costs, the route choice sets needed to be created at each iteration step in the calibration process. After a certain amount of evaluations the search procedure is stopped and gives a set of non-dominated solutions. The solution which results in the lowest absolute throughput difference is again used as input for a new run of evaluations.

4.5 VALIDATION

In order to check if the model produces reasonable results in terms of air cargo flows, the model is validated upon reported air traffic data and air cargo capacity data.

4.5.1 *Validation data*

BTS data

The Bureau of Transportation Statistics (BTS) is part of the United States Department of Transportation and reports transportation data, which also includes air traffic related data. From this source we can therefore gather domestic and international market data reported by U.S. and foreign air carriers which contains reported carrier, origin and destination and freight per flight (Transportation Statistics, 2014).

Capacity data

Besides the Trade Database, Seabury also offers other database products to its customers. The *Capacity Database* includes scheduled airport to airport air cargo capacity data per airline, on top of the existing Innovata schedules. This capacity data is available in tonnes, ATKs (Available Ton Kilometers) and number of flights per week per route (Seabury Consulting, 2014a). A sample extract of this capacity data is given in Table 6.

4.5.2 *Model quality*

The process of validation to confirm that the model is capable of reproducing the airport throughput and air cargo flows is performed

carrier	aircraft type	origin airport	destination airport	capacity (tonnes)
CX	Boeing B747-8F	ANC	HKG	13,121
SQ	Boeing B747-400F	SIN	HKG	5,481
CK	Boeing B777-200LRF	LAX	PVG	2,913
QR	Boeing B777-300ER	DOH	BKK	2,306
EY	Boeing B777-200LRF	AUH	PVG	1,964
CX	Airbus A330-300	HKG	TPE	1,295
AA	Boeing B777-200ER	MIA	EZE	1,283
JL	Boeing B767-300ER	HND	KMQ	1,216
AF	Boeing B747-400F	CDG	MEX	2,043
DL	Airbus A330-300	SEA	NRT	427
JL	Boeing B787-8	NRT	BKK	1,126
...	...	...	...	...

Table 6: Extract sample of capacity data (January 2014)

with the data explained above. In order to proof that the model is capable to capture reality a smaller case study is created. After we can confirm that the output of the model correlates with the observed validation data, we proceed by applying it on a global scale.

4.6 summary

This chapter has described all the data used for model. The O/D air trade data is presented for the demand side and the network data is presented for the supply side. The parameters which mainly attribute to the time and cost calculation in the model are given. Furthermore, the calibration methodology is described in detail with the supporting data. Lastly, the validation of the model is shortly discussed.

5 US - BRAZIL

Before applying the model on a global scale, we want to perform some validation and verification checks to ensure the ability and robustness of the model. Therefore, first a case study is set-up where trade is distributed from the U.S. regions (US) towards Brazil (BR). The purpose of this is to illustrate the capabilities of the model to reproduce two specific outputs, namely (1) air cargo flows from airport to airport and (2) yearly airport throughput.

5.1 SET-UP

In order to perform validation and verification checks on the model the US-BR analyse needs a set-up. Hereby we have to make some assumptions. In this US-BR case we assume that:

1. The majority of O/D trade between US regions and Brazil is transported via direct flights
2. The interference of O/D trade trade towards Brazil from other countries than the US on the US-BR trade lanes, is very low. Interference is defined as the possibility that the routing of air trade imports towards Brazil from other countries will flow via the US.

The combination of those two assumptions would mean that air cargo flows from US airports towards BR airports only consists of trade between the US regions and Brazil. In that case, the US-BR could be seen as an isolated trade market which is suitable for further analysis and to prove the capabilities of the model before going global. To support the assumptions, a detour analysis and a capacity share analysis is done.

5.1.1 *Detour analysis*

The stated assumptions are firstly supported by a data analysis whereby the detour factor is calculated between the main trade regions of Brazil and Brazil. The detour factor of a route is equal to the indirect great circle distance divided by the direct great circle distance between two airports. The use of detour factors for analyzing airline networks in terms of connectivity or for network design problems has been done before (Boonekamp and Burghouwt, 2016; Kuby and Gray, 1993). Results from the detour analysis is presented in Table 7. In this table the origin regions of the Brazilian trade import are shown including the percentage of total Brazilian imports which they are accountable for. Detour factors are computed from the main airport of the countries within the regions, with a stopover at an US airport. Guarulhos International Airport, the main airport of Brazil located at São Paolo, is used as destination airport. The detour factor range is presented in the last column. For example, the direct route from Amsterdam (AMS) to São Paolo has a distance of 9,752 km. The indirect option via the US is the most optimum via New York (JFK) which has a flight distance of 13,500 km. This together results in a detour factor of 1.38.

From the table one can notice that the majority of Brazilian air trade import from European countries, which account for 33% of Brazil total imports, can be assumed to be flown directly. Detour factors for flights between Europe and Brazil with a stop in the US are around 1.50. This means that flight related costs, which is included in the model in terms of flight distance, will be multiplied 1.50 times, which will decrease the attractiveness of these options. When a maximum direct flight distance for a commercial flight of 14,500 km is applied, the Europe region is well within this range (Calder, 2017). On top of that, research shows that forwarders prefer direct routings over indirect routings (Chu, 2014). So it is highly unlikely that there is interference with the O/D air trade between US and Brazil coming from Europe. The same holds for trade from Middle East & South Asia (MESA) region with detour factors between 1.40 and 1.60.

Possible interference comes particularly from Asia Pacific (APAC) countries like South Korea, Japan and China, which account for 31.5% of Brazil total imports. Because of the flight distance, trade from those countries towards Brazil are forced to have a 1-stop flight. A stop at US grounds ranges between detour factors of 1.00 (NRT - JFK - GRU) and 1.44 (SIN - JFK - GRU).

This detour analysis strengthen the previously stated assumption to some extend, but did not ruled out completely the interference on US-BR trade lanes. Especially the APAC region has to be investigated in more detail.

5.1.2 *Inbound capacity analysis*

A second part of the analysis to support the two assumptions is based on the inbound cargo capacity share towards Brazil. For this, the Seabury Cargo Capacity Database is used (see Section 4.5.1). Results from this analysis are presented in Table 8, where again the origin regions for the trade towards Brazil are shown in the first column. North America accounts for 38% of all WB passenger and freighter, and NB freighter capacity directly towards Brazil. This high figure strengthens the first assumption where was stated that the majority of O/D trade between US and BR is transported via direct flights. Also

ORIGIN REGION	% OF BR TRADE IMPORT	DETOUR FACTOR RANGE (VIA US)
Europe	32.9%	1.38 - 1.55
Asia Pacific	31.5%	1.00 - 1.44
North America	27.0%	N/A
Latin America	5.8%	N/A
MESA	2.5%	1.40 - 1.60
Africa	0.3%	N/A

Table 7: Detour factor ranges for O/D trade market towards Brazil

the fact that the direct capacity share from European countries is high, confirms the high detour factors from previous section.

In order to strengthen the second assumption about the low interference of O/D trade trade towards Brazil from other countries than the US, we shift our focus towards the indirect capacity share towards Brazil, which is presented in the third column. Europe and Asia Pacific account for 9% and 3% of total inbound cargo capacity of Brazil. And because only 1.2% and 2% of this is from scheduled airline services which is routed via the US, there can be concluded that the capacity towards Brazil from Asia Pacific countries with a stop in the US is minor.

From both data analysis, we can conclude that potential interference of O/D trade is does exists but is minimized within the trade market between the US regions and Brazil and therefore can be used as a proper case study to check if the model is able to reproduce yearly airport throughput and air cargo flows in an accurate way.

5.1.3 *US-BR Case study*

The set-up of this case study includes six airports in the US, six airports in Brazil and more than 2,000 scheduled flight services. Domestic routes within the US and Brazil and international flights departing in the US and arriving in Brazil are included. For each US region the largest airport in terms of airport throughput has been chosen. Only for US North West SEA airport is chosen instead of the larger ANC airport, because ANC airport is mainly used for transpacific stop-overs. For Brazil the six biggest airports in terms of international Wide Body capacity is chosen. The centroid for Brazil is relocated from the capital city Brasilia to the city of São Paolo, because this location captures more the cargo CG of Brazil (see subsection 4.2.1) Trucking between the US regions and within Brazil is allowed, as well as transshipments at the available airports in the set. A geographical

ORIGIN REGION	Capacity share DIRECT	INDIRECT	VIA US (CARRIER)
Europe	33%	9%	1.2% via MIA (Centurion)
Asia Pacific	0%	3%	2% via ANC, LAX or MIA (Korean)
North America	38%	0%	-
Latin America	9%	0%	-
MESA	4%	1%	0%
Africa	2%	0%	-

Table 8: Direct and indirect capacity share for O/D trade market towards Brazil

map of this case study is presented in Figure 9. The trade data between the US regions and Brazil is presented in Table 9.

ORIGIN COUNTRY	DESTINATION COUNTRY	TRADE (TONNES)
US South East	Brazil	87,489
US North East	Brazil	12,481
US South Central	Brazil	8,734
US South West	Brazil	6,021
US North Central	Brazil	3,825
US North West	Brazil	1,161

Table 9: Trade data from US regions towards Brazil

5.2 CALIBRATION

For model fitting purposes, the scale parameter μ and the airport impedance parameter A_p are adapted. In order to calibrate the model, observed throughput is needed for the 12 airports. But as mentioned before, ACI only reports observed airport throughput as yearly totals without a split from which region the throughput is coming from. This data does not hold for this case study where we only consider trade leaving the six US airports and arriving at the six Brazilian airports. Therefore we have to determine the observed airport throughput in another way.

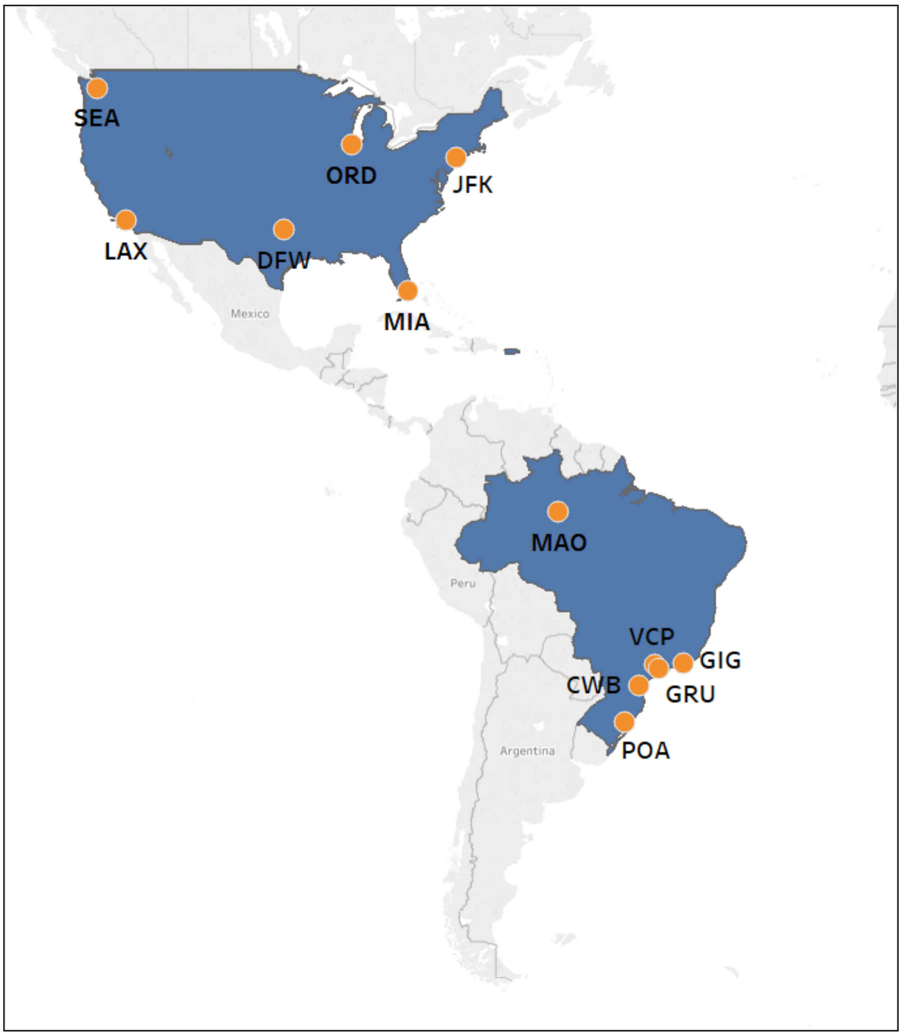

Figure 9: Geographical map of the US-BR case study including 12 airports

5.2.1 *Airport throughput*

For the US airports we use the reported trade from the US region where the airport is the centroid of, as throughput figure. For example, JFK airport is the centroid of US North East, so US North East - Brazil trade will be used as throughput for this particular airport. To support this assumption all flights from the US region are allocated to the main airport. This means that a BOS - GRU flight will be changed to a JFK - GRU flight.

For the Brazilian airports we again use capacity data from the Seabury Air Cargo Capacity Database. To determine the observed throughput of the Brazilian airport in this case study we apply a *fair share calculation*. This means that the inbound WB capacity share is used to calculate the share the Brazilian airports will have in terms of throughput. The formula is stated as follows:

$$\frac{Cap_{US-BR_i}}{Cap_{BRtotal}} = \frac{TP_{US-BR_i}}{TP_{BRtotal}} \tag{7}$$

where Cap_{US-BR_i}, inbound WB capacity from the US to Brazilian airport i; $Cap_{BRtotal}$, total inbound WB capacity Brazilian airports; TP_{US-BR_i}, Brazilian airport i throughput from US; $TP_{BRtotal}$, total throughput Brazilian airports. The only unknown in this formula is the throughput of the Brazilian airports when only considering US as an export country. Because observed throughput includes both import as export data, the percentage of US-BR import is applied as a last step. the The result of this calculation is presented in Figure 10.

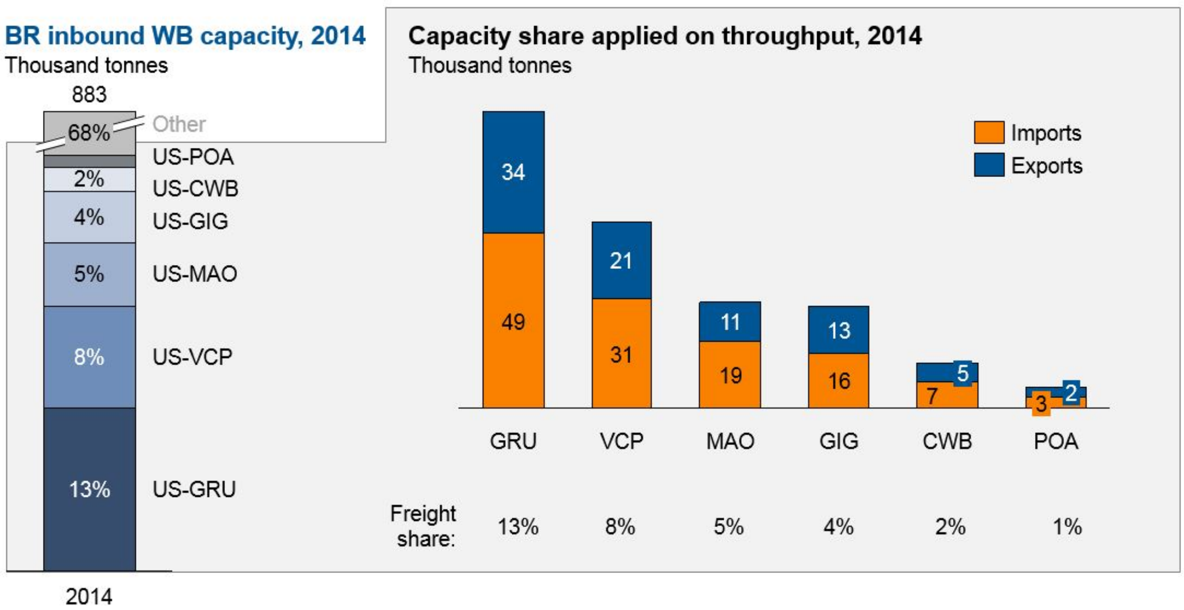

Figure 10: Fair share calculation applied on Brazilian airports

Table 10 presents an overview of the twelve airports and the estimated throughput. We can see that the totals are almost equal and therefore we can conclude that seems like a fair estimation for this case study. This because in such an isolated set-up, the US exports should be rather equal to the BR imports.

5.2.2 *μ calibration*

After determining the 'observed' airport throughput, the calibration of the model can start. As described in subsection 4.4, the purpose of the calibration is to minimize the difference between the observed and calculated flows on an airport level. Firstly, the scale parameter μ being adapted and the model performs the steps from the framework.

AIRPORT	THROUGHPUT	AIRPORT	THROUGHPUT
MIA	87,489	GRU	48,989
JFK	12,481	VCP	30,729
DFW	8,734	MAO	18,755
LAX	6,021	GIG	15,548
ORD	3,825	CWB	7,476
SEA	1,161	POA	3,419
Total	119,711	Total	124,915

Table 10: 'Observed' airport throughput for the US-BR case study

After the flows are assigned the absolute throughput difference is calculated. The result is presented in Figure 11. We have found an optimal value of 0.0006.

The World Container Model found an optimal value of 0.0045 which is 7.5 times higher. The reason for this is because the cost of the routes in the choice set of our case study does vary to a higher extend than in the maritime case. The direct flight routes towards GRU and VCP are very attractive in terms of generalized costs compared to the other options. To assure that there is some variance in the assignment of the flows, the μ-value is lower to increase the impact of the generalized costs on the distribution.

On top of that, the average size of a choice set in the WCM is lower than in the US-BR case study. Where the amount of available ports per country is between 1 and 25 in the WCM, with an average of 4 ports per country, the U.S. regions and Brazil both have each 6 available airports. More available (air)ports means a larger choice set and means a lower μ-value to increase the variance within the set.

Another logit scale parameter value found in literature is in another research about freight traffic assignment in a multimodal network (Maia and Couto, 2013). Here they used a value of 0.1 and 0.5 for μ, but the network they used was a relative simple one, with only five centroids and a couple of road and rail links. For this reason, those values are difficult to compare to the more complex case study we created.

5.2.3 A_p calibration

After the calibration of the scale parameter, we proceed by adapting the airport impedance parameter A_p. Lower and upper boundaries for the parameter are set on 0 and $2,000$. Within these constraints, the model has enough freedom to find impedance parameters for the airports to fit the model. A total of 10 runs have been performed,

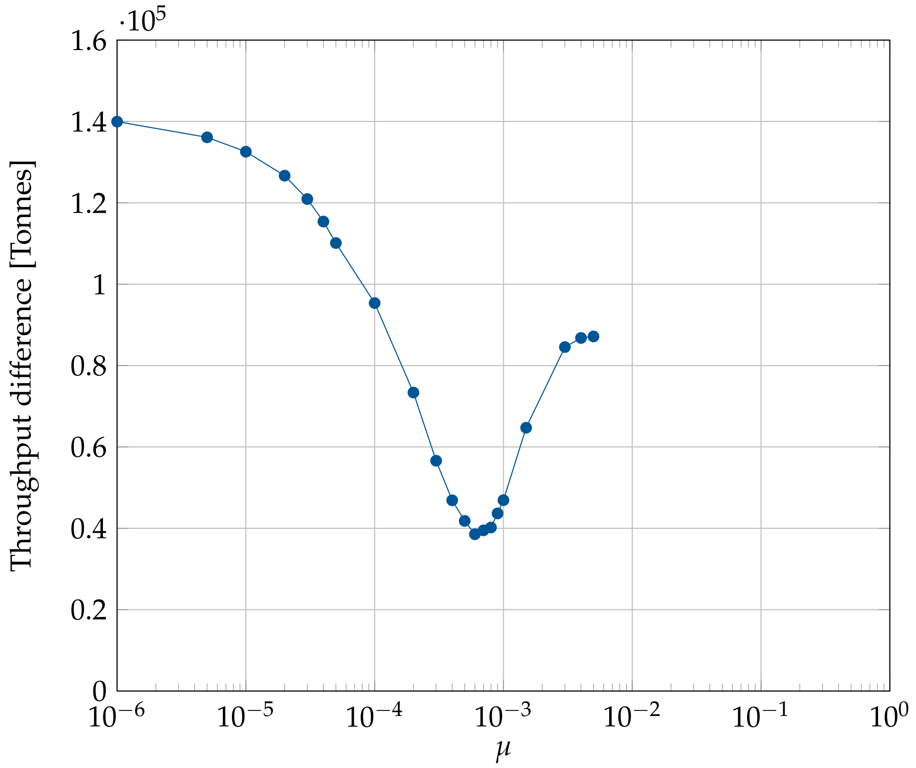

Figure 11: Comparison of the absolute difference between observed and calculated throughput for different scale parameter μ (logarithmic scale)

whereby each run consisted of 10,000 iterations. As can be seen in Figure 12, the algorithm converged very quickly in the first runs and is flatting out after the third run. A minimum absolute throughput difference of $34,546$ tonnes is obtained, which an improvement of more than 10% compared to the difference before the A_p calibration.

5.3 RESULTS

Figure 13 shows the model output. The blue lines are proportional to the volume of air traffic. The size of the pie charts visualize the amount of throughput (dark grey) and transshipment (light grey) at each airport. The map already shows the amount of transshipment is minimum in this case. We utilize the java geo-visualization library Unfolding (Nagel et al., 2013) to visualize the results of the model.

Table 11 presents the calibrated A_p-values for each airport in the model. VCP, the airport which mainly serves Campinas, a municipality in São Paulo, has the highest assigned airport impedance value, where MAO, located approximately 2,700 km to the North West of São Paolo has the lowest value. Hereby one has to keep in mind that a high A_p-value for an airport means a high impedance and therefore a low

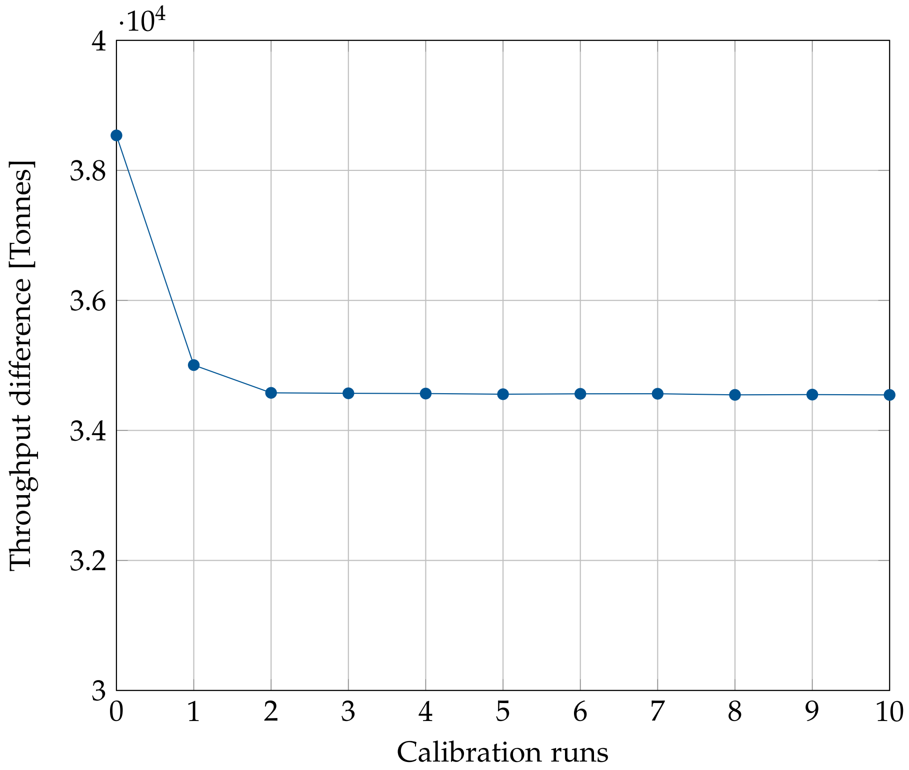

Figure 12: Convergence in throughput difference after different calibration runs

level of attractiveness, where a low A_p-value means a low impedance and a high value of attractiveness.

AIRPORT CODE	CITY	COUNTRY	A_p-VALUE
VCP	Campinas	Brazil	1,954
CWB	Curitiba	Brazil	1,751
GIG	Rio de Janeiro	Brazil	1,736
GRU	São Paulo	Brazil	1,450
LAX	Los Angeles	US South West	1,370
JFK	New York	US North East	1,165
ORD	Chicago	US North Central	929
MIA	Miami	US South East	723
SEA	Seattle	US North West	356
DFW	Dallas/Ft. Worth	US South Central	146
POA	Porto Alegre	Brazil	83
MAO	Manaus	Brazil	16

Table 11: Main airports and countries captured by second filter

Figure 13: Air cargo flows between the US and Brazil

5.4 ANALYSIS

In this section the results from the US-BR case study are analyzed. First, in the validation part, the output is compared with observed data in order to check to what extend the model is able to capture reality. Both modeled airport throughput as airport to airport traffic flows are compared with observed data. Secondly, in the verification part, the model is tested on its robustness. A sensitivity analysis is performed where some parameters are changed to see how the model reacts.

5.4.1 *Validation*

The results show that the model is very accurate in reproducing the observed airport throughput in this US-BR case study. Figure 14 show a coefficient of variation of 95% for the twelve airports in the model. To analyze the correlation in an optimum way the Brazilian airports are marked orange. One can clearly notice that especially the throughput of the US airports is reproduced very accurately since they all are located on the linear black line. Because the observed throughput for the US airports is set equal to the trade of each region, it is maybe more straightforward that the model is able to produce an accurate throughput estimation. This despite the fact that we allowed the model to use RFS towards the airports in the other regions. The Brazilian airports show a correlation which is lower. The airport of GRU, located at São Paolo, is slightly overestimated by the model, even as GIG, the airport from Rio de Janeiro. VCP and POA are estimated very accurately. The airport with the highest difference between observed and estimated flows is Eduardo Gomes International Airport (MAO). Even though, as a result of the calibration, the MAO airport has a low A_p-value, the model is not able to reproduce the throughput for this airport. The airport throughput is underestimated, which probably has to do with the fact that this airport is the furthest away from the centroid of Brazil which is located at São Paolo. This makes this airport unattractive because of high trucking costs, compared to the other five airports which are located more near centroid of Brazil.

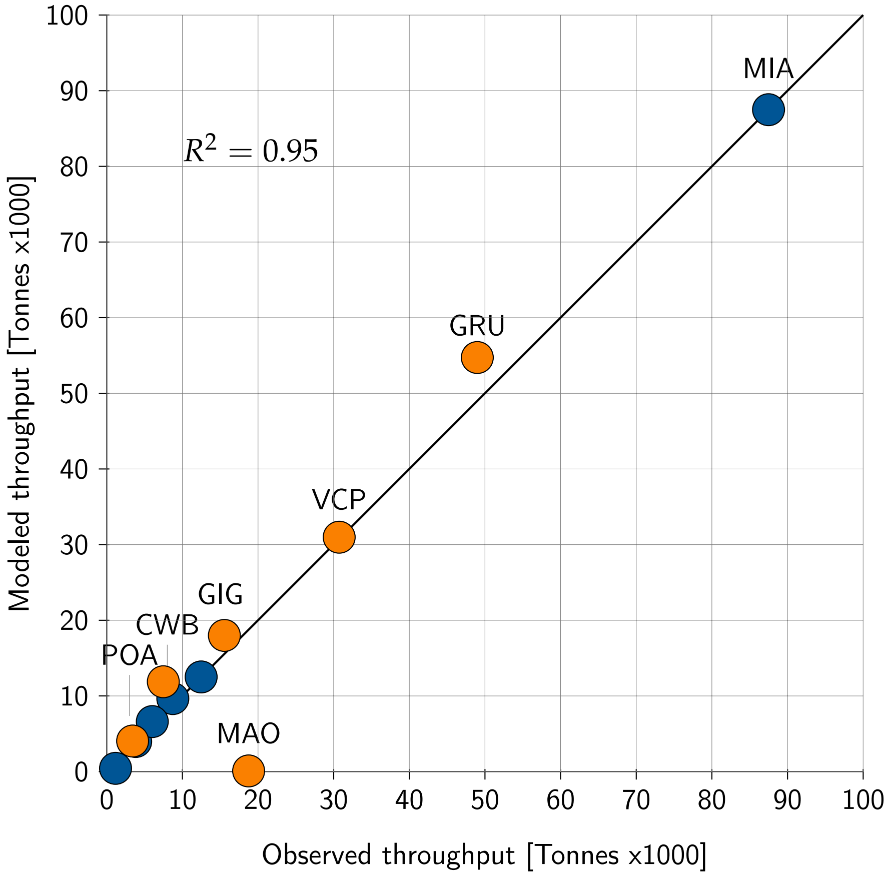

Figure 14: US-BR case study: Calibration result for modeled vs observed airport throughputs (Brazilian airports in orange)

Besides the ability of reproducing airport throughput, we also want to illustrate the capability of reproducing airport to airport flows. To validate the estimated flows by the model with, BTS is used for observed data, which can be seen in Figure 15. The correlation shows a coefficient of variation of 72% for the 36 airport pairs (six times six) in the model. One can see that especially for the air cargo traffic flows up to 10,000 Tonnes there is some variation between the modeled flows and the data from BTS. The more larger flows reported by BTS are mostly originated in MIA, which is rather logical since we saw that MIA is the major throughput airport to transport trade from the US to Brazil. The flow between MIA and GRU is overestimated by the model. This has probably something to do with the fact that the model has difficulties with capturing the flow towards MAO and therefore it will put these tonnes on the most attractive route. Related to that, the route MIA-MAO is underestimated the most, which is in line with the airport throughput we saw. The largest flow reported by the BTS data between MIA and VCP is also underestimated. A more detailed overview of the BTS reported traffic flows and the modeled ones are presented in Table 12.

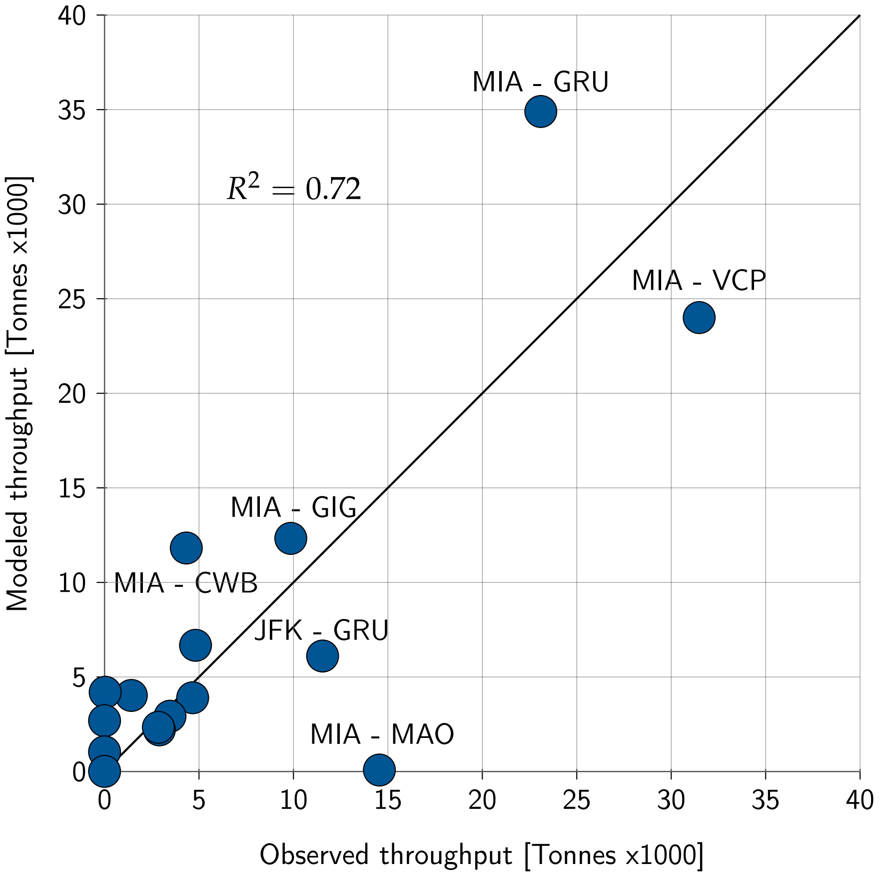

Figure 15: US-BR case study: Calibration result for airport to airport flows

5.4.2 *Verification*

In this subsection the now calibrated and validated model is used as a starting point. In order to check if the model if robust, we perform a sensitivity analysis by changing some parameters.

Adapt airport impedance

Firstly the airport impedance of GRU is changed, where the current value of $1,450$ is adapted. By decreasing the A_p-value, the airport will become more attractive for the model to assign flows to. Therefore the hypothesis is that the airport throughput will decrease when increasing the value.

In Figure 16 the modeled for all the airports is presented when changing the airport impedance of GRU over a range between 0 - $10,000$. The throughput of GRU is represented by the orange line. The calibrated A_p-value is shown by the dotted line. The figure shows a clear increasing trend when the impedance of this airport is decreased. At a A_p-value of 0 the model puts more air cargo flows towards GRU, where it increased from 55 tonnes with 44% to 79 tonnes. When the impedance value is increased, the orange line gradually decreases,

ROUTE	OBSERVED TRAFFIC FLOWS (TONNES)	ESTIMATED TRAFFIC FLOWS (TONNES)
MIA-VCP	31,482	24,003
MIA-GRU	23,096	34,900
MIA-MAO	14,550	80
JFK-GRU	11,543	6,108
MIA-GIG	9,859	12,328
DFW-GRU	4,816	6,674
LAX-GRU	4,665	3,902
MIA-CWB	4,328	11,819
ORD-GRU	3,464	2,925
JFK-GIG	2,901	2,203
DFW-GIG	2,847	2,337
MIA-POA	1,419	4,023
JFK-VCP	33	4,194
LAX-VCP	0	2,688
ORD-GIG	0	1,042
LAX-MAO	0	9

Table 12: Air cargo flows on routes from US regions towards Brazil (only direct routes displayed)

which confirms the previously stated hypothesis. At the same time the throughput of the other Brazilian airports increase because they become relative more attractive for trade. The airports which mainly increasing are presented by the blue lines. The throughput of MIA is very stable in this analysis. Trade from this airport is flowing towards Brazil via other routes than towards the main airport. The same holds for the airports of JFK and LAX. For the US airport, the throughput of ORD shows the biggest decline (-10%), which is displayed by the red line. This airport has only two direct routes towards Brazil which one of them now gets less attractive. All of the results presented and discussed is a logical result of changing the impedance value of GRU.

Secondly the airport impedance of MIA is changed. Figure 17 shows the result when the A_p-value is changed along a range from 0 - $10,000$. The current value is 723, which is shown by the dotted line. When running the model with this change we see that this the throughput of MIA is decreasing when the impedance gets higher . Where air cargo flows which usually used MIA as connecting airport to transport trade to Brazil, are now a little bit shifted towards DWF and JFK. However, the orange line is not as steep as we saw in the previous case. Even when the impedance is set at $10,000$, MIA is still the third US

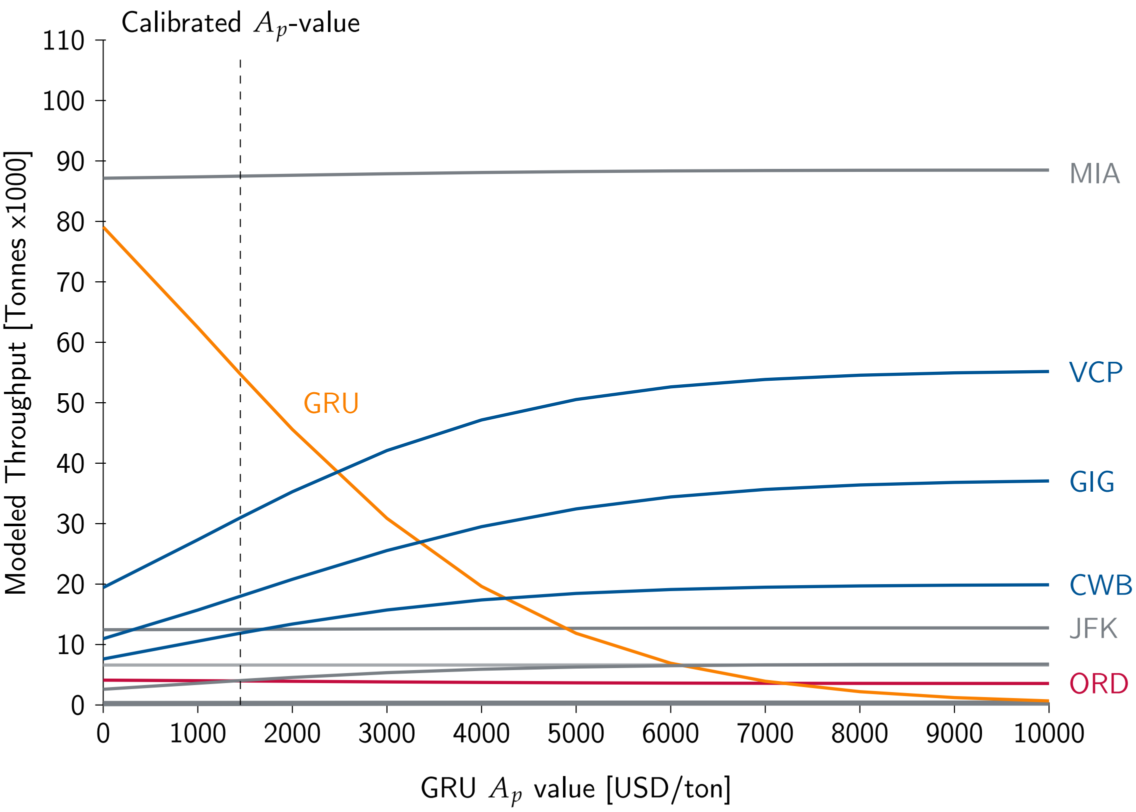

Figure 16: Airport throughput estimation after adaption of A_p-value of GRU

airport in terms of throughput after JFK and DFW. The geographical favorable position related to Brazil plays the most important role in this. The Brazilian airports which are mainly dependent on MIA as main throughput airport are the most affected, which is presented by the red lines. GRU is on its turn, profits from this change, because it is better connected with direct flights from the other US airports. All of the results presented and discussed is a logical result of changing the impedance value of MIA.

Adding a route

As a second part of the sensitivity analysis an airline service route is added to the network. In the current model there is no direct connection between US North West and Brazil. Trade from this region is transported via one of the other regions with a domestic flight, where it has to be transshipped at one of the other US airports. Therefore we did not saw any direct estimated flow between SEA and Brazil in section 5.4.1. To see how the output of the model will react on the addition of a new route operated by a potential new airline in the market, we add the route SEA-GRU to the service network.

After running the model the air cargo flows are assigned as presented in Table 13. On the added direct route SEA-GRU the model

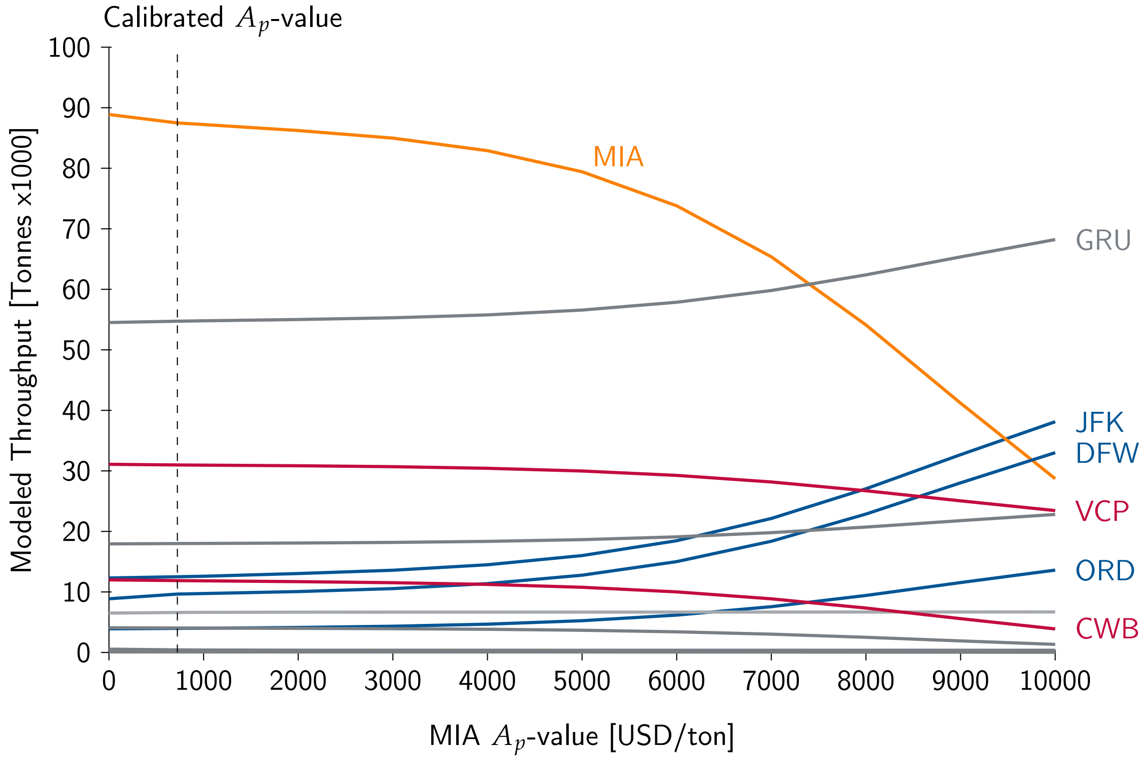

Figure 17: Airport throughput estimation after adaption of A_p-value of MIA

assigns a total of 1,203 Tonnes. The model clearly captures the opportunity to use the direct service option to transport the trade. On the routes LAX-GRU and LAX-VCP there is both a decrease of 10% of assigned flow. From this we can conclude that in the standard model, without the direct SEA-GRU route, trade from USNW was mostly flown towards LAX. Other air cargo flows which decreased are ORD-GRU and ORD-GIG, so also ORD was used as transshipment airport to transport trade from Seattle to Brazil. The R-squared for airport to airport flows remained at the value of 0.72.

Closing a route

In this exercise we will see how the model reacts when a route is removed from the network. The route between New York (JFK) and Viracopos (VCP) which accounts for almost 5,000 Tonnes of yearly trade between USNE and Brazil is closed.

After running the model the air cargo flows are assigned as presented in Table 14. Obviously, the model is not allocating trade on the JFK-VCP route, which is now decreased to zero. Routes to GRU and GIG from JFK are now increased by 45% in terms of tonnes of trade. Those routes are increased by respectively 2,800 and 1,000 tonnes. Allocation of the trade, which was previously on the JFK-VCP route, on these routes can be seen as a reasonable result from the model since the airports of GRU and GIG are closely located to GRU. The

ROUTE	Estimated traffic flows (Tonnes)		CHANGE
	BEFORE	AFTER	(%)
MIA - GRU	34,900	34,899	
MIA - VCP	24,003	24,009	
MIA - GIG	12,328	12,328	
MIA - CWB	11,819	11,814	
DFW - GRU	6,674	6,628	
JFK - GRU	6,108	6,107	
JFK - VCP	4,194	4,193	
MIA - POA	4,023	4,022	
LAX - GRU	3,902	3,497	-10%
ORD - GRU	2,925	2,904	
LAX - VCP	2,688	2,409	-10%
DFW - GIG	2,337	2,320	
JFK - GIG	2,203	2,202	
ORD - GIG	1,042	1,034	
MIA - MAO	80	80	
LAX - MAO	9	8	
SEA - GRU	-	1,203	-

Table 13: Air cargo flows on direct routes from US regions towards Brazil when direct SEA-GRU service is added

R-squared for airport to airport flows increased to 0.74 because BTS does not report air traffic flows on the removed JFK-VCP route.

Adapting α

The last exercise of the verification of the US-BR case study is the adaption of α. Hereby we focus the analysis on the Brazilian airports since this parameter is mainly influencing the costs of trucking transport. Where the US airports are also the 'country' centroid, trucking will not be applicable that much. At Brazil, traded goods arriving at one of the airports have to truck towards the country centroid which is located in São Paolo. With a current value of $4,750$ USD/day/ton, α will be adapted along a range from $1,000$ - $10,000$. Analysis before $1,000$ will not be analyzed because transport of goods will most likely be performed via other modes of transport like rail and sea.

Figure 18 presents the modeled throughput while adapting α. A low α means that the total costs of a route will be determined more by distance. A high VOT means that the time it takes to transport goods from A to B will be more valuable. Therefore, the more you move to

ROUTE	Estimated traffic flows (Tonnes)		CHANGE
	BEFORE	AFTER	(%)
MIA - GRU	34,900	34,972	
MIA - VCP	24,003	24,059	
MIA - GIG	12,328	12,354	
MIA - CWB	11,819	11,843	
DFW - GRU	6,674	6,688	
JFK - GRU	6,108	8,883	45%
JFK - VCP	4,194	-100%	-
MIA - POA	4,023	4,031	
LAX - GRU	3,902	3,902	
ORD - GRU	2,925	3,069	5%
LAX - VCP	2,688	2,689	
DFW - GIG	2,337	2,342	
JFK - GIG	2,203	3,203	45%
ORD - GIG	1,042	1,093	
MIA - MAO	80	81	
LAX - MAO	9	9	

Table 14: Air cargo flows on direct routes from US regions towards Brazil when direct JFK-VCP service is removed

the left of the x-axis we call it a distance-led assignment, and the more we move to the right we call a time-led assignment.

What has to be kept in mind when analyzing this graph is that the adaption of α is done *after* the calibration of the model. This means that the assigned impedance parameter values have more influence on the route cost at lower values of α. This can immediately be seen at the lines of the airports of Manaus (MAO) and Porto Alegre (POA), which show a steep growth in attractiveness when the model is more distance-led. During the model fitting especially those airports have been assigned with a relative low impedance which makes them now more attractive. For the same reason Campinas (VCP) decreases in modeled throughput, because it have been assigned with the highest impedance value after calibration. The airports of São Paolo (GRU) and Campinas (VCP) increase the most when the model is more time-led. The fact that both airports are located the closest to the centroid of Brazil and therefore have relative low trucking costs, results that these airports increase in attractiveness. They are also in the most favorable position when in comes to direct routes from the U.S. airports.

We can conclude that when taking into account that the modeled throughput, apart from changing α, also is influenced by the A_p-values of each airport, the graph presented shows a logical output.

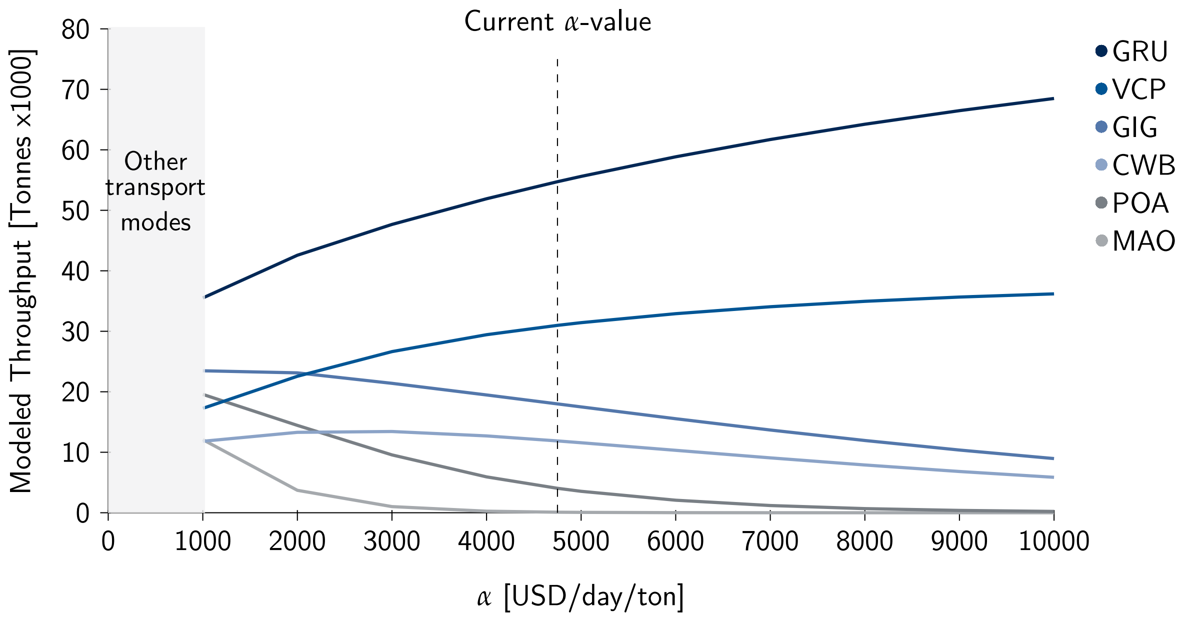

Figure 18: Airport throughput estimation after adaption of α

5.5 SUMMARY AND CONCLUSIONS

This chapter described the US-BR case study which had the purpose to prove the capabilities of the model to reproduce air cargo flows from airport to airport and yearly airport throughput. First, a detour factor an capacity share analysis was performed from which we could conclude that potential interference of O/D trade is minimized within the trade market between the US regions and Brazil. Therefore it could be used as a proper case study.

Secondly, the case study was set-up, which includes six US and six Brazilian airports and a service network to distribute the trade between those countries. For calibration purposes, assumptions had to be made to determine the observed throughput data for the twelve airports. After calibration, validation analysis showed that the model is able to reproduce the airport throughput very accurate, with a R-squared of 95%. Also the air cargo flows between the airports are estimated rather accurately.

Furthermore, verification of the model has proven that the model is robust when parameters are changed. From this we could conclude that the model is able to estimate air cargo flows with an acceptable accuracy and can be used to apply on a global scale.

6 FULL WORLD

After the proven capabilities of the model in the US-BR case study, including the validation and verification, in this chapter the model is applied on a global scale. The Full World model first needs a structured set-up in order to capture the global air cargo market, which includes total trade, airport throughput and network capacity. Secondly, the model is calibrated and the results are analyzed.

6.1 SET-UP

In order to create a Full World model a set of airports, countries and flight services is created which serve as the main input data of the model. Taking all WB, passenger plus freighter operations, and NB freighter operations in 2014 results in a set of +11,000 distinct flight routes, 816 airports and 203 countries. Unfortunately the model gives an *out of memory error* when the choice sets are generated. Therefore the size of the set has to be decreased, while still capturing the (1) air trade between countries, (2) airport throughput and (3) network capacity. Hereby we noted some rules for the model:

- Trade between countries is only captured when both origin country as the destination country are included in the set
- A country has to have an airport which can serve as a main gateway for import and exporting air trade
- Network capacity of an airport includes all air cargo capacity, inbound and outbound, flown by WB freight, passenger, and NB freight scheduled services in 2014

A total of three selection criteria are created to set-up a combination of airports and countries which can serve as a feasible input set for the model to run on a global scale.

Selection 1 - Throughput > 5,000 Tonnes

The first selection created is a minimum threshold for the reported airport throughput. The threshold is set on 5,000 Tonnes, which results in a total amount of 282 airports. The criteria assures that almost all the major airport, which serves as country gateways for international global trade, are included in the set. To put this threshold into perspective, airports which report a throughput around 5,000 tonnes are for example: PDX (Portland, USNW), 6,384 Tonnes; GEO (Georgetown, Guyana), 6,082 Tonnes; MPM (Maputo, Mozambique), 5,165 Tonnes. The 282 airports selected by this threshold capture already:

- 91% of network capacity
- 99% of airport throughput
- 96% of air trade

Selection 2 - Trade addition > 5,000 Tonnes

The second selection is based on a threshold of a minimum trade addition of 5,000 tonnes. This means that all airports, which are located in a country which is not yet represented by set of 282 airports and this country has a total trade (import + export) above 5,000 tonnes, will be added to the set. The main purpose of this selection of airports is to add countries which are not represented by an airport yet, which means that trade of that specific country is not captured yet. The top 10 of airports in terms of trade captured by this filter are shown in Table 15. A total of 21 airports where added to the set. These airports where not selected with the first criteria because, throughput is only reported in totals (international + domestic), or is not reported at all . Observed throughput for the latter case is based on historical data or a capacity share.

AIRPORT CODE	CITY	COUNTRY	TRADE CAPTURED (TONNES)	NETWORK CAPACITY (TONNES)
DAC	Dhaka	Bangladesh	184,147	460,034
CPH	Copenhagen	Denmark	155,232	280,532
CCS	Caracas	Venezuela	78,359	195,488
LAD	Luanda	Angola	62,151	116,666
EBL	Erbil	Iraq	62,053	105,851
GUA	Guatemala City	Guatemala	60,160	63,663
GYD	Baku	Azerbaijan	47,736	292,918
KBL	Kabul	Afghanistan	25,034	53,365
TAS	Tashkent	Uzbekistan	23,763	81,432
MRU	Port Louis	Mauritius	21,133	115,560

Table 15: Top 10 airports in terms of trade captured by second selection

Selection 3 - Add connection airports

The third and final selection is applied is to capture the airports which are located in a country which is already in the current set, but could serve as an important connection airport in the global network. In terms of connectivity we mainly selected on the network capacity of the airport. If reported, also the total throughput is used as a guideline to investigate if airports could be added to the selection. Again, these airports where not selected with the first selections because international throughput is not accurately reported on a yearly basis by ACI or that there is no clear split between international and domestic handled freight. Observed throughput is based on historical data or a capacity share. Top 10 captured airports in terms of network capacity by this selection are shown in Table 16. A total of 15 airports where added to the set.

AIRPORT CODE	COUNTRY	CITY	TRADE CAPTURED (TONNES)	NETWORK CAPACITY (TONNES)
SYD	Sydney	Australia	-	1,013,936
CGO	Zhengzhou	China North	-	451,130
YVR	Vancouver	Canada	-	384,454
CKG	Chongqing	China West	-	270,947
CPT	Cape Town	South Africa	-	133,267
DUB	Dublin	Irish Republic	-	124,190
YYC	Calgary	Canada	-	110,608
TSA	Taipei	Taiwan	-	106,966
DLC	Dailan	China North	-	84,641
MSP	Minneapolis	US North Central	-	77,381

Table 16: Top 10 airports in terms of network capacity captured by third selection

Those three selection criteria together create a set of 318 airports representing 148 countries and a network consisting out of +7,500 distinct flight routes which capture:

- 96% of network capacity
- 99% of airport throughput
- 98% of air trade

With the generated set the model is able to run the Full World in less than one minute without generating any errors.

6.2 CALIBRATION

In order to calibrate the model, observed airport throughput from ACI is used. On top of that, to be able to use the multi-objective optimization which also calibrates for transshipments per airport, we would like to have the observed transshipments for each airport in the set. Unfortunately, this data is not included within the ACI data. Therefore we decide to use public data sources and academic literature to estimate yearly transshipment in tonnes for the top 14 cargo airports in terms of yearly throughput. Those 14 airports represent almost 50% of global airport throughput which will be a proper dataset to calibrate the model. Table 17 shows the transshipment rates we found and the related amount in tonnes.

AIRPORT CODE	THROUGHPUT (TONNES)	TRANSSHIPMENT RATE	TRANSSHIPMENT (TONNES)
HKG	4,376,349	54.0%	2,363,228
ICN	2,474,152	50.1%	1,239,550
DXB	2,367,574	85.0%	2,012,438
PVG	2,334,368	40.0%	933,747
TPE	2,072,602	58.5%	1,211,612
NRT	2,043,372	19.4%	396,414
FRA	2,007,318	1.5%	30,110
CDG	1,858,482	1.5%	27,877
SIN	1,843,800	50.0%	921,900
ANC	1,787,287	38.0%	679,169
MIA	1,739,005	44.0%	765,162
AMS	1,633,195	1.5%	24,498
LHR	1,497,701	1.5%	22,466
BKK	1,193,300	42.4%	505,817

Table 17: Transshipment in tonnes for the Top 14 cargo airports (Air Cargo World, 2017; Chung, 2013; Emirates, 2017; FRA Airport, 2015; Ohasi, 2005; Miami International Airport, 2014; Seabury Analysis)

6.2.1 μ calibration

As described in the US-BR case, firstly the scale parameter μ being adapted. The result of the absolute throughput difference for different values of μ is presented in Figure 19. An optimal value of 0.00015 is found. This value differs from the μ found in the US-BR case study, where the value was equal to 0.0006 . This d with the

What again can be noticed that the average size of a choice set in the Full World model is larger than in the US-BR case study. The 148 countries in the model have an amount of available ports per country between 1 and 56. This increase of choice sets again means a lower μ-value to increase the variance within the set.

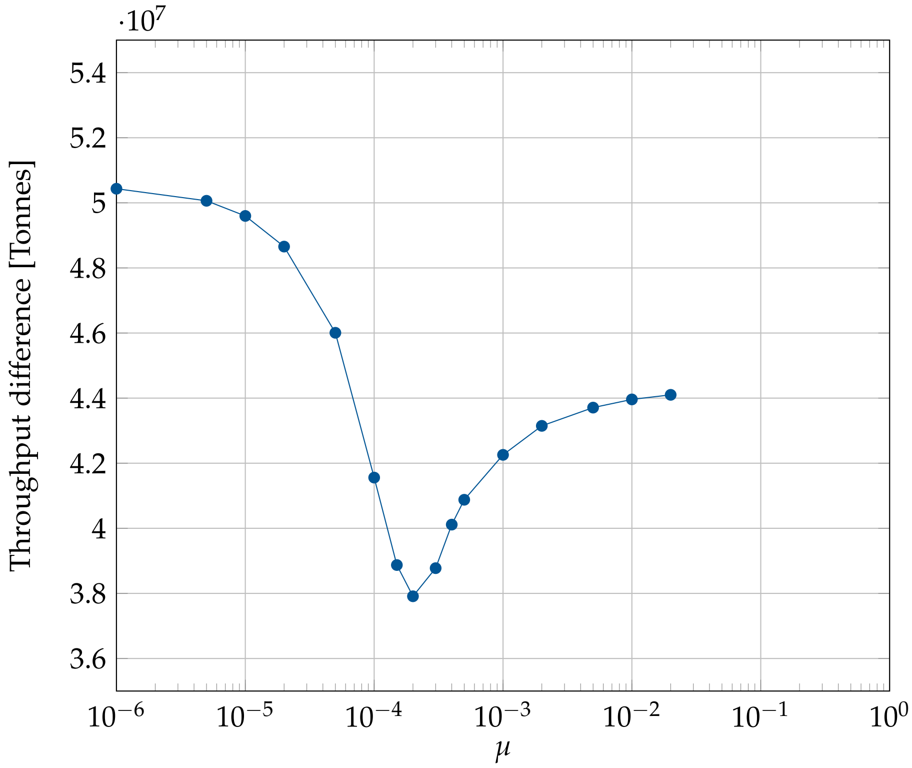

Figure 19: Comparison of the absolute difference between observed and calculated throughput for different scale parameter μ (logarithmic scale)

6.2.2 A_p *calibration*

After the calibration of the scale parameter, we proceed by adapting the airport impedance parameter A_p. Other than at the US-BR case, this time we did not choose to start the calibration with only setting lower and upper boundaries for A_p. In order to improve the process of fitting the model, a *normalized impedance parameter* is calculated for every airport by normalizing the accuracy of the modeled throughput compared to the observed throughput. This is done with the following formula:

$$N_p = \frac{TPR_p - min(TPR)}{max(TPR) - min(TPR)} \tag{8}$$

where N_p, the normalized impedance parameter of airport p; TPR, the throughput ratio of airport p, which is defined as:

$$TPR_p = \frac{mTP_p}{oTP_p} \tag{9}$$

where mTP_p, the modeled throughput of airport p; oTP_p, the observed throughput of airport p. The calculated N_p value will scale the total airport set with values in between 0 and 1. In order to let the value be able to be a proper part of the route choice and therefore be able to differentiate the impedance values of the airports, the N_p value is multiplied by $5,000$. For comparison, a general transpacific route (PVG - JFK) has around $2,000$ of distance-related costs and $15,000$ time-related costs.

Running the model with the calculated N_p values already resulted in a decrease of more than 10 million tonnes of absolute throughput difference which is a decrease of 28 %. This first step of fitting the model by changing the airport impedance parameter is clearly shown in figure 20. After this, a total of 10 runs have been performed, whereby each run consisted of 1,000 iterations. After the first run the algorithm converged gradually towards the minimum obtained absolute throughput difference of $26,483,719$ tonnes, which an improvement of more than 31% compared to the value before the A_p calibration.

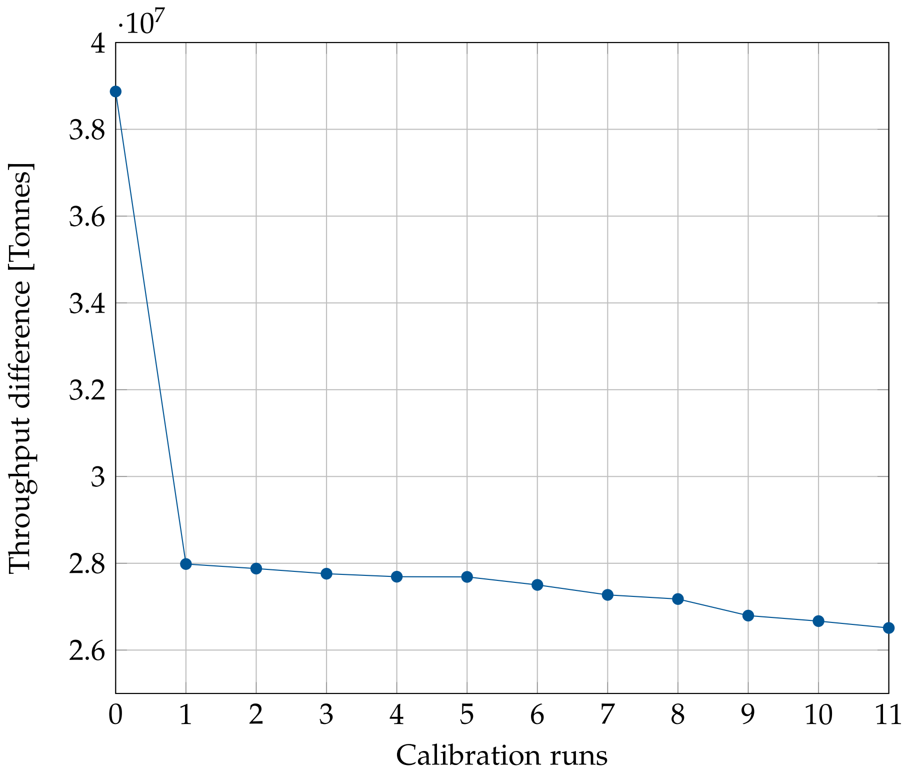

Figure 20: Convergence in throughput difference after different calibration runs

6.3 RESULTS

Figure 21 shows the model output on a global scale. The blue lines show the volume of air traffic of flows above $15,000$ tonnes a year. The size of the pie charts visualize the amount of throughput (dark grey) and transshipment (light grey) at each airport. It is clearly shown that part of the throughput of some large airports is made out of transhipments, like for example at Tokyo (NRT), Paris (CDG) and Hong Kong (HKG). The Transatlantic and Transpacific trade flows are clearly represented. Furthermore, the connecting hubs of Dubai (DXB) for the European - Asian market, Anchorage (ANC) for the Transpacific market and Miami (MIA) for the North America - Latin American market, are shown. On the map one can also see the inland flows from Brazilian airports to/from Brasilia and the inland flows in India to/from New Delhi. Because flows in some regions are below 15,000 tonnes a year, those are not shown on the map (e.g. Africa, islands).

Table 18 presents the top and bottom five calibrated A_p-values for the airports in the model. HGH, an airport located near the city of Shanghai in the China East region, has the highest assigned airport impedance value, where a total of five airports have assigned the lowest value of zero. A total overview of all the airports including A_p-values can be found in Appendix B.

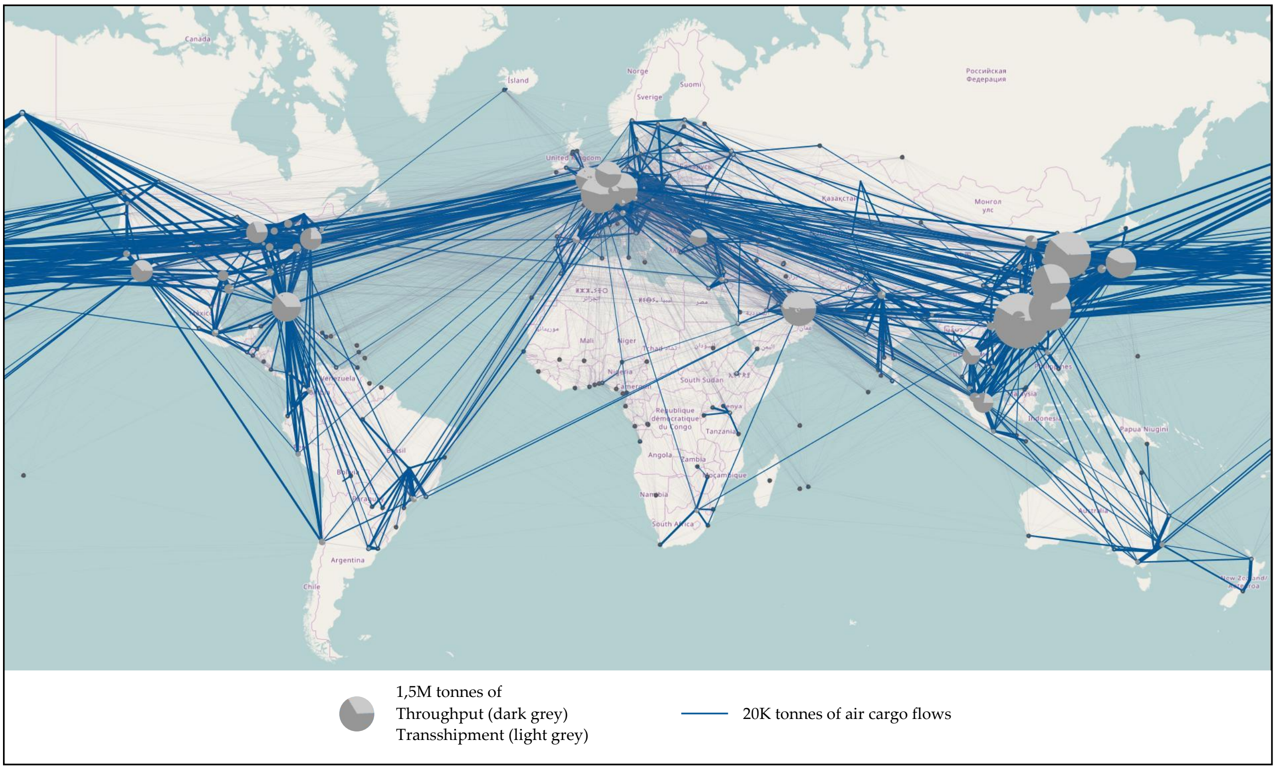

Figure 21: Global air cargo flows between countries (>15 K Tonnes/year)

AIRPORT CODE	CITY	COUNTRY	A_p-VALUE
HGH	Hangzhou	China East	19,596
BEG	Belgrade	Serbia	13,481
PEW	Peshawar	Pakistan	11,253
HKT	Phuket	Thailand	11,125
NKG	Nanjing	China East	10,844
...	...	...	...
FDF	Fort-de-France	Martinique	0
ANC	Anchorage	US North West	0
SIN	Singapore	Singapore	0
DXB	Dubai	United Arab Emirates	0
HKG	Hong Kong	Hong Kong	0

Table 18: Top and bottom 5 calibrated A_p-values per airport

6.4 ANALYSIS

After the model converged it resulted in a fit which explains 88% of the amount of variation between airport throughput volumes (Figure 22). With this high R-squared value, we can conclude that the model is able to reproduce airport throughput rather accurately. On the correlation graph we can identify group of airports which have the same characteristics. These airports will be discussed in the rest of this section.

Top 20 airports

Top 20 airports in terms of observed throughput have also the highest estimated throughput, as can be seen in the correlation graph. A more detailed overview of the estimated throughput accuracy for those airports is presented in Table 19. The majority of the TPR, the parameter introduced in the previous section which represents the modeled throughput divided by the observed throughput of an airport, is within an accuracy range of 40%, with Los Angeles (LAX), New York (JFK) and Chicago (ORD) even within 5%.

The airports of Anchorage (ANC) and Leipzig (LEJ) show the lowest ratios. Both airports are known as hubs of integrator flights for DHL, FedEx and UPS. A further analysis for these integrator hubs is performed later in this section.

Other airports with a relative low TPR are the airports of Abu Dhabi (AUH), Doha (DOH) and Singapore (SIN). Because both AUH and DOH are located near Dubai International Airport (DXB), which is

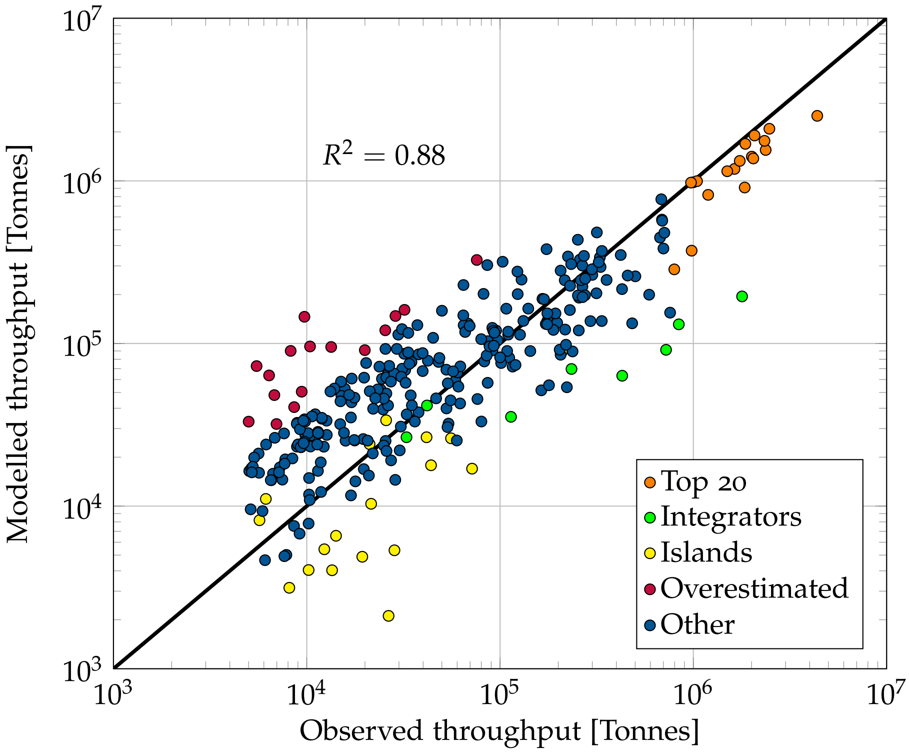

Figure 22: Full World: Calibration result for modeled vs observed airport throughputs (Note: logaritmic scale)

number three in terms of total cargo throughput, it maybe experience some competition in terms of attractiveness. SIN has a ratio of 0.49. This has mainly to to with the low value of estimated transshipment at the airport which is lower that 30% of the observed transshipment figure. It seems that the model is not able to capture the transshipment value of the airport of Singapore. The use of the shortest path algorithm can be the reason for the underestimation of airports which are connecting trade markets by transshipments. Also the fact that the focus of the calibration lay at the throughput and not the transshipments, could be a factor.

Integrator hubs

Looking again at the airport throughput correlation graph, one can notice a clear range of underestimated airports in the higher segment in terms of observed throughput. Those four airports marked in green, can be identified as the integrator hubs (ANC, LEJ, CGN, SDF). For analyzing purposes, also other airports with known integrator activity are marked green (EMA, BGY, ONT, VIT). Integrators, like DHL, FedEx and UPS are the main visitors of these airports and so determine the network capacity.

In Table 20 an overview is given of the eight airports including the network capacity and amount of routes. As explained in section

AIRPORT CODE	CITY	COUNTRY	OBSERVED THROUGHPUT 2014 (TONNES)	TPR
HKG	Hong Kong	Hong Kong	4,376,349	0.57
ICN	Seoul	South Korea	2,474,152	0.85
DXB	Dubai	UAE	2,367,574	0.65
PVG	Shanghai	China East	2,334,368	0.76
TPE	Taipei	Taiwan	2,072,602	0.92
NRT	Tokyo	Japan	2,043,372	0.67
FRA	Frankfurt	Germany	2,007,318	0.70
CDG	Paris	France	1,858,482	0.91
SIN	Singapore	Singapore	1,843,800	0.49
ANC	Anchorage	US North West	1,787,287	0.11
MIA	Miami	US South East	1,739,005	0.76
AMS	Amsterdam	Netherlands	1,633,195	0.73
LHR	London	United Kingdom	1,497,701	0.77
BKK	Bangkok	Thailand	1,193,300	0.69
LAX	Los Angeles	US South West	1,044,835	0.95
JFK	New York	US North East	1,002,569	0.98
DOH	Doha	Qatar	980,114	0.38
ORD	Chicago	US North Central	972,803	1.00
LEJ	Leipzig/Halle	Germany	840,987	0.16
AUH	Abu Dhabi	UAE	797,069	0.36

Table 19: *TPR* results for top 20 cargo airports

6.1, the calculated network capacity includes all air cargo capacity, inbound and outbound, flown by WB freight, passenger, and NB freight scheduled services in 2014. This does not include charter and integrator flights because they are not in the service data used from Innovata. This is the reason why the integrator hubs have such a low capacity and a relative low amount of routes in the model. The airport of Memphis (MEM), home to the FedEx Express global hub, is not even selected in the set-up stage because it is not included in one of the routes.

In order to compare the network capacity and amount of routes, comparable airports in terms of observed throughput are also listed in the table and are highlighted in blue. It is clearly shown that those airports where the presence of integrator operations is less dominant have a higher network capacity and amount of routes.

A relative low amount of route options means that the probability that those airports are used for trade flows will decrease. This is the main reason why those airport are underestimated in general. The reason why Ontario (ONT) and Vitoria (VIT) are exceptions in this case is because they are not that big in terms of observed throughput and thereby easier to estimate.

The missing network capacity as a result of the lacking of charter and integrator data also play an important role in terms of connectivity where ANC for example, is one of the main connector airports for transpacific trade. The absence of integrator and charter routes can therefore also effect other modeled throughput.

AIRPORT CODE	OBSERVED TP (TONNES)	NETWORK CAPACITY (TONNES)	ROUTES	INTEGRATOR(S)	TPR
ANC	1,787,287	1,872,944	151	UPS, FedEx, DHL	0.11
AMS	1,633,195	2,002,438	314	-	0.73
ORD	972,803	1,641,058	221	-	1.00
LEJ	840,987	88,053	14	DHL	0.16
CGN	722,584	33,306	12	UPS, FedEx	0.13
MXP	457,424	725,355	175	-	0.57
SDF	428,599	40	2	UPS	0.15
EMA	234,242	29,232	19	DHL	0.30
BGY	113,946	582	1	DHL	0.31
ONT	41,872	48	2	UPS	0.99
VIT	32,816	28	2	DHL	0.81

Table 20: *TPR* results for integrator hubs

'Island' airports

Airports which also are underestimated by the model but are in the lower segment in terms of observed airport throughput can be identified as 'island' airports, which are mapped in Figure 23. A more detailed overview of the throughput ratio is presented in Table 21. The reason of this underestimation for this specific group can be found in the fact that these airports do not have the advantage of RFS towards available airports in neighbouring countries. This results in less options for traded goods to reach or depart the islands. Another reason is the use of the shortest path algorithm. The probability that the route IST - MLE - CMB (Istanbul - Maledives - Colombo) will be added to the choice set is lower than the direct option IST - CMB.

Airport which are located on an island but are not very underestimated ($TPR > 0.60$) by the model form the exemptions in this group. The majority of these airport are profiting from an geographical position which makes them attractive for transshipments. For example, the main airport of Iceland in Reykjavik (KEF), has a throughput ratio of 0.64 but this island is located on a favorable location on the transatlantic route between Europe and North America.

Figure 23: 'Island' airports: low throughput airports which are underestimated by the model

Overestimated airports

As presented in the correlation graph, most of the dots are located above the black 45 degree line and therefore are overestimated by the model. This automatically means that those airports have a TPR value of more than one. Where TPR values of the overestimated airports range till 15.01, in this section the airports with a TPR value above 4.00 are discussed (Table 22). What can be noticed from the 16 airports which are above this threshold is that most of them are located in the US or China, two countries which are represented in the top when looking at total trade (Figure 24). Also Japan and Korea are accounting for a big part of imports and exports of global traded goods. A second observation from this table is that the overestimated airports are relatively small in terms of observed throughput, with more than half of them below $10,000$ tonnes.

There seems to be a relation between the available airports of a country, so airports within its country plus the ones which can be

airport code	city	country	observed tp 2014 (tonnes)	tpr
FDF	Fort-de-France	Martinique	9,913	0.01
PTP	Pointe A Pitre	Guadeloupe	10,996	0.02
GUM	Guam	Guam	12,359	0.06
RUN	Saint Denis	Reunion	26,503	0.08
BWN	Bandar Seri Begawan	Brunei Darussalam	28,425	0.19
MLE	Male	Maldives	71,658	0.24
BGI	Bridgetown	Barbados	19,380	0.25
STI	Santiago	Dominican Republic	13,522	0.30
SEZ	Mahe Island	Seychelles	8,124	0.39
KIN	Kingston	Jamaica	10,224	0.40
POS	Port Of Spain	Trinidad & Tobago	43,948	0.41
CMB	Colombo	Sri Lanka	208,673	0.43
PPT	Tahiti	French Polynesia	12,328	0.44
MLA	Malta	Malta	14,170	0.46
SDQ	Santo Domingo	Dominican Republic	55,753	0.47
PUJ	Punta Cana	Dominican Republic	21,532	0.48
KEF	Reykjavik	Iceland	41,669	0.64
MRU	Mauritius	Mauritius	21,133	1.14
LCA	Larnaca	Cyprus	25,738	1.31
SSG	Malabo	Equatorial Guinea	5,695	1.44
MBJ	Montego Bay	Jamaica	6,143	1.80

Table 21: *TPR* results for 'island' airports (lowest on top)

reached via RFS, and the amount of overestimated airports (Table 23). Where for example Germany, which has a second position in terms of total trade, has no overestimated airports with a TPR higher than 4.00. Besides the nine airports Germany has within its own borders, due to the trucking possibility, it has an extra of 44 other European airports to distribute its trade. This is probably an advantage for the fitting of the model of assigning an impedance value which results in properly reproducing the airport throughput.

The two above observations from above combined is a logical result if the current methodology of the model is being followed. Trade will be distributed along the shortest paths which are linked to the available airports of the O/D countries. If the amount of available airports is relatively low and the trade importing and/or exporting the country is relatively high, the distribution is facing overestimation problems. As an addition to that, minimizing the sum of absolute difference between observed and modeled airport throughput forces the model to focus on the more larger airports, to get to the most optimal result. Minimizing a relative big difference (for the large airports) results in a higher convergence than minimizing a relative small difference (for the small airports). Overestimation is a result from that method, because overestimating smaller airports will less 'harm' the absolute throughput difference than underestimating larger airports.

Figure 24: Overestimated airports: low throughput airports which are overestimated by the model

CG exception airports

The last group of airports to be discussed is not highlighted in the correlation graph but which is worth analyzing, are the CG exception airports. In section 4.2.1 the assumption that the cargo center of gravity (CG) lies at the capital city of a country was discussed. Here we presented a total of ten country exceptions were big cargo airports in terms of throughput would move that GC away from the capital city. Hereby the hypothesis is that those airports which are located near the GC but are not the main airport of the country in terms of throughput will be overestimated. In table 24 is shown what the TPR value is of the airports which are not the main gateways of the country in terms of throughput, but are located near the centroid of the country. For six of the ten countries the throughput is too low to pass the selection criteria in the set-up of the Full World model. The other four show a clear overestimation with Berlin (TXL), the main international airport of Berlin, with the highest TPR. Moving the CG of those airports will result in a better estimation of airport throughput.

AIRPORT CODE	CITY	COUNTRY	OBSERVED TP 2014 (TONNES)	TPR
WUH	Wuhan	China North	9,727	15.01
SJC	San Jose	US South West	5,502	13.20
SHA	Shanghai	China East	8,256	10.91
PDX	Portland	US North West	6,384	9.97
PUS	Busan	South Korea	10,397	9.22
XIY	Xi'an	China West	13,378	7.11
LAS	Las Vegas	US South West	6,799	7.08
TPA	Tampa	US South East	5,006	6.61
KMQ	Komatsu	Japan	9,416	5.37
MFM	Macau	Macau	28,769	5.13
IND	Indianapolis	US North Central	32,045	5.03
CLT	Charlotte	US South East	25,479	4.73
AUS	Austin	US South Central	8,605	4.72
MID	Merida	Mexico	6,975	4.58
SHE	Shenyang	China North	19,969	4.56
DTW	Detroit	US North Central	75,810	4.30

Table 22: TPR results for overestimated airports (TPR > 4.00, highest on top)

RANK TOTAL TRADE	COUNTRY	AVAILABLE AIRPORTS	$tpr > 4.00$
1	China East	21	1
2	Germany	53	0
3	Japan	21	1
4	US North Central	15	2
5	China South	17	0
6	US South East	18	2
7	US North East	14	0
8	South Korea	17	1
9	United Kingdom	40	0
10	India	18	0
11	China North	26	2
12	Netherlands	47	0
13	US South Central	4	1
14	US South West	7	2
15	Italy	46	0

Table 23: Relation between main trade countries and their available airports in the model to distribute trade

COUNTRY	CAPITAL CITY	AIRPORT CODE NEAR CAPITAL	TPR
Germany	Berlin	TXL	3.46
Italy	Rome	FCO	1.04
Turkey	Ankara	ESB	2.36
Vietnam	Hanoi	HAN	1.27
Brazil	Brasilia	BSB	-
Canada	Ottawa	YOW	-
Kazakhstan	Astana	AST	-
Iraq	Bagdad	SDA	-
Morocco	Rabat	RBA	-
Tanzania	Dodoma	DOD	-

Table 24: TPR results for airports near GC of exception countries

6.5 summary and conclusions

This chapter described the Full World model, a model which estimates global air cargo flows based on O/D trade data. First, a set of airports, countries and flight services was created by defining three selection criteria. Those selection criteria created a set of 318 airports, representing 148 countries and a network consisting out of +7,500 distinct flight routes. With this it captured 96% of network capacity, 99% of airport throughput and 98% of air trade and was able to estimate the air cargo flows on a global scale within one minute.

After converging of the model due to calibration, it resulted in a fit which explains 88% of the amount of variation between airport throughput volumes. With this high R-squared value, we could conclude that the model is able to reproduce airport throughput rather accurately.

Correlation analysis

By analyzing the correlation graph which presented the modeled throughput compared with the observed throughput, we identified four group of airports, namely, top 20 airports, integrator airports, 'island' airports and overestimated airports. We also discussed the airports which are located near the assumed cargo center of gravity. We can draw the following conclusions for those groups:

- **Top 20 air cargo airports** are rather accurately reproduced in terms of throughput, but underestimated in general. Because the focus in this model was less towards the amount of transshipments, which is an important part of the throughput for those airports, this can be part of the reason why.

- **Integrator airports** are typically underestimated. This is mainly related with the fact that integrator and charter services are not included in the scheduled airline service schedule which is used as main input for the supply side of the model.

- **'Island' airports** are generally underestimated because they are not attractive in terms of hinterland flows. The potential of RFS to distribute flows from or towards neighbour countries is very low because they are geographically isolated.

- **Overestimated airports** are generally be found in countries where the amount of available airports is relatively low and the trade importing and/or exporting the country is relatively high. The smaller the airport, the easier the overestimation by the model.

- **CG exception airports** are overestimated. Moving the CG of those countries towards a more realistic location will result in a better estimation of the airport throughput.

Shortest path & Calibration method

Most of the conclusions stated above are due to two main factors, namely (1) the use of the shortest path algorithm and (2) the method of calibration.

By using the shortest path algorithm, only one route between an airport pair will be added to the choice set. This narrows down the amount of possibilities to distribute the flows and will capture transshipment and island airports to a lesser extend. The number of direct routes and available airports for the distribution of trade is a determining factor in the current model to reproduce airport throughput when a shortest path algorithm is used.

The calibration method of minimizing the sum of absolute gaps between observed and modeled airport throughput forces the model to focus on the more larger airports, for the most efficient approach. This results in a general overestimation of smaller airports, because NSGA-II also minimizes the difference between observed and modeled transshipments. Because in the current model only for the top 14 the (estimated) observed transshipment is inserted, the calibration will be forced towards a value of zero for the rest of the airports, which may not be the case.

In general we can conclude that, despite the fact that the model is able to produce airport throughput in a rather accurate way, some quick wins can be gained by making adaptions in the model and/or calibration method.

7 CONCLUSION AND RECOMMENDATION

This chapter concludes the research performed in this report. The first section provides answers on the research questions posted at the start of the report. The second section shows the recommendation for further study.

7.1 CONCLUSIONS

The research question defined in Chapter 1.1 is: *How can the air cargo flow per flight leg on a worldwide basis be estimated from aggregated traffic data?* To answer this main research question a set of four sub questions were answered first, which are stated below.

1. *Which data regarding air cargo transport is available in order to estimate air cargo flows?*
   Seabury Consulting was the perfect partner in order to fulfill the need of an accurate and granular input of data. The *Seabury Global Trade Database* covers approximately 99% of world trade transported via air, which serves as ideal data for the demand side to estimate air cargo flows. On the supply side, the flight schedules from Innovata was used, with the only disadvantage that is does not include integrator and charter operations.

2. *Which solution technique will be used to estimate the air cargo flow on a flight leg basis?*
   The fundamentals of the methodology used in the World Container Model served as a solution technique to estimate the global air cargo flows. The methodology used was also very suitable from the perspective of the available demand data and supply data. On top of that, the academic research done with the WCM strengthens the fact that it is a proven method which can be used on a global scale.

   The model determines for each O/D country pair a choice set of routes by a shortest path algorithm from airport to airport, based on actual scheduled airline services. The possibility of intermediate transshipments and trucking connections are included. By using a path size logit model together with a generalized cost function, the route choice probabilities are calculated and applied to the O/D trade data for the distribution of the flows.

3. *How and which data should be used in order to test the model on its quality?* The international handled freight per airport from ACI was available. This observed throughput was the main data source of calibrating the model and to improve the quality of the model. Furthermore, BTS and Seabury's Cargo Capacity Database where used for analyzing and validation purposes.

4. *Which steps have to be taken to develop the model further?* The main objective of this research, *'contribute to the development of a model which estimates the air cargo flows'* is accomplished. This work gives possibilities for future research on a number of topics. Possible improvements and extensions are recommended in the next section.

Based on the answers on the sub questions, the main research question can be answered. The answer is practically the Air Cargo Flow model itself including the methodology described in this report.

The novelty in this model lies in the fact that this is the first model in academic literature that estimates air cargo flows in such a detailed airport-to-airport level. Both on the demand side of the model as on the supply side very granular data is used. O/D air trade data on country-to-country level is the main demand input, where route choices are based on scheduled flight services reported by the airline, serve as supply input.

The work presented give airlines and other stakeholders in the aviation industry sufficient insight into the now largely unknown air traffic flows of the competitive air cargo market. This helps them to understand the behavior of competitor airlines, balance the capacity and for route development in order to increase the profitability. On top of that, it will also help them with market analysis, pricing strategies and new route analysis to improve their network.

7.2 RECOMMENDATIONS

This first step in the development of a model which estimates air cargo flows on a global scale, creates numerous opportunities to extend and improve the model, but also for research which is more airport focused. This further research, which now becomes available by the introduction of this version of an unique air cargo flow model, will allow us to study the current and future state of air cargo flows on a global scale.

- **k-shortest path** This research include the shortest path algorithm which creates limitations in how realistic the model is at the moment. Including the k-shortest path algorithm will make the allocation of flows more realistic and will increase the reliability of the calibration process.

- **Capacity \ Frequency** Links and nodes are assumed to have unlimited capacity in this model. The unconstrained network capacity might lead to overestimating the capacity of those links and nodes. A certain flight route, for instance, might be unable to satisfy the whole freight flow as showed in the model. The addition of capacity constrains on routes or trade lanes will improve the quality of the model. Adding flight frequency to capture route attractiveness can also be an option to avoid overestimation.

- **Integrator & charter operations** Flight services used in the model does not include charter and integrator flights. This leads to a relative low network capacity and amount of routes for typical integrator hubs, resulting in the fact that the model is not able to reproduce the airport throughput for these airports. Unfortunately, sufficient data about the flight services of integrator and charter carriers is not available, but including these will improve the estimated throughput of integrator hubs and indirectly the airports connected to those.

- **Transshipment data** This model does only include airport transshipment data for the top 14 cargo airports. Including transshipment data for all the airports in the model will be beneficial for the accuracy of the estimation of airport throughput. Especially when using the multi-objective optimization which minimizes the difference between observed and modeled throughput but also transshipments.

- **Metropolitan airports** The overestimation of the the relative smaller airports in terms of observed throughput can possibly be solved by creating so called 'metropolitan airports', which group airports within a certain area together. By doing this, the output data will be less granular but the objective of estimating air cargo flows on a global scale, can still be accomplished.

All of the recommendations stated above will be subject for future research in order to improve the quality of the model. Besides this, further research into the airport impedance parameters assigned after calibration is also recommended. Especially when the model is extended with the k-shortest path and a capacity constrain, the quality of the A_p-values will increase. All relevant, measurable and hidden service characteristics of airports, such as fuel costs, airport charges, handling costs, congestion costs, etc. are included in this parameter and can be investigated if and how they are related to each other. On top of that, one can do research into the correlation between airport characteristics known from previous research and public data, and the assigned impedance value by the model.

ACI (2014). "ACI World Airport Traffic Report." In: URL: http://www.aci.aero/Publications.

Air Cargo World (2017). "Cathay Pacifc Cargo has busy first year." In: URL: http://aircargoworld.com/cathay-pacific-cargo-has-busy-first-year/.

Airbus (2015). "Global Market Forecast 2015." In: URL: http://http://www.airbus.com/company/market/forecast/.

Beckmann M. McGuire, C. B. and C. B. Winston (1956). "Studies in the Economics of Transportation." In: *Yale University Press, New Haven, CT*.

Bell, M.G.H. (1995). "Alternatives to Dials logit assignment algorithm." In: *Transportation Research Part B* 29.4, pp. 287–295.

Ben-Akiva, M. and M. Bierlaire (1999). "Discrete Choice Methods and their Applications to Short Term Travel Decisions." In: *Handbook of Transportation Science. Kluwer*.

Boeing (2014). "World Air Cargo Forecast 2014-2015." In: URL: http://www.boeing.com/assets/pdf/commercial/cargo/wacf.pdf.

Boonekamp, T. and G. Burghouwt (2016). "Measuring connectivity in the air freight industry." In: *Journal of Air Transport Management*.

Burrell, J.E. (1968). "Multiple route assignment and its application to capacity restraint." In: *4th Int. Symp. on the Theory of Trafic Flow, Karlsruhe.*

Calder, S. (2017). "The world's longest flight prepares for take-off: Doha to Auckland will cross 10 time zones." In: URL: http://www.independent.co.uk/travel/news-and-advice/worlds-longest-flight-doha-auckland-qatar-airways-london-heathrow-a7560531.html.

Campbell, M.E. (1950). "Route selection and traffic assignment." In: *Highway Research Board Bulletin* 1.

Cascetta E. Marzano, V. Papola-A. and R Vitillo (2013). *A multimodal elastic trade coefficients MRIO model for freight demand in Europe. In: Ben-Akiva, M., Meersman, H., & E. Van de Voorde. Freight Transport Modeling*. Emerald Group Publishing Limited, Bingley.

Chang, Y.H. and C.H. Yeh (2007). "A survey and optimization-based evaluation of development strategies for the air cargo industry." In: *International Journal of Production Economics* 106.2, pp. 550–562.

Chu, Hsing-Chung (2014). "Exploring preference heterogeneity of air freight forwarders in the choices of carriers and routes." In: *Journal of Air Transport Management* 37, pp. 45 –52.

Chung Tae-won Han, Jong-khil (2013). "Evaluating Competitiveness of Transshipment Cargo in Major Airports in Northeast Asia:

Airport Branding." In: *The Asian Journal of Shipping and Logistics* 29.43, pp. 377–394.

Crainic T.G. Florian, M. and J.E. Leal (1990). "A model for the strategic-planning of national freight transportation by rail." In: *Transportation Science* 24.1, pp. 1–24.

Crepinsek, M., S.-H. Liu, and M. Mernik (2013). "Exploration and exploitation in evolutionary algorithms: A survey." In: *ACM Comput. Surv.* 45, pp. 1–33.

Daganzo, C. and Y. Sheffi (1977). "On Stochastic Models of Traffic Assignment." In: *Transportation Science* 11.3, pp. 253–274.

Dial, R.B. (1977). "Probabilistic multipath traffic assignment model which obviates path enumeration." In: *Transportation Science* 5, pp. 83–111.

Dijkstra, E.W. (1959). "A note on two problems in connection with graphs." In: *Numerische Mathematik* 1, pp. 269–271.

Emirates (2017). "SkyCentral - Facts & Figures." In: URL: http://www.skycargo.com/english/about-us/dwc/dwc.aspx.

FRA Airport (2015). "Air Traffic Statistics." In: URL: FRA_air-traffic-statistics-2015.

Feng, Bo., Y. Li, and H. Shen (2015). "Air cargo operations: Literature review and comparison with practices." In: *Transportation Research Part C* 56, pp. 263–280.

Frank, M. and P. Wolfe (1956). "An algorithm for quadratic programming." In: *Naval Res. Logist. Quart* 3, pp. 95–110.

Guelat J. Florian, M. and T. G. Crainic (1990). "A multimode multiproduct network assignment model for strategic-planning of freight flows." In: *Transportation Science* 24.1, pp. 25–39.

Halim, R. A. and M. D. Seck (2011). "the Simulation-based Multi-objective Evolutionary Optimization (SIMEON) framework. Conference on Winter Simulation 43th Annual Meeting, 2011 Phoenix, Arizona. Winter Simulation Conference." In:

Heinitz, F. M. and P. A. Meincke (2013). "Modeling Framework of Origin and Destination Air Cargo Routing." In: *Transportation Research Record* 2336, pp. 83–90.

Hoogendoorn-Lanser, S., R. van Nes, P. Bovy, and Trb (2005). "Path size modeling in multimodal route choice analysis." In: *Travel Demand 2005* 1921, pp. 27–34. ISSN: 0361-1981. URL: <GotoISI>://WOS:000236203700004.

IATA (2014). "IATA Cargo Analysis Archives." In: URL: http://www.iata.org/publications/economics/Pages/caarchives.aspx.

IATA (2015). "IATA Cargo Strategy 2015-2020." In: URL: https://www.iata.org/whatwedo/cargo/Documents/cargo-strategy.pdf.

ITF (2015). "ITF Transport Outlook 2015." In: URL: /content/book/9789282107782-en.

Innovata (2014). "Flight Schedules Database." In: URL: http://www.innovata-llc.com/timetables/.

Jong, G.C. de, M. Kouwenhoven, J. Bates, P. Koster, E. Verhoef, L. Tavasszy, and P. Warffemius (2014). "New SP-values of time and reliability for freight transport in the Netherlands." In: *Transportation Research Part E* 64, pp. 71–87.

Jourquin, B. (2005). "A multi-flow multi-modal assignment procedure applied to the European freight transportation networks." In: *Studies in Regional Science* 35, pp. 929–945.

Jourquin, B. and M. Beuthe (1996). "Transportation policy analysis with a geographic information system: The virtual network of freight transportation in Europe." In: *Transportation Research Part C-Emerging Technologies* 4.6, pp. 359–371.

Kuby, M. J. and R. G. Gray (1993). "The hub network design problem with stopovers and feeders: The case of Federal Express." In: *Transportation Research Part A* 27.1, pp. 1–33.

LeBlanc L.J.and Morlok, E.K. and W.P Pierskalla (1975). "An efficient approach to solving the road network equilibrium traffic assignment problem." In: *Transportation Research* 9.5, pp. 309–318.

LeBlanc, L. J., R. V. Helgason, and D. E. Boyce (1985). "Improved efficiency of the Frank-Wolfe algorithm for convex network programs." In: *Transportation Science* 19.4, pp. 445–462.

Maia, L. C. and A. F. do Couto (2013). "An innovative freight traffic assignment model for multimodal networks." In: *Computers in Industry* 64.2, pp. 121–127.

Miami International Airport (2014). "Cargo Hub." In: URL: www.miami-airport.com/pdfdoc/MIA_Cargo_Brochure.pdf.

Nagel, T., J. Klerkx, A. Vande Moere, and E. Duval (2013). "Unfolding – A Library for Interactive Maps. In: HOLZINGER, A., ZIEFLE, M., HITZ, M. & DEBEVC, M. (eds.)" In:

Newton, S. (2009). "WORLDNET Final Report." In: URL: http://www.transport-research.info/sites/default/files/project/documents/20130205_103358_77747_D11_WORLDNET_Final_Report.pdf.

Ohasi Hiroshi Tae-Seung, Kim Tae Hoon-Oum Chunyan Yu (2005). "Choice of Air Cargo Transshipment Airport: An Application to Air Cargo Traffic to/from Northeast Asia." In: *Journal of Air Transport Management.* 11.3, pp. 149–159.

Ortuzar, J. and L.G. Willumsen (2011). *Modelling Transport*. John Wiley & Sons, New York.

Patriksson, M. (1994). *The Traffic Assignment Problem - Models and Methods*. VSP, Utrecht, Netherlands.

Reed, P., J.B. Kollat, and V.K. Devireddy (2007). "Using interactive archives in evolutionary multiobjective optimization: A case study for long-term groundwater monitoring design." In: *Environ. Model. Softw* 22.1, pp. 683–692.

Seabury Consulting (2014a). “Air Cargo Capacity Database.” In: URL: http://www.seaburycargo.com/ibmcognos/cgi-bin/cognosisapi.dll.

Seabury Consulting (2014b). “Global Trade Database.” In: URL: http://www.seaburycargo.com/ibmcognos/cgi-bin/cognosisapi.dll.

Sheffi, Y. (1985). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall, Englewood Cliffs, NJ.

Southworth F. & Peterson, B.E. (2000). “Intermodal and international freight network modeling.” In: *Transportation Research Part C-Emerging Technologies* 8.1-6, pp. 147–166.

Sun Y. Lang, M. and D. Wang (2015). “Optimization models and solution algorithms for freight routing planning problem in the multi-modal transportation networks: A review of the state-of-the-art.” In: *TOCIEJ The Open Civil Engineering Journal* 9, pp. 714–723.

Tavasszy, L. A., M. Minderhoud, J.-F. Perrin, and T. Notteboom (2011). “A strategic network choice model for global container flows: specification, estimation and application.” In: *Journal of Transport Geography* 19, pp. 1163–1172.

Tavasszy, L.A. (2006). *Freight Modelling– An overview of international experiences*. Washington DC: TRB.

Thomas, R. (1991). *Traffic Assignment Techniques*. Avebury Technical. The Academic Publishing Group, Aldershot.

Transportation Statistics, US Bureau of (2014). “BTS Transtats Air Carriers : T-100 Segment (All Carriers).” In: URL: https://www.transtats.bts.gov/DL_SelectFields.asp?Table_ID=293&DB_Short_Name=Air%20Carriers.

Tsamboulas D. Moraitis, P. (2007). “Methodology for estimating freight volume shift in an international intermodal corridor.” In: *Transportation Research Record 2008*, pp. 10–18.

Wardrop, J.G. (1952). “Some theoretical aspects of road traffic research.” In: *Proc. Institution of Civil Engineers* Part II.1, pp. 325–378.

Wong, W., A.M Zhang, Y.V. Hui, and L.Leung (2009). “Optimal Baggage-Limit Policy: Airline Passenger and Cargo Allocation.” In: *Transportation Science* 43.3, pp. 355–369.

Zhang, A.M., Y.V. Hui, and Leung L. (2004). “Air cargo alliances and competition in passenger markets.” In: *Transportation Research Part E-Logistics and Transportation Review* 40.2, pp. 83–100.

Part I

Appendix

A

APPENDIX A: GLOBAL TRADE DATABASE

ASIA & OCEANIA	AMERICAS	EUROPE	
China	Brazil	Austria	Lithuania
Japan	Chile	Belgium	Luxembourg
Korea	Canada	Bulgaria	Malta
Hong Kong	Canada	Cyprus	Netherlands
Taiwan	USA	Czech Republic	Norway
Philippines	Colombia	Denmark	Poland
Indonesia		Estonia	Portugal
Malaysia		Finland	Romania
Singapore		France	Slovakia
New Zealand		Germany	Slovenia
Australia		Greece	Spain
		Hungary	Sweden
		Ireland	Switzerland
		Italy	United Kingdom
		Latvia	

Table 25: Direct sourcing countries for the Trade Database

ASIA	MESA	AMERICAS	AFRICA	EUROPE
Mongolia	Egypt	Argentina	Kenya	Russia
Thailand	India	Bolivia	South Africa	
Vietnam	Israel	Costa Rica		
	Pakistan	Cuba		
	Saudi Arabia	Ecuador		
	Turkey	Mexico		
	U.A.E.	Panama		
		Peru		
		Venezuela		

Table 26: Indirect sourcing countries for the Trade Database

APPENDIX B: AIRPORT IMPEDANCE

AIRPORT CODE	CITY	COUNTRY	A_p-VALUE
ABJ	Abidjan	Ivory Coast	458
ACC	Accra	Ghana	240
ADD	Addis Ababa	Ethiopia	61
ADL	Adelaide	Australia	2,148
AGT	Ciudad del Este	Paraguay	995
AKL	Auckland	New Zealand	336
ALA	Almaty	Kazakhstan	435
ALG	Algiers	Algeria	704
AMD	Ahmedabad	India	407
AMM	Amman	Jordan	1,766
AMS	Amsterdam	Netherlands	7
ANC	Anchorage	US North West	0
ARN	Stockholm	Sweden	739
ASB	Ashgabat	Turkmenistan	261
ASU	Asuncion	Paraguay	335
ATH	Athens	Greece	959
ATL	Atlanta	US South East	1,887
AUH	Abu Dhabi	United Arab Emirates	32
AUS	Austin	US South Central	9,407
BAH	Bahrain	Bahrain	180
BAQ	Barranquilla	Colombia	1,598
BCN	Barcelona	Spain	359
BEG	Belgrade	Serbia	13,481
BEY	Beirut	Lebanon	1,717
BFI	Seattle	US North West	4,968
BGI	Bridgetown	Barbados	441
BGY	Milan	Italy	131
BHX	Birmingham	United Kingdom	6,000

Table 27: Overview A_p-values per airport (1/10)

AIRPORT CODE	AIRPORT NAME	COUNTRY NAME	A_p-VALUE
BKK	Bangkok	Thailand	11
BKO	Bamako	Mali	1,051
BLQ	Bologna	Italy	240
BLR	Bengaluru	India	43
BNE	Brisbane	Australia	764
BOD	Bordeaux	France	7,000
BOG	Bogota	Colombia	50
BOJ	Burgas	Bulgaria	2,000
BOM	Mumbai	India	71
BOS	Boston	US North East	112
BRU	Brussels	Belgium	290
BSL	Basel/Mulhouse	Switzerland	597
BUD	Budapest	Hungary	58
BWN	Bandar Seri Begawan	Brunei Darussalam	27
BZV	Brazzaville	Rep. of the Congo	2,084
CAI	Cairo	Egypt	162
CAN	Guangzhou	China South	378
CCJ	Kozhikode	India	581
CCS	Caracas	Venezuela	43
CCU	Kolkata	India	1,874
CDG	Paris	France	10
CEB	Cebu	Philippines	1,760
CGK	Jakarta	Indonesia	132
CGN	Cologne/Bonn	Germany	41
CGO	Zhengzhou	China North	442
CHC	Christchurch	New Zealand	834
CKG	Chongqing	China West	9,244
CKY	Conakry	Guinea	541
CLO	Cali	Colombia	3,000
CLT	Charlotte	US South East	4,109
CMB	Colombo	Sri Lanka	153
CMN	Casablanca	Morocco	1,670
CNS	Cairns	Australia	7,546

Overview A_p-values per airport (2/10)

AIRPORT CODE	AIRPORT NAME	COUNTRY NAME	A_p-VALUE
COK	Kochi	India	219
COO	Cotonou	Benin	1,961
CPH	Copenhagen	Denmark	266
CPT	Cape Town	South Africa	576
CRK	Angeles/Mabalacat	Philippines	500
CTS	Sapporo	Japan	7,339
CTU	Chengdu	China West	890
CUN	Cancun	Mexico	9,927
CVG	Cincinnati	US North Central	797
CWB	Curitiba	Brazil	1,131
DAC	Dhaka	Bangladesh	139
DAR	Dar Es Salaam	Tanzania	421
DEL	Delhi	India	150
DEN	Denver	US North Central	1,410
DFW	Dallas	US South Central	1,393
DKR	Dakar	Senegal	232
DLA	Douala	Cameroon	82
DLC	Dalian	China North	5,244
DME	Moscow	Russia	661
DMK	Bangkok	Thailand	10,000
DMM	Dammam	Saudi Arabia	542
DOH	Doha	Qatar	59
DPS	Denpasar-Bali	Indonesia	1,045
DTW	Detroit	US North Central	1,583
DUB	Dublin	Irish Republic	360
DUR	Durban	South Africa	4,138
DUS	Duesseldorf	Germany	810
DWC	Dubai	United Arab Emirates	12
DXB	Dubai	United Arab Emirates	0
EBB	Entebbe	Uganda	185
EBL	Erbil	Iraq	107
EDI	Edinburgh	United Kingdom	6,235
EDL	Eldoret	Kenya	2,806

Overview A_p-values per airport (3/10)

AIRPORT CODE	AIRPORT NAME	COUNTRY NAME	A_p-VALUE
EMA	Nottingham	United Kingdom	109
ESB	Ankara	Turkey	9,677
EVN	Yerevan	Armenia	2,602
EWR	New York	US North East	5,887
EZE	Buenos Aires	Argentina	73
FCO	Rome	Italy	488
FDF	Fort-de-France	Martinique	0
FIH	Kinshasa	Dem. Rep. of the Congo (Zaire)	852
FOC	Fuzhou	China South	6,718
FRA	Frankfurt	Germany	10
FRU	Bishkek	Kyrgyzstan	964
FUK	Fukuoka	Japan	1,505
GDL	Guadalajara	Mexico	132
GEO	Georgetown	Guyana	903
GIG	Rio De Janeiro	Brazil	184
GLA	Glasgow	United Kingdom	1,862
GMP	Seoul	South Korea	6,743
GOT	Gothenburg	Sweden	1,960
GRU	Sao Paulo	Brazil	87
GUA	Guatemala City	Guatemala	919
GUM	Guam	Guam	15
GVA	Geneva	Switzerland	5,077
GYD	Baku	Azerbaijan	1,961
GYE	Guayaquil	Ecuador	819
HAM	Hamburg	Germany	391
HAN	Hanoi	Vietnam	999
HEL	Helsinki	Finland	334
HGH	Hangzhou	China East	19,596
HHN	Frankfurt	Germany	149
HKG	Hong Kong	Hong Kong	0
HKT	Phuket	Thailand	11,125
HND	Tokyo	Japan	1,379
HNL	Honolulu	US North West	105

Overview A_p-values per airport (4/10)

AIRPORT CODE	AIRPORT NAME	COUNTRY NAME	A_p-VALUE
HRE	Harare	Zimbabwe	1,113
HYD	Hyderabad	India	549
IAD	Washington	US North East	369
IAH	Houston	US South Central	799
ICN	Seoul	South Korea	72
IFN	Esfahan	Iran	183
IKA	Tehran	Iran	608
IND	Indianapolis	US North Central	808
ISB	Islamabad	Pakistan	3,569
IST	Istanbul	Turkey	305
JED	Jeddah	Saudi Arabia	194
JFK	New York	US North East	516
JNB	Johannesburg	South Africa	92
KAN	Kano	Nigeria	408
KBL	Kabul	Afghanistan	610
KBP	Kiev	Ukraine	754
KEF	Reykjavik	Iceland	1,622
KGL	Kigali	Rwanda	506
KHH	Kaohsiung	Taiwan	5,246
KHI	Karachi	Pakistan	226
KIN	Kingston	Jamaica	201
KIX	Osaka	Japan	689
KMG	Kunming	China West	5,209
KMQ	Komatsu	Japan	8,058
KRT	Khartoum	Sudan	322
KTM	Kathmandu	Nepal	3,277
KTW	Katowice	Poland	564
KUL	Kuala Lumpur	Malaysia	50
KWI	Kuwait	Kuwait	438
LAD	Luanda	Angola	645
LAS	Las Vegas	US South West	7,825
LAX	Los Angeles	US South West	25
LBU	Labuan	Malaysia	1,638

Overview A_p-values per airport (5/10)

airport code	airport name	country name	A_p-value
LBV	Libreville	Gabon	1,169
LCA	Larnaca	Cyprus	81
LCK	Columbus	US North Central	9,819
LED	St Petersburg	Russia	6,972
LEJ	Leipzig/Halle	Germany	48
LFW	Lome	Togo	379
LGG	Liege	Belgium	125
LGW	London	United Kingdom	2,238
LHE	Lahore	Pakistan	842
LHR	London	United Kingdom	10
LIM	Lima	Peru	264
LIS	Lisbon	Portugal	123
LOS	Lagos	Nigeria	172
LPA	Gran Canaria	Spain	1,136
LTN	London	United Kingdom	2,462
LUN	Lusaka	Zambia	3,489
LUX	Luxembourg	Luxembourg	178
LYS	Lyon	France	2,077
MAA	Chennai	India	55
MAD	Madrid	Spain	632
MAN	Manchester	United Kingdom	2,198
MAO	Manaus	Brazil	85
MBJ	Montego Bay	Jamaica	219
MCO	Orlando	US South East	8,134
MCT	Muscat	Oman	307
MDE	Medellin	Colombia	1,184
MED	Madinah	Saudi Arabia	3,210
MEL	Melbourne	Australia	155
MEX	Mexico City	Mexico	20
MFM	Macau	Macau	2,004
MGA	Managua	Nicaragua	171
MHD	Mashhad	Iran	70
MIA	Miami	US South East	258

Overview A_p-values per airport (6/10)

AIRPORT CODE	AIRPORT NAME	COUNTRY NAME	A_p-VALUE
MID	Merida	Mexico	1,500
MLA	Malta	Malta	212
MLE	Male	Maldives	146
MMX	Malmo	Sweden	599
MNL	Manila	Philippines	159
MPM	Maputo	Mozambique	7,705
MRS	Marseille	France	652
MRU	Mauritius	Mauritius	1,245
MSE	Manston-Kent	United Kingdom	8,096
MSP	Minneapolis/Saint Paul	US North Central	3,153
MST	Maastricht	Netherlands	447
MTY	Monterrey	Mexico	2,000
MUC	Munich	Germany	510
MVD	Montevideo	Uruguay	2,071
MXP	Milan	Italy	187
NBO	Nairobi	Kenya	100
NCE	Nice	France	5,851
NDJ	N'Djamena	Chad	304
NGO	Nagoya	Japan	3,324
NKG	Nanjing	China East	10,844
NRT	Tokyo	Japan	7
NTE	Nantes	France	5,998
OAK	Oakland	US South West	2,006
OKA	Okinawa	Japan	67
ONT	Ontario	US South West	274
OPO	Porto	Portugal	39
ORD	Chicago	US North Central	551
ORY	Paris	France	281
OSL	Oslo	Norway	344
OTP	Bucharest	Romania	181
OUA	Ouagadougou	Burkina Faso	1,287
OVB	Novosibirsk	Russia	1,069
PBM	Paramaribo	Surinam	211

Overview A_p-values per airport (7/10)

AIRPORT CODE	AIRPORT NAME	COUNTRY NAME	A_p-VALUE
PDX	Portland	US North West	5,790
PEK	Beijing	China North	1,643
PEN	Penang	Malaysia	1,418
PER	Perth	Australia	68
PEW	Peshawar	Pakistan	11,253
PHC	Port Harcourt	Nigeria	1,117
PHL	Philadelphia	US North East	762
PIK	Glasgow	United Kingdom	3,687
PNH	Phnom Penh	Cambodia	342
PNR	Pointe-Noire	Rep. of the Congo	133
POA	Porto Alegre	Brazil	9,983
POM	Port Moresby	Papua New Guinea	404
POS	Port Of Spain	Trinidad & Tobago	43
PPT	Tahiti	French Polynesia	71
PRG	Prague	Czech Republic	520
PTP	Pointe A Pitre	Guadeloupe	1
PTY	Panama City	Panama	292
PUJ	Punta Cana	Dominican Republic	2,187
PUS	Busan	South Korea	7,935
PVG	Shanghai	China East	41
RGN	Yangon	Myanmar	1,422
RIX	Riga	Latvia	1,012
RUH	Riyadh	Saudi Arabia	5
RUN	Saint Denis de la Reunion	Reunion	11
SAH	Sana'a	Yemen	210
SAL	San Salvador	El Salvador	128
SAN	San Diego	US South West	9,546
SAP	San Pedro Sula	Honduras	1,238
SAW	Istanbul	Turkey	1,539
SCL	Santiago	Chile	111
SDF	Louisville	US North Central	257
SDQ	Santo Domingo	Dominican Republic	19
SEA	Seattle	US North West	310

Overview A_p-values per airport (8/10)

AIRPORT CODE	AIRPORT NAME	COUNTRY NAME	A_p-VALUE
SEZ	Mahe Island	Seychelles	71
SFO	San Francisco	US South West	1,629
SGN	Ho Chi Minh City	Vietnam	364
SHA	Shanghai	China East	7,793
SHE	Shenyang	China North	2,621
SHJ	Sharjah	United Arab Emirates	106
SIN	Singapore	Singapore	0
SJC	San Jose	US South West	4,369
SJO	San Jose	Costa Rica	661
SNN	Shannon	Irish Republic	2,091
SOF	Sofia	Bulgaria	959
SQQ	Siauliai	Lithuania	875
SSA	Salvador	Brazil	5,828
SSG	Malabo	Equatorial Guinea	714
STI	Santiago	Dominican Republic	54
STN	London	United Kingdom	1,937
STR	Stuttgart	Germany	6,005
SUB	Surabaya	Indonesia	197
SVO	Moscow	Russia	1,099
SVX	Yekaterinburg	Russia	651
SYD	Sydney	Australia	831
SYZ	Shiraz	Iran	760
SZB	Kuala Lumpur	Malaysia	9,941
SZX	Shenzhen	China South	221
TAO	Qingdao	China North	590
TAS	Tashkent	Uzbekistan	2,033
TBS	Tbilisi	Georgia	1,926
THR	Tehran	Iran	728
TLL	Tallinn	Estonia	317
TLS	Toulouse	France	356
TLV	Tel Aviv-Yafo	Israel	594
TNR	Antananarivo	Madagascar	957
TPA	Tampa	US South East	6,385

Overview A_p-values per airport (9/10)

AIRPORT CODE	AIRPORT NAME	COUNTRY NAME	A_p-VALUE
TPE	Taipei	Taiwan	82
TRV	Thiruvananthapuram	India	64
TSA	Taipei	Taiwan	285
TSN	Tianjin	China North	938
TUN	Tunis	Tunisia	2,072
TXL	Berlin	Germany	1,231
TZX	Trabzon	Turkey	247
UIO	Quito	Ecuador	131
URC	Urumqi	China West	472
VBS	Verona	Italy	5,999
VCE	Venice	Italy	1,172
VCP	Sao Paulo	Brazil	117
VIE	Vienna	Austria	726
VIT	Vitoria	Spain	350
VKO	Moscow	Russia	8,295
VVI	Santa Cruz	Bolivia	1,336
WAW	Warsaw	Poland	209
WDH	Windhoek	Namibia	207
WUH	Wuhan	China North	6,324
XIY	Xi'an	China West	4,358
XMN	Xiamen	China South	2,535
YVR	Vancouver	Canada	587
YYC	Calgary	Canada	393
YYZ	Toronto	Canada	1,691
ZAZ	Zaragoza	Spain	83
ZRH	Zurich	Switzerland	580

Overview A_p-values per airport (10/10)