FORECASTING CONTAINER CARGO THROUGHPUT IN PORTS

Master thesis Mathilde Jansen (386267)
Erasmus University Rotterdam
Erasmus School of Economics
Urban, Port and Transport Economics
Supervisor Dr. M. Nijdam
September 2014

Preface

This report is my final effort in order to graduate the Master Urban, Port & Transport Economics at the Erasmus University Rotterdam. It is written commissioned by Witteveen+Bos to give guidelines for their container cargo forecasting practices.

I would like to thank my supervisor at the EUR, Michiel Nijdam, for his guidance and feedback. Also, I would like to thank my supervisors at Witteveen+Bos: Bert Burgers for his detailed feedback and Ben-Jaap Pielage for making sure I would not lose sight of the bigger picture. Acknowledgements go out to Lóránt Tavasszy (TNO and TU Delft) and Aernoud Willeumier (Port of Rotterdam) who managed to make time to provide me with useful information.

Mathilde Jansen

Rotterdam, September 2014

Abstract

At the time of writing this thesis, no universal method for making forecasts on container cargo throughput in seaports is conducted by consultancy firms and other interested parties. This report gives an overview of the different models used for freight modelling. These models are based on the four step approach that is used in modelling demand for passenger transport: production and attraction, distribution, modal split and assignment to the transport network. Based on this overview, essential elements of a good forecasting method are identified. These are a model to determine the generation and distribution of freight flows and a model to explain port choice and competition. The input and output of these models should be evaluated using expert and commodity specific knowledge, resulting in possible modification of the results. Case studies on different port categories - i.e. transhipment, import-export, Greenfield, Brownfield, well-developed and less-developed economies - are used to illustrate the factors that explain the fluctuations of container cargo throughput volumes in practice. These findings have possible implications for the optimal set of variables and/or port choice model used. The purpose of this thesis is to provide a solid basis to develop an integrated and universal forecasting method.

Table of contents

1 Introduction

In the past decades, globalization caused a huge increase in trade volumes. Because of this increase, new port areas had to be developed. This process is still going on, as can be seen from the development of Maasvlakte II at the Port of Rotterdam for example. Investment decisions on these kind of expansion projects are usually based on port traffic and cargo throughput forecasts. These forecasts are made for at least 25 years ahead, since ports generally have a long technical and economic lifetime, and investments made in port infrastructure have a long pay-back period (De Langen, Van Meijeren & Tavasszy, 2012a). Until now, no universal method is developed to do these forecasts. The main approach used for existing ports is based on trend extrapolation techniques that rely on historic data. However, such models give little insight in the variables that drive cargo throughput in ports other than economic growth. For Greenfield ports it is noted that trend extrapolation is not possible, thus requiring an alternative approach.

The purpose of this thesis is understanding the variables that affect container cargo throughput. The analysis is made for this specific type of cargo, because in contrast to bulk cargo containers can be used for a wide range of consumer, construction and industry related commodities and clients. This makes its development subject to trends and general economic development. Mass containerization in the past decades, for example, caused a major growth in container throughput in ports. Therefore, many port expansion or development projects are about developing (additional) container terminals (De Langen, Nijdam & Van der Lugt, 2012b). It is assumed that in the future, more and more commodities are being containerized (Havenga & Van Eeden, 2011) causing the container shipping industry to expand even more. Havenga & Van Eeden (2011) also predict a maturing in the containerization trend, simply because on a given point in time every commodity that can be shipped in containers shall be shipped in containers.

The focus is not solely on the variables. Different modelling techniques are evaluated to prepare a general pragmatic approach for container cargo throughput forecasting in ports. These methods are generally based on the four step approach that is also used in passenger transport modelling: production & attraction, distribution, modal split and route choice. The evaluation of these techniques serve as the basis for identifying the essential elements of a good container cargo forecasting approach.

To verify whether the drivers mentioned in the literature are indeed the factors that determine container throughput volumes, the developments of throughput volumes in several ports are studied. Drivers of growth and decline are identified for two opposites in three different port categories, namely transhipment vs. import-export, Greenfield vs. Brownfield and less-developed vs. well-developed economy. As a result, a recommendation is given on the model used for each port.

The thesis is structured as follows. In chapter 2, an overview is given on the literature on freight transport modelling. Variables and modelling techniques are described and serve as a basis for chapter 3. Herein the requirements that a good forecasting should meet are identified and the mathematical relations between variables in models that are assumed suitable are described. In chapter 4, the development of port throughput in the past years are described for different port categories and disruptions in the trends will be explained. An assessment is made on how the modelling approach should differ per port category. The thesis concludes with an overall conclusion, an advice for Witteveen+Bos, the company that requested this research, and recommendations for further research, as this thesis should serve as a basis for the development of forecasting models.

2 Freight transport modelling

The aim of this chapter is to give an overview of the methods that are now available on freight forecasting. The overall accepted method that is derived from the passenger transport modelling literature is described. Models used in practice and the variables and data that are required for these models are explained.

In the literature, little attention is given to freight transport compared to passenger transport modelling. Most of the freight transport models that are developed nowadays are based on the four step approach of passenger transport models. However, some important differences have to be kept in mind, like the diversity of decision makers and items being transported as well as the limited availability of data in freight transport forecasting. According to De Jong, Gunn and Walker (2004) the four steps in a freight transport model are (1) production and attraction, (2) distribution, (3) modal split and (4) assignment. In the production and attraction step, quantities of goods that are produced and demanded in certain regions are determined. It is considered that in the origin, the goods are produced and that those goods are attracted by the destination. Therefore the step is called production & attraction. In the second step, the quantities defined in the first step are translated into transport flows between the origin and destination zones. After that, in the third step, these transport flows of goods are allocated to transport modes. In the last and fourth step, these flows per mode are assigned to the transport network (infrastructure). It is also possible to combine the last two steps and assign the transport flows to modes and the network directly. This four step approach is broadly accepted by researchers. However, Tavasszy (2006) argues that a fifth, logistic layer is required, situated between step 2 and 3, because consumer preferences are changing causing changes in freight demand and logistics. In this step, the locations of distribution centres are determined to complete the network (L. Tavasszy, personal communication, July 11, 2014). In the next subsections the steps are discussed in more detail. Subsequently, models and methods that can be used to complete each step are described.

2.1 Models for production and attraction

Models for production and attraction illustrate the quantities of goods that are transported between two regions. These goods start at the production zone and are transported to the attraction zone. Zones can be countries, a group of countries or a region within a country. This is up to the researcher to decide. Attraction exists when there is demand for a certain product that is produced in another region. This causes a flow of a goods from zone to zone.

2.1.1 Time series models

There are four models for production and attraction that are used in practice. The first are trend and time series models, wherein historical trends are extrapolated into the future. Time series models exists in a wide range, from simple growth factor models to complex Autoregressive Moving Average models (ARMA). However, the latter model is mainly suitable for short-term forecasting (De Jong et al., 2004). Therefore, time series models with an explanatory variable, i.e. a transport multiplier, are used more often. This is not the same as a regression analysis with explanatory variables. Because here a trend of a particular explanatory variable is described, while in the latter a certain value is forecasted and used in a regression model (M. Nijdam, personal communication, May 27, 2014).

A transport multiplier can be based on a lot of different data, like growth in sales, gross domestic product (GDP) or stock market analysis (Gosasang, Chandraprakaikul & Kiattisin, 2010). GDP is mostly used as transport multiplier because there is an easily demonstrable relationship between GDP and port throughput, as the latter is a function of import and export, which are both a function of GDP (Van Dorsser, Wolters & Van Wee, 2011). In case GDP is applied, the transport multiplier reflects the relationship between GDP growth and transport growth. For example, between 1991 and 1996 the multiplier for transport of containerized goods was about 3. This means container cargo transport grew 3 times faster than GDP. Eventually, the multiplier has to move in the direction of 1 because port throughput in relation to GDP growth cannot continue to grow forever. The high multiplier for containerized goods in the 90s is a result of the mass containerization in the past decades (De Langen et al., 2012b), also the increasing shipment of intermediate goods resulted in increasing container cargo. For example, cars are not produced in one location, but the parts are produced in various different countries and shipped to the place of assembly.

It is important to motivate the value(s) used for GDP in the forecasts. GDP forecasts are usually based on neo-classical growth theory. This theory states that economic growth is a product of three main drivers: labour force growth, capital accumulation and technological progress (or factor productivity growth). In general, one works with forecast of IMF (International Monetary Fund) for GDP growth. However, this forecast is only for 8 years ahead and therefore the trend in GDP growth has to be extrapolated after this eight years period (UNESCAP, 2007), to obtain a forecasts for the necessary 25 years ahead in case of container cargo forecasting. The GDP forecasts from the international financial authority IMF are not easily disputed, which makes this forecasts attractive to use for container cargo forecasting. So, the issue here is what percentage for GDP growth forecasts should be used, and even more practical, what percentage GDP growth forecast should be used for the period IMF does not cover. Unfortunately, scientific literature presents little information on how to deal with this issue.

Although time series models have the advantage that they require only a few data, there are also some drawbacks (De Jong et al., 2004). Because the forecast is based on only one variable (or a historical trend), no insight is gained in other events or variables that can affect the throughput volumes in a particular port. Like policy effects for example. These insights can be of great value to understand the trend that is forecasted (De Langen et al., 2012a). In contrast with this, the literature also suggests that in making forecasts it can be dangerous to put in many variables containing more details to make the model "better", a failure called *false reduction.* This is because the more variables included in a forecasting model, the more uncertainty there is about the validity of the model (Bankes, 1993). This is particularly the case in cause-and-effect models that will be discussed later.

Also, in the example of GDP as variable a stable relationship between GDP and growth of port throughput is assumed. Forecast models that rely only on trend forecasts do not account for shocks in the economy, like the oil crisis and the mass containerization that made the model over- and underestimate the throughput, respectively. As can be seen in these examples, the fact that the model does not capture disruptions in the future, especially affects throughput of specific types of cargo (oil and containers in the examples). Therefore, it is important to establish models that suit one type of cargo instead of all freight throughput in a port (De Langen et al., 2012a). Additionally, it is important to note that the beginning of the containerization trend, i.e. the invention of the container, could not be included in a model. However, now that containerization is going on for a while, it should be accounted for in forecasts of container throughput in ports. It is expected that first the effect was very strong, but after some years the containerization rate will lose effect. This is because eventually, everything that could be shipped in containers shall be shipped in containers. However, it is important to note that containers are, like almost every product, subject to innovations. For example, new climate control systems allow us to ship flowers in containers overseas or transport bananas on container vessels instead of reefer carriers. The issue with innovations is that, most of the times, they cannot be predicted and so their effects are also a surprise (A.E. Willeumier, personal communication, June 25, 2014). Of course, a qualitative analysis can be made on possible developments and their effects in the future. Havenga and Van Eeden (2011) make an effort to forecast the containerization rates of different commodities in 2040. They include this containerization trends in a model to forecast container cargo traffic in South-Africa. The expectation is that predominantly commodities like fruit, vegetables and processed chemicals will be containerized by 100% in the future. Commodities like wood, iron and steel will not reach further containerization because of weight complexities. These expectations are based on containerization rates in 2009 and industry expert's judgement (Havenga & Van Eeden, 2011). Lastly, historical trend models are not suitable for long term forecasts, which is necessary for port development. Even if a time series model is combined with a causal relationship, like the relation between GDP and cargo growth, the forecast that comes out is not reliable in the long term (Van Dorsser et al., 2011). This is because in the long term structural economic changes and fluctuations in transport flows are inevitable and mere trend extrapolation does not account for these disruptions. Therefore, forecasts made by time series model should serve as a starting point for making a forecast.

2.1.2 System dynamics models

Second, there are system dynamics models wherefore parameters are obtained from existing literature and trial and error, instead of being estimated statistically. System dynamic models are basically schemes that illustrate relationships between different variables using arrows. These arrows stand for so called feedback from one variable on the other. Feedbacks can be negative, i.e. making sure the variable moves to its long-term equilibrium, and positive, i.e. accelerate growth. It models variables that both affect *each other*. So not just in one direction, but back and forth. Therefore the relations cannot be tested one by one but have to be tested as the system as a whole. The input of the model can be either quantitative or qualitative data. This makes the model both better and worse, because it is more valuable if managerial experience, intuition and knowledge can be included in the model. This is especially the case for forecasting models, because external shocks can be accounted for in system dynamics models in contrast to econometric (time series) models. But at the same time, this feature makes the model less trustworthy, because it cannot be tested statistically. So, although this method is useful to map causalities between variables, the results cannot be tested statistically and thus econometric models are still needed to verify the significance of the variables using historical data (Dikos, Marcus, Papadatos & Papakonstantinou, 2006).

2.1.3 Regression analysis on cross-sectional data

Another way to explain the flows of cargo from origin to destination is by use of regression analyses on cross-sectional data. De Jong et al. (2004) mention zonal trip rate models. Cross-sectional data on transport volumes from/to each zone under consideration is classified. The output is then a number of homogenous zone types from which zonal trip rates are derived. However, until now this method is only applied to urban regions and thus road transport only (De Jong et al., 2004). This makes the model not very usable in this specific case of forecasting container cargo throughput in ports.

In other scientific literature different methods are proposed under the more general heading "cause-and-effect models". The purpose of cause-and-effect models is to determine a relationship between several variables and cargo throughput in ports. So this method is more specifically related to a particular port, that can be seen as a zone. An example of a cause-and-effect model that is used frequently is regression analysis. A regression analysis helps to describe data, estimate parameters and verify relations that arise from economic logic. A regression model can be linear or non-linear. Previous research has shown that a non-linear regression is the preferable regression model for forecasting cargo throughput (Gosasang et al., 2010).

There are several variables used in cause and effect models. The most popular one is GDP, as also mentioned above. Many variables are suitable for regression analyses. However, they can be divided into two broad categories for now: variables that explain trade generation and variables that explain trade distribution. The first category of variables is of importance for the first step of production and attraction. The main rationale behind this category is that the demand for port services is derived from demand for import, export and transhipment. The variables in this category describe macroeconomic conditions of a country of interest and general demographic characteristics like the size of the population. Below, a list is given of several macroeconomic variables that are mentioned in the literature. Also, the rationale behind them is pointed out to clarify their relation with container cargo throughput.

- Import prices (Coto-Millán, Baños-Pino & Castro, 2005), the higher the import prices, the lower the transport volume.
- Trade value (Seabrooke et al., 2003), because demand for port services is an outcome of demand for imports and exports.
- Population (Seabrooke at al., 2003; Gosasang et al., 2010), the larger the population, the higher demand, which causes increasing trade flows.
- Exchange rates (Gosasang et al., 2010), unfavourable exchange rates discourage trade which will make the trade flows decline.
- Interest rates (Gosasang et al., 2010), interest rates partly determine the value of GDP and GDP is a good indicator of port throughput because this is determined by imports and exports, which are a function of GDP.
- Inflation rates (Gosasang et al., 2010), inflation rates also partly determine the value of GDP.
- Trade between biggest trade partner of the country and the country itself (Hui, Seabrooke & Wong, 2004), this determinant is derived from the variable 'trade value', that is mentioned

above. The trade between two biggest trade partners can be used as a representative for the complete regional trade.

Next to the above described variables that can be expressed quantitatively, there are also macroeconomic parameters, i.e. macroeconomic events, that affect port cargo throughput but whereof no historical data exists. These factors cannot be incorporated in the model and have to be treated differently. Seabrooke et al. (2003) treat specific macroeconomic events like liberalization of trade between China and Taiwan (relevant for their specific case) as external shocks. A qualitative assessment on the effects of these events is given in the paper. Qualitative assessment as a supplement of the forecasting technique is treated more extensively later in this chapter.

The above mentioned possible variables are input to several (mathematically) different cause-and-effect models. Below, three ways in which data is used to come to a certain forecast are described.

Classical regression

The classical regression model is commonly used in practice. By measuring the co-movement of variables, this model detects causal relationships. Applying a classical regression model for making forecasts can be a problem, because the outcomes are only valid if the variables are stationary, i.e. their value does not depend on the point in time at which it is measured. Variables that follow the same trend over time are indicated as closely related, while this is not necessarily true. For example, when you treat nonstationary variables like US export rates and the life expectancy of Australian males as stationary and perform a classical regression, the results show that the variables are highly correlated. This is reflected in a high fit (R^2) and high F- and t-values. However, when thinking logically this relation does not make any sense. One can also perform a Dickey-Fuller (ADF) test to test whether the variables are stationary. This event, where unrelated variables are estimated to have a causal relationship, is called *spurious regression* (Hui et al., 2004).

Error correction model

One way to deal with nonstationary data is estimating the first difference of variables and structure a model with that data. In that case, not the original values are included in the model but the changes in value between two periods. A major drawback of this is that the model only considers the short-run adjustments related to how the difference in one variable correlates with the changes in the other. Hence, it ignores the long-term relationship between variables (Hui et al., 2004).

To avoid this problem, the error correction model (ECM) can be used. This is a differenced model that contains an error correction term, that predicts short-term adjustments of the dependent variable. The underlying idea of ECM is that a possible disequilibrium in the short run corrects itself over time, creating a path that fluctuates around the long-run equilibrium (Hui et al., 2004). So ECM is only valid if variables have a true relationship in the long-run (Van Dorsser et al., 2011). A co-integration test can be used to test whether such a relationship exists (Hui et al., 2004).

Neural networks

Another type of cause-and-effect models is the neural network model. The methodology of neural networks (NNs) is based on the structure of the human brain (Lam et al., 2004). This is reflected in the fact that NNs can learn from environment and, in that way, improve their performance (Gosasang et al., 2010). The model consists of several components, that are illustrated in the figure below. There are three layers: the input layer, the hidden layer (optional) and the output layer. Each layer is connected with the other by so called neurons. These are simple processing elements that are connected by weights. Each neuron receives weighted input from other neurons and then sends its output to neurons in another layer (Lam et al., 2004).

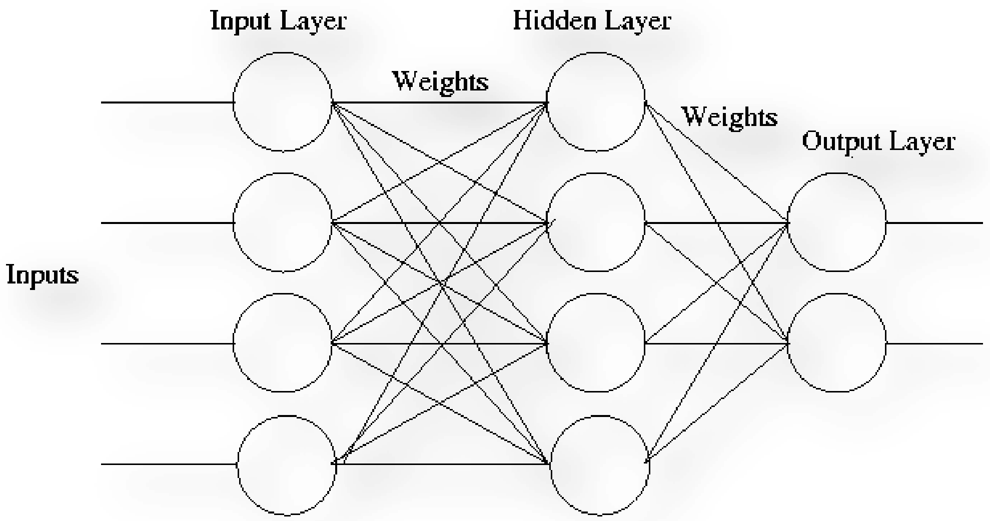

Figure 1: Neural network model. Source: wordassocation1.net.

Research on NNs shows that these models perform better than regression analyses. NNs seem to be able to follow volatilities better than regression analyses. The latter can only follow the middle of the trend. For forecasting cargo throughput it is better to use NNs because sometimes the volumes depend on explanatory variables that are very volatile over time. In short, neural network models give more trustworthy results when forecasting cargo throughput, than regression analyses (Lam et al., 2004). However, although in terms of results a neural network model outperforms other cause-and-effect models, it is difficult to use in practice, because it is a very complicated mathematical model for which specified knowledge is required. Also, a lot of data are required which are generally hardly available for ports.

Type of model	Advantages	Disadvantages
Classical regression	Easy to use in practice	Does not capture the real causalities because variables are nonstationary
Differenced regression	Deals with nonstationary variables	Only considers short term adjustments
Error correction model	Deals with nonstationary variables	Co-integration between variables is required, otherwise the model is useless
Neural networks	Model is able to learn from environment and thus improve its performance More trustworthy results than the models above, because it can follow volatilities better	A lot of data is needed as input Difficult to use in practice

Table 1: Advantages and disadvantages of the different cause and effect models. Based on Hui et al. (2004), Gosasang et al. (2010), Van Dorsser et al. (2011) & Lam et al. (2004).

2.1.4 I/O models

Finally, I/O models are described. These are macroeconomic models based on input/output tables. These are tables that display what each sector of the economy delivers to other sectors, expressed in money units. Final demand by consumers, import and export are also included in these tables. Next to these multisectoral I/O tables, multiregional ones exist. These are better suited to freight forecasting and do this in two ways. One is by use of fixed technical and trade coefficients, which means the level of production and the pattern of trade are extrapolated into the future. A second way to use a multiregional I/O table for freight transport forecasting is to use elastic technical and trade coefficients. Usually a multinomial logit model is estimated in which the fraction that is consumed in a specific region A of the production in sector S in region B depends on the total production in sector S in region B and the generalized transport costs. So, in this way distribution and production are elastic to changes in transport costs and time. It is also possible to use a multisectoral I/O model for the entire country and regionalize it afterwards, as in the Dutch TEM-II and Swedish SAMGODS model. It is important to keep in mind that the input of an I/O model is in money units and has to be transformed into tonnes before one can proceed with the next step (De Jong et al., 2004). Although this method is clear and transparent, it also takes a lot of time and effort, because import and export flows have to be identified for a lot of regions and after that I/O tables have to be developed.

Type of model	Advantages	Disadvantages	Variables
Time series	Limited data requirements (but for many years)	Little insight into causality and limited scope for policy effects Not perfectly suitable for long term forecasts	None Mostly GDP
System dynamics	Limited data requirements Can give land-use	No statistical tests on parameter values	Any variable that could impact container cargo

	interactions External and policy effects variables can be included		volumes and the other variables, like GDP, exchange rates, population, competition etc.
Zonal trip rate model	Limited data requirements	Little insight into causality and limited scope for policy effects	Transport volumes from/to each zone
Regression analysis	Designed to illustrate causal effects	Difficult to find the proper functional form	Macroeconomic variables
Input-output	Link to the economy Can give land-use interactions Policy effects in case of elastic coefficients	Need I/O-table, preferably multiregional Restrictive assumptions in case of fixed coefficients Need conversion from values to tonnes Need to identify import and export trade flows	The amount (in money units) that a sector delivers to other sectors Final demand Imports Exports

Table 2: Freight transport models for production and attraction. Based on De Jong et al. (2004).

2.2 Models for distribution

The distribution of freight transport is based on the output of the "production and attraction" step (the quantities of trade) and a measure of transport resistance, expressed as generalized transport costs.

2.2.1 Gravity model

Usually, a gravity model is used for this second step. This is also the case in the European TRANSTOOLS model. A gravity model is based on the gravity theory in physics. It is assumed that there is an attraction between two objects (countries) and the greater the mass (economy) of the objects, the greater the attraction. Mutual distance (transport costs) causes resistance of trade (Polder, 2000). So the freight movement between two regions is a function of the outcomes in the production & attraction step for both regions separately, divided by the generalized transport cost. It is useful to consider these generalized transport cost, and in particular the proportion of it, because for low value products transport costs make up a larger part of total cost than for high value products (Ortúzar & Willumsen, 1994). In general, goods that are transported in containers are high value products (Coto-Millán et al., 2005), so the transport cost make up a small part of the total costs related to the product. The gravity model is illustrated in figure 2, to give a better view on the variables and relation between them.

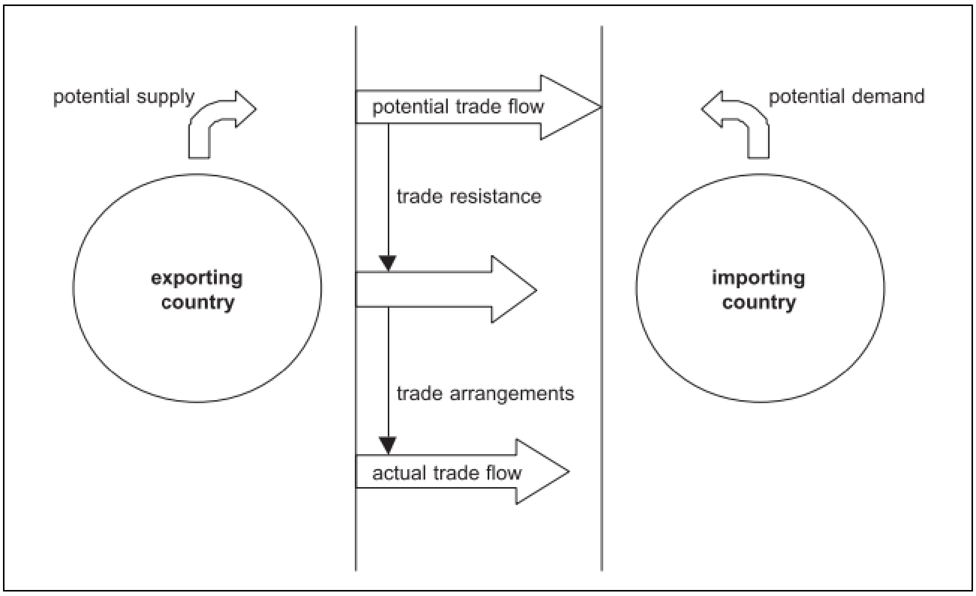

Figure 2: Graphical representation of the gravity model. Source: Polder (2000).

The gravity model is a clear model that is easy to understand and use. Data requirements are limited, which is positive. However, policy effects and other explanatory variables than trade resistance and trade arrangements cannot easily be included in the model. These variables can play a role in the next step that covers port choice. Overall, the model is considered useful in forecasting freight transport.

2.2.2 Multiregional I/O model

It is also possible to use the multiregional I/O model for both step one and two (De Jong et al., 2004). A more extensive discussion on I/O models can be found in the previous section.

Comparing the two methods, it is clear that the gravity model is easier to use but I/O models can account for policy effects in case of a multinomial logit model with elastic coefficients. The most software packages, however, include a gravity model for the distribution of freight flows. This shows that this model is strongly preferred by researchers.

Type of model	Advantages	Disadvantages	Variables
Gravity	Limited data requirements Some policy effects through transport cost function	Limited scope for including explanatory variables and policy effects Limited number of calibration parameters	Potential supply Potential demand Trade resistance Trade arrangements
Input-output (I/O)	Link to the economy Can give land-use interactions Policy effects in case of elastic coefficients	I/O table required, preferably multiregional Restrictive assumptions if fixed coefficients Need conversion from values to tonnes	The amount (in money units) that a sector delivers to other sectors Final demand Imports Exports

Table 3: Freight transport models for distribution. Based on De Jong et al. (2004).

2.3 Models for modal split

In the third step, the flows of goods from the two previous steps are distributed to different modes. Modal split concerns the distribution of cargo between transport modes, i.e. marine, road, rail and inland waterway (IWT). There are several kinds of models that deal with modal split. De Jong et al. (2004) mention seven models in their paper, those are all briefly described below.

The first are *elasticity models*. These models use elasticities that are adopted from other models or expert judgement. The elasticities show the effects of changing a single variable and are mainly used when very little data are available, or for a quick first analysis.

Aggregate modal split models use data on shares of a mode in a specific zone and put these in a binomial or multinomial logit model. So the output is a share instead of an absolute number and thus the elasticities in the model are conditional on the quantity demanded. Although little data is required as input for this model, it has a weak theoretical basis and there is little insight into causality.

Next, *neoclassical models* are based on microeconomic theory. The share of a particular mode in the total cost is usually the explanatory variable in these kind of models. So the share in total costs instead of the share in transport volume, which is the relevant variable in the greater four step system, is important in neoclassical models. This difference makes it hard to incorporate neoclassical models' output in the greater transport model system.

Models that are also hard to incorporate in the four step system are *direct demand models*. This is because transport routes for a specific mode are predicted directly by these models, while in the greater system the total demand over all modes is used.

Disaggregate modal split models use data from surveys conducted under shippers or commodity surveys. These data are used in a multinomial or nested logit model. The approach is based on passenger transport models: the utility functions of such models are adjusted to a profit function that is suitable for freight transport. Disaggregate modal split models have more advantages than drawbacks. It is possible to include many causal variables in the model and there is a solid theoretical basis: microeconomic theory. However, disaggregated data is required as input, which means surveys have to be conducted. This is a time consuming way of data collection.

A *microsimulation model* concerns trips for freight transported by trucks. It is actually a part of a two stage model of which the first part determines flows between regions and the second part, a microsimulation model, focuses on urban vehicle trip patterns. This makes this kind of model not interesting for the purpose of this paper.

The most extensively discussed models by De Jong et al. (2004) are *multimodal network models*. These models predict mode and route choice at the same time. An optimal combination of modes for a particular route is found using a cost minimization algorithm. It is important to note that the unit of observation is an origin-destination combination and not a specific firm, so the scale is more global. In software packages (SMILE, SCENES) the multimodal network is designed based on the endogenously specified locations of distribution centres. A huge drawback of multimodal network models is that they give little insight into causality. However, the model has a good theoretical basis and limited data requirements.

The World Container Model

An example of a multimodal network model is the World Container Model (WCM). Based on the assumption that attractive ports in terms of fuel cost, handling cost, congestion cost etc., get more container cargo throughput, Tavasszy et al. (2011) developed the WCM. This is a discrete choice model that predicts the container flows in the world and the effect of different scenarios on the distribution of these flows. The authors state that the distribution of container flows among the world's seaport network depends on the geographical location of the origin and destination of the containers, the cost of overseas and inland transport, the forwarding organisation's preferences regarding merchant or carrier haulage, the logistics characteristics of the goods in terms of unit value for example, the available facilities and services (for example infrastructure), and the decision agent's characteristics and strategies (Tavasszy et al., 2011). The assumption made in the model is that route choices are made with a profit maximizing objective. As one can see, the underlying focus is more on the micro-economic than on the macro-economic level.

The Port Competition Model

Another model that can be compared to the WCM, is the Port Competition Model developed by Veldman and Bückmann (2003). This is also a multinomial logit model that deals with port choice, and thus competition. Demand choice models for modal choice and routing choice are used. The probability of choosing a specific routing, defined by the choice for a port and an inland transport mode, is dependent on the utility of this routing. The utility of the routing is derived from the shipping costs, the transit time, the frequency of service and a dummy variable illustrating the preference of shippers/receivers for a specific routing.

Comparing the models it becomes clear that there is not one model that stands out immediately. The neoclassical and direct demand models are dismissed because they are hard to incorporate into the four step approach. The microsimulation approach is focussed on urban freight transport, while the focus of this study is on transport between ports. Because of its weak theoretical basis and its little insight into causality and policy effects, the aggregate modal split model is also not the most preferable model. Hence, a choice has to be made between the elasticity based approach, the disaggregate modal split and the multimodal network model. The latter two are more thorough and have a solid theoretical basis, while the first is better suitable to draw a quick picture. This makes the disaggregate modal split and the multimodal network model the best options for this step. It depends on the preference of the researcher and the availability of data - as the first needs a big and the second needs a small amount of data - which model should be used.

Type of model	Advantages	Disadvantages	Variables
Elasticity based	Very limited data requirements Fast in application	Elasticities may not be transferable Only impact of single variables, no interactions between variables	Freight transport flows, aggregate transport networks, vehicle usage/specifications
Aggregate modal split	Limited data requirements	Weak theoretical basis Little insight into causality Limited scope for policy effects	Market share of a mode
Neoclassical	Limited data requirements Theoretical basis	Hard to combine in larger transport model system	Budget share of a mode in the total cost
Direct demand	Limited data requirements	Hard to combine in larger transport model system	Ratio travel time given mode to best mode Absolute travel time best mode
Disaggregate modal split	Theoretical basis Opportunity to include many causal variables and policy measures	Need disaggregate data (surveys)	?
Microsimulation approach	Many behavioural choices Links to theory included	Large data requirements or many assumptions on distributions Focus on urban vehicle transport	Observed data on distributions
Multimodal network (WCM & PCM)	Limited data requirements Theoretical basis Possible to include elastic demand and policies affecting generalized transport costs	Little insight into causality Mostly done with fixed demand	Transport time Terminal cost Transport cost Other port specific variables

Table 4: Freight transport models for modal split. Based on De Jong et al. (2004).

2.4 Models for assignment

In the previous step, freight in tonnes is converted into vehicle-units. In the last step these vehicle-units are assigned to networks. Theoretically, this means the truck, rail or inland waterway trips are assigned to routes that consist of links between these different modes. In practice, many models do not even incorporate the separate assignment step and most models only consider assignment to trucks. In this case, this is often done together with the assignment of passenger road traffic. This is because trucks usually make up a small proportion of total road traffic. It is also possible to replace the assignment step by a multimodal network that is described in the previous section (De Jong et al., 2004). The fact that a separate assignment model focuses merely on road traffic, makes it less relevant for this research. Therefore it is recommended to use the multimodal network approach so that the last two steps of the system can be covered jointly. However, it is important to keep in mind that there is one major drawback on the use of this model in step 4. Because an optimization mechanism (in the shape of a cost minimization algorithm) is used in this step, sometimes the output contains illogical mode-route solutions. Indeed travel choices, are in addition to cost also based on transport time.

Type of model	Advantages	Disadvantages	Variables
Separate assignment	Mode choice model can be disaggregate Allows interaction with passenger trips if freight and passenger trips are assigned together	No interaction between demand and assignment can be unrealistic Transport chains are difficult to incorporate	?
Multimodal network	Substitution takes place between mode-route combinations Chains with different modes on a route can be handled	Little scope for controlling the optimization process (sometimes the cost minimization objective causes illogical mode-route solutions because of omitted factors)	Transport time Terminal cost

Table 5: Models for the assignment step in freight transport models. Based on De Jong et al. (2004).

Other variables that affect port choice

It can be argued that general transport cost is not the only factor that explains port choice. There are many other variables that affect the competitiveness of a port and in this way have an effect on the route assignment of freight transport flows. These variables are usually mentioned in the literature as if they serve as input for regression analyses (the regression models described above). If a port has a better profile than their regional competitors, it is more likely that (for example) shippers will choose to ship their cargo through this port instead of the others. Ports compete based on transport relevant characteristics such as geographical location, connectivity, capacity, the capability to accommodate larger vessels (e.g. depth, quay length, etc.) (Garatt, 2006) and access to major markets in the hinterland (Polo & Guittérez, 2006).

Both Garatt (2006) and Polo & Guittérez (2006) point out that transport connectivity is the key feature of these four determinants. According to Garatt (2006) this is because 23% of container movements are transhipments. A port is more suitable as a transhipment hub if it is well connected with a lot of other ports (Hoffmann, 2012) and because containers are now more and more transhipped it is important for ports to be well connected. The connectivity can be measured by the liner shipping connectivity index (LSCI) developed by UNCTAD, which takes into account the deployment of container ships and container carrying capacity (TEU), the number of liner shipping companies, liner services and average and maximum vessel sizes (Polo & Guittérez, 2006). Polo and Guittérez (2006) support the assumption that connectivity is a key feature for ports by describing a vicious cycle. They state that a high LSCI means that the port serves containerized trade and thus, can attract regular lines. These regular lines facilitate trade and so increase the competitiveness of the port. To enhance this competitiveness, the port increases its transport offer and thereby attracts new cargo throughput, which gives the port a high LSCI.

Next to connectivity, location is very important for a port because when it is situated close to major markets this is favourable for the transport costs, which makes the port more competitive. However, the port may be close to major markets but it is also important that there is easy access to these markets. So the infrastructure of the hinterland is also a very important factor in determining the competitiveness of a port. Ports can also compete on costs in terms of import/export tariffs, prices of maritime transport services (Coto-Millán et al., 2005) and port fees (Hui et al., 2004). A straightforward way to include of a nearby port in a cause and effect model is to include the port throughput of the competing port in the analysis (Hui et al., 2004). Hui et al. (2004), who forecasted the container throughput in the port of Hong Kong using an Error Correction Model, included the throughput of the competing port of Shenzhen at time *t* in the model and regressed it on the throughput in Hong Kong at time *t*.

Also, fuel price should be mentioned as a variable (Gosasang et al., 2010). This is one of the most important drivers according to A.E. Willeumier of Port of Rotterdam (personal communication, June 25, 2014). The oil price determines the volumes and routes of trade. When oil prices are high, volume will eventually decline. Especially when the bunker price at a certain port is high in comparison to other (nearby) ports, shipping lines will try to avoid this port and throughput volumes will go down.

2.5 The additional logistics layer

Tavasszy (2006) argues that a fifth logistics step is needed to complete the freight transport modelling process, because logistics are changing and have impact on the distribution of freight. In the past decades it can be seen that transport costs have declined substantially. This is caused by the fact that supply chains are more efficient, which is made possible by globalization and the improvement and developments of information technology. Because of these cost and efficiency improvements in terms of logistics services, we are now heading to a time wherein "mass-customized logistics services" are the standard (Tavasszy, Ruijgrok & Davydenko, 2012). In this step the locations of distribution centres should be determined. This can be done by incorporating an inventory location model that helps to give a more realistic estimate of interaction costs, and so decreases the probability of bias in volume forecasts. Also, it helps to forecast the effect of changes in logistics services.

An approach to account for logistics cost can be to extend the gravity model, which is done by Hausman, Lee and Subramanian (2005) and Hummels & Schaur (2012). Not only transport costs are of influence of the probability of trade, but logistics costs (including transport costs) act as a resistance to trade. In both researches it is found that logistics cost have a significant effect on trade.

According to Tavasszy et al. (2009), to incorporate logistics choices in freight models it is necessary to use the Generalized Cost concept. This concept includes "all costs that are involved in overcoming time and space" (Tavasszy et al., 2009). It is assumed that companies have the objective to minimize these generalized cost. Tavasszy et al. (2009) propose this concept has to include shipment size, speed, value density, demand uncertainty, scale economies and network synchronization to cover logistics choices. A disadvantage of using the Generalized Cost concept is that it does not take into account trade-offs between variables. Models like the World Container Model and the Port Competition Model do take into account these trade-offs. They do, however, do not take into account the locations of distribution centres. Hence, the models should be adjusted to suit this layer or new models have to be developed.

2.6 Qualitative assessment

In the previous sections it became clear that some variables are hard to include in the model. This is predominantly caused by the fact that a lot of data is needed to make a variable usable. In this event, qualitative methods can also be used to forecast freight flows. This is especially done when historical data is not available. If these data are missing, a qualitative analysis and prediction of the variables can give insight in the movement and the influence of the variable in the future. The method relies on human, or expert, judgement on certain issues or the likeliness of an event to occur in the future.

There are different issues that cannot be accounted for quantitatively. Firstly, the economic structure of the base country of the port. This means a distinction is made between a manufacturing based and a knowledge based economy. In a knowledge based economy, the focus is on the service sector while in a manufacturing based economy the focus is on the production of (consumer) goods. For example, Hong Kong is now one of the biggest container ports, because a lot of manufacturing activities take place in the region. But the economy is now heading to a more knowledge based one, which will eventually mean that imports and export will decline (Seabrooke et al., 2003). This example makes clear that for this category, only the movement towards another economic structure is important to account for in a qualitative analysis. However, it should be noted that it is assumed a relationship between the type of the economy and income per capita exists. If income per capita rises, residents get higher educated and a gradual shift towards a more knowledge based economy can occur. This means that instead of the more widely assumed effect of a raise in income per capita is an increase in imports of consumer, it affects the import of semi manufactures and exports of manufactured products in a negative way. In this sense it is possible to account for the effect of a knowledge based economy in a mathematical model, by incorporating a variable for income per capita and a dummy variable for knowledge based economy.

Second, it is important to mention the influence of the market structure of the port sector and the associated market power of terminal operating companies (TOCs). These market players are expanding more and more globally. This means that they get more market power and thus the ports develop according to the strategic interests of these TOCs instead of that of the port. However, this factor is very hard to include in a model, because companies react to certain events in their environment that are unable to predict (Seabrooke et al., 2003). So, a qualitative assessment on this topic has to be made.

Little attention is given to the expert judgement method in the literature, probably because it is not a scientific and verifiable way to deal with forecasts. If, however, attention was paid to this method the approach was not broad or systematic enough to build on for further research (De Langen et al., 2012a). Below, the development of the Port Compass 2030 for the Port of Rotterdam is given as an example how cargo throughput is forecasted in practice.

An example: Port Compass 2030

For the development of the Port Compass 2030 for the Port of Rotterdam, a combination of a top-down and bottom-up approach is used to forecast cargo throughput. This means that macroeconomic modelling (top-down) is supplemented with expert judgment (bottom-up) to overcome the shortcomings of quantitative models in forecasting freight flows. Top-down approaches do not account for disruptions, policy effects and industry specific trends. Qualitative analyses on these points, made by experts in the field, can be used to learn about their possible effects.

For the Port Compass 2013 of the Port of Rotterdam, four scenarios where picked that were developed by Dutch planning agencies and the European Commission. The scenarios serve as input for the quantitative model (Transtools). The first scenario is *Global Economy* and stands for strong international cooperation and development of the private sector. It can be characterized as the scenario that predicts strong economic growth. The second scenario is called *European Trend* and is used to analyse European policy effects. It is basically a continuation of the present situation. The third scenario has to do with *Alternative Energy and Sustainability*, and affects mainly the flows of different commodities because alternatives are found for coal and other conservative sources of energy. The fourth scenario is the Low Growth scenario. It predicts slow growth in Europe and stagnation of global trade flows.

The quantitative model that is used is the Transtools model. It uses three steps in its calculations, based on the four steps described in the previous part of this chapter. In the first step, regional freight generation and the interregional distributions of freight flows are calculated resulting in an interregional trade table in tonnes/year. This number is derived from determining how much freight will leave a particular region (this is where production takes place) and how much freight will enter another region (this is where consumption takes place). A link is made between the "production" and "consumption" region, using a gravity model, resulting in a freight flow. In the second step, mode choice is taken into account using an aggregate multinomial logit model based on transport times and tariffs. Thirdly, network assignment takes place. This is done by translating freight flows in tonnes to number of transport units and assigning those units to the infrastructure network. This results in information about the intensities of traffic, vehicle type and fuel use (De Langen et al., 2012a). There are several drawbacks to this model. The most important one is that commodity specific trends and developments are not accounted for in the model. Next to that, as with every other forecast method, effects of external shocks in the environment (i.e. climate change) are not included in the model. De Langen et al. (2012a) therefore suggest to validate the results of the

TRANSTOOLS model with expert judgement and manipulate the results according to the statements of those experts, as is done for the Port Compass 2030. Also, the Transtools model can be labelled as a “blackbox”, it is not exactly clear what factors are included in the model. Some factors seem to be in it, but are not, and other factors are not expected to be included but are (L. Tavasszy, personal communication, July 11, 2014). This makes the software not very useful to use in the remainder of this research.

The Port of Rotterdam, however, did use the Transtools model and modified its results according to information obtained by experts. More than 50 consultants were interviewed, employed by the Port of Rotterdam authority and multinational firms that deal with transportation of goods. Also, governmental agencies were consulted. According to the ideas of these experts, the input or the output of the Transtools model was adjusted, resulting in the final forecasting outcomes.

For containers the outcomes of this research are that the multiplier between GDP and container volumes that is used by Transools should be lower in the future. This is because container freight consists for a large part out of intermediate goods. These volumes were very high due to global sourcing but are expected to mature in the future. Also, consumer goods were transported from low wage countries, a phenomenon that is considered temporary. Most important, the high container volumes were also due to mass containerization, a trend that is likely to mature in the coming decades. Further, transhipment operations will multiply, which is due to ever increasing ship sizes. The U.K. is considered to be an important transhipment location, therefore more throughput is added to the results for this country (De Langen et al., 2012a).

	Advantages	Disadvantages
Expert judgement	Gives more insight and awareness of disruptions that cannot be included in the model	Data is not measurable and cannot be verified

Table 6: Advantages and disadvantages of the expert judgement method. Based on De Langen et al. (2012a).

3 Essential components of a good forecasting approach

The aim of this chapter is to identify the essential elements of a good forecasting approach. A general forecasting approach is proposed, based on the models described in the literature. It should have all the characteristics of a good freight forecasting method that are identified in paragraph 3.1. A stepwise approach is given to come to a forecast of container cargo throughput in ports.

3.1 Characteristics of a good forecasting method

From the previous chapter it becomes clear that a good cargo forecasting method has several characteristics. First, in the maritime industry data is often not available in large quantities and detail. This makes it better to use a model that does not require a lot of data. Also because using a lot of different data makes the forecasts less reliable, since it is more prone to error (Bankes, 1993). Second, the model should make clear how trade flows are generated. This is based on information about supply and demand in different regions. Based on this information, supply and demand (or production and consumption) regions are linked to each other to show how these trade flows are distributed over the world. Here, macroeconomic variables are used to forecast freight flows. Third, factors that determine the competitiveness of a port have to be taken into account. Because not only the existence of trade explains the cargo throughput in ports but also the attractiveness of the port. This is because many ports in the world serve the same hinterland and port specific characteristics determine which port gets the biggest share of the trade flow that is destined for this hinterland. Finally, preferably the model accounts for shocks in the economy and trends and developments in the industry. It has shown that this is very difficult to include in a mathematical model. Therefore, it is also possible to analyse and – if it is desired - adjust the results of quantitative models by use of expert judgement on these trends and shocks.

3.2 The generation and distribution of trade

Flows of trade form the basis of a freight transport model. Therefore, as a first step, generation and distribution of trade flows should be determined. As mentioned before, GDP is an important indicator of trade. This is supported by the papers of Stack (2009), Hausman et al. (2005) and De Groot, Linders, Rietveld & Subramanian (2003). The results of these researches state that 70 percent of bilateral trade is explained by GDP, GDP per capita and distance between countries. This arises the assumption that the gravity model is a good tool to forecast freight flows between countries. The assumption is confirmed by the fact that the majority of freight transport modelling software uses the gravity model for the freight distribution module (Europe & Brinckerhoff, 2002).

First, GDP should be forecasted in order to use this model. There are organizations that are dedicated in doing GDP forecasts. Therefore, it is recommended to base the forecast on their findings. An example of an organization that brings out GDP forecasts is the IMF. Their forecasts are widely used and considered acceptably reliable. IMF forecasts go only 8 years ahead. Port throughput volumes must be forecasted for at least 25 years forward. The trend of the IMF should be extrapolated into the future in order to get this 25 year forecast of GDP. This can be done by continuing the path of the recent past into the future as is done in the UNESCAP report 2007 on container traffic forecast. It is also possible to use the forecasts of the Organization for Economic Co-operation and Development (OECD) as these forecasts are made for 50 years ahead. However, this dataset is less extensive in terms of country specific forecasts. Next OECD-members, that are mainly high income economies, only a few non-members are present in the economic outlook dataset.

The GDP forecasts that are obtained serve as input for the gravity model that estimates the distribution of trade flows. This model builds on the assumption that trade volumes depend on both the importing and the exporting country's GDP (Stack, 2009; Hausman et al., 2005; De Groot et al., 2003). So, a gravity model is used that includes the overall GDP and the GDP per capita of the importing and exporting country, based on the research of De Groot et al. (2003). The model is expressed mathematically is shown below.

$$\ln T_{ij} = \beta_0 + \beta_1 \ln Y_i + \beta_2 \ln y_i + \beta_3 \ln Y_j + \beta_4 \ln y_j + \beta_5 \ln D_{ij} + \beta_5 A_{ij} + \varepsilon$$

With

T_{ij} aggregate merchandise exports from country *i* to *j*

$Y_{i,j}$ total GDP for country *i* resp. *j*

$y_{i,j}$ GDP per capita for country *i* resp. *j*

D_{ij} distance between country *i* and *j*

A_{ij} trade arrangement dummy

As mentioned above and can be seen in the model, 70 percent of bilateral trade is determined by GDP, GDP per capita and the distance between the countries. A trade arrangement dummy is included to account for factors that reduce the resistance of trade, measured by distance. Note that this is a very basic version of the model. It is also possible to include more dummy variables that tell something about characteristics that are shared by the two countries, like religion, border, primary language etc. (De Groot et al., 2003). Also, adjustment can be made to change the explanatory power of the model. Please refer to Polder (2000) for an extensive description on this. It should be noted that the gravity model is complicated in case a port also handles import and export cargo for neighbouring countries. Especially for countries that do not have a seaport connection on their own. This makes that the macroeconomic conditions in these countries are also important to predict freight flows. The model should be estimated using OLS.

3.3 Assignment of container flows to the port network

In the second step, the container flows that are determined in the previous step should be assigned to the maritime transport network. To do this, first a maritime networks should be mapped, containing nodes and links between different ports and (main) sea routes. For example, Tavasszy et al. (2011) build their transport network based on weekly container liner routes that are provided by container shipping lines. Information on liner schedules and transport fees that are publically available were used to construct a database containing 800 maritime services and 437 ports. Eventually, the network was developed by using one land node per country (origin and destination), port nodes and maritime nodes. As a result the transport network was illustrated as shown below.

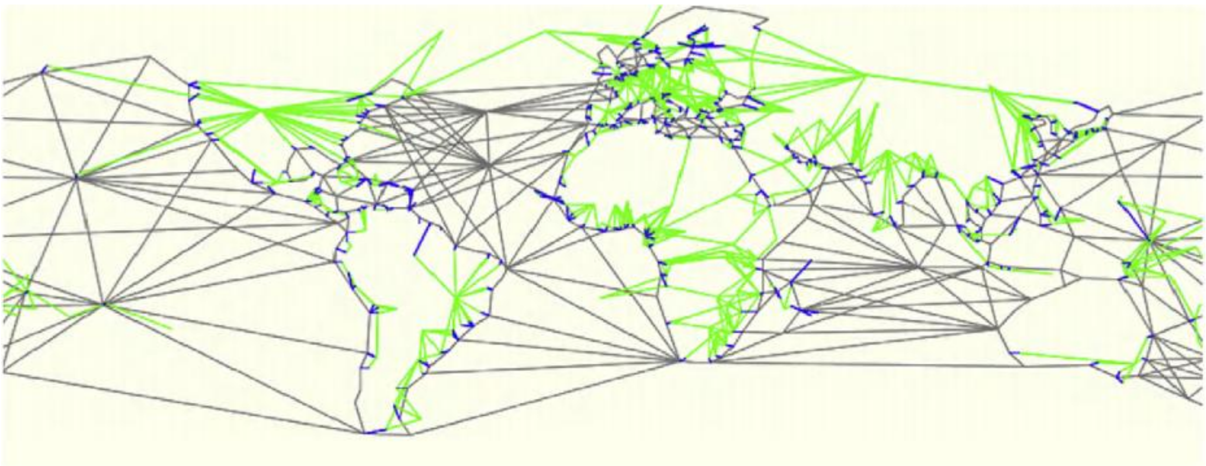

Figure 3: Maritime network with country access links o and from ports. Source: Tavasszy et al. (2011).

Next, port competition should be included in the approach. In the previous chapter, two port choice models are discussed briefly: the World Container Model and the Port Competition Model. Here, a more extensive (mathematical) description is given. The objective of incorporating a port competition model is to account for characteristics of a port that determine its competitive position with respect to other ports in the region.

There are several differences between the World Container Model (WCM) and the Port Competition Model (PCM) causing preference for one of the two models in particular situations. The WCM is focused more on the maritime side of the transport chain, while the PCM also strongly focusses on hinterland transport and mode choice next to seaborne trade. For example for a Greenfield port or a transhipment port, the WCM is more suitable because it has a global focus and predicts the share of a certain port in the global network. Also, it incorporates transhipment explicitly. On the other hand, for import-export ports, hinterland connections are very important. Hence, the PCM is more suitable for these kind of ports. It could be useful to use both models for forecasting throughput volumes if a focus on both the maritime network and the hinterland network is preferred.

3.3.1 The World Container Model

Using the WCM, container flows all over the world can be predicted for different scenarios. In contrast to what the model suggests, this model is also suitable to predict the throughput volumes in individual ports based on a worldwide network approach (L. Tavasszy, personal communication, July 11, 2014). The assumption made in the model is that the distribution of container flows among the world's seaport network depends on the geographical location of the origin and destination of the containers, the cost of overseas and inland transport, the forwarding organisation's preferences regarding merchant or carrier haulage, the logistics characteristics of the goods in terms of unit value for example, the available facilities and services (for example infrastructure), and the decision agent's characteristics and strategies (Tavasszy et al., 2011).

Based on this a simple logit route choice model using path enumeration was extended into a path size logit model[1], wherein the choice probability of the route is the dependent variable. Because a profit maximizing objective is assumed, the probability of choice for a certain route depends on the costs of a route (and the path size overlap variable, to assure that overlapping routes will not be overestimated), expressed mathematically as follows.

$$P_r = \frac{e^{-\mu(C_r + lnS_r)}}{\sum_{h \in cs} e^{-\mu(C_h + lnS_h)}}$$

With S_r the degree of path overlap, expressed as follows:

$$S_r = \sum_{\alpha \in \Gamma_r} \left(\frac{z_\alpha}{z_r}\right) \frac{1}{N_{ah}}$$

[1] This was done because in the simple logit model, the routes where bundled and therefore the routes that overlap (are in the same bundle) would be overestimated. So, a path size overlap variable is added to the model. This is a variable that depicts the degree of overlap based on the sum of the relative lengths of the routes, divided by the number of times a link is found in alternative routes (Tavasszy et al., 2011).

α_r the link in route *r*

Γ_r the 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

And the cost of route, C_r:

$$C_r = \sum_{p\in r} A_p + \sum_{l\in r} c_l + \alpha \cdot \left(\sum_{p\in r} T_p + \sum_{l\in r} t_l \right)$$

A_p total cost of transhipment at port *p*

c_l total cost of transportation over link *l*

T_p time spent during transhipment at port *p*

t_l time spent during transportation over link *l*

α value of transport time (USD/day/ton)

Modal split is not explicitly expressed in the model. Instead the developers chose to use a "mode-abstract formulation", which enables the researcher to use a more detailed underlying multimodal network. The parameter A_p is made of several characteristics that measure the service level at the port, like fuel costs, handling costs, congestion costs, etc. Another important parameter in the model is the value of time. This variable indicates the preference of shippers for fast and more expensive or slow and cheaper transport. This parameter contains all the characteristics of the good, that are also mentioned above. Next a set of route choices is determined using a shortest path algorithm to decide what is the preferred route between the origin country and the destination country, via the ports in the country. This can be pictured as follows.

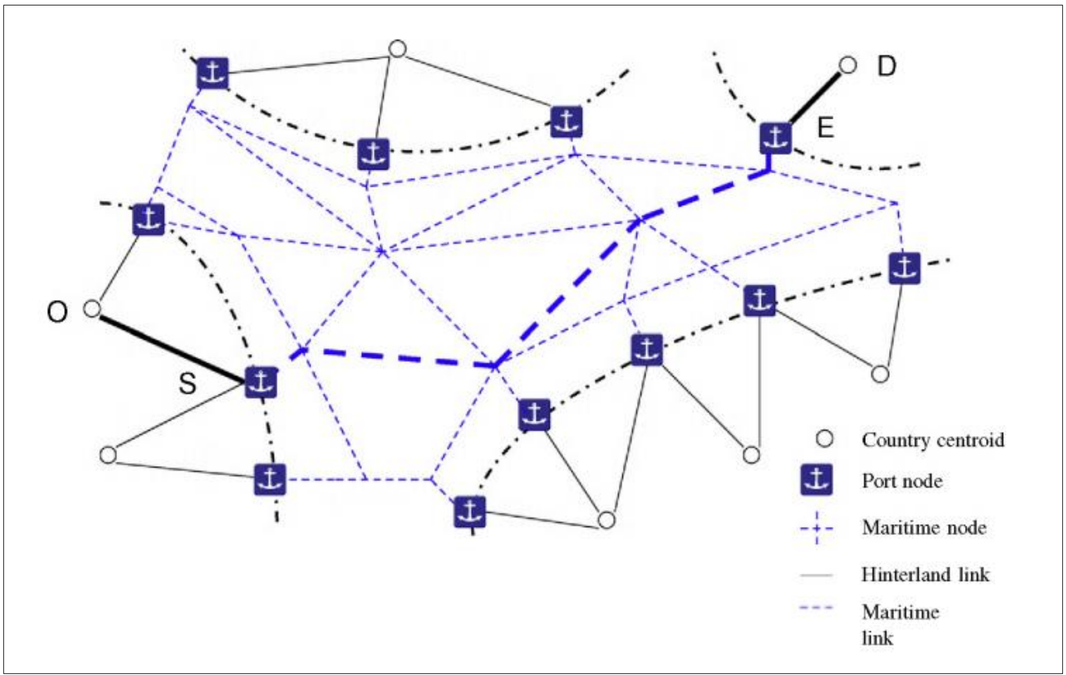

Figure 4: Routes between an OD-country pair. Source: Tavasszy et al. (2011).

Where O and D stand for origin and destination, respectively, and S and E are the ports of these countries. The shortest route is the thick scattered line (Tavasszy et al., 2011).

In their paper, Tavasszy et al. (2011) use the WCM to estimate the effect of different scenarios that illustrate a change in the transport system, like opening a shipping route over the North instead of via the Suez canal. Using the WCM, the authors show the changes in the proportion of total container flows that comes into a port, per scenario. It is an interesting feature that the model can account for possible changes and show the effect of these changes. The forecasting of effects of changes is very hard to do with a model and therefore, qualitative assessments are used. The WCM gives the opportunity to analyse the effects of changes in the transport system quantitatively.

3.3.2 The Port Competition Model

Based on demand choice models, the PCM forecasts the relative market share of a port in a specific region as it comes to container throughput. The model depends heavily on the four step approach that is described in the previous chapter and is usually used in the transport modelling literature. For the convenience of the model, the steps are shaped into certain "moments of choice" as follows:

1. Production and attraction is the choice to buy or sell a specific good and the amount of the good.
2. For distribution the choice is to buy from a specific supplier and sell to a certain group of customers.

3. As a last step a choice is made for a specific transport alternative in terms of routing and mode.

So, in this last step a choice has to be made for a mode for the overland transport from origin to a seaport of departure. Next, a choice is made for an ocean shipping service followed by choosing a port and a mode for the hinterland transport to the final destination. In the paper that discusses the PCM, the focus is on the North Sea ports, which causes the authors to restrict the analysis to the last end: the ocean route, the port of call and the hinterland transport. This is done because choices made at the origin in terms of transport costs and opportunities do not affect the choices made in Western Europe (Veldman & Bückmann, 2003).

The assumption is that the choice for a specific combination of sea transport, port and hinterland transport is made based on costs and quality of services. These can include transit time or frequency of service, which are already quantitative variables. There are also a lot of factors that determine port of choice that are not quantifiable, such as reliability of service and responsiveness to customers' needs. The logit model is expressed mathematically as follows.

$$P_m(m = r|r = 1 \dots M) = \frac{e^{U_m}}{\sum_{r=1}^{r=M} e^{U_r}}$$

Where,

P_m probability of choosing routing *m* from all possible routings, r = 1…M

U_m utility resulting from choosing route *m*

m routing index

The utility function U_m is derived as follows.

$$U_m = \alpha_{0m} D_m + \alpha_1 C_m + \alpha_2 T_m + \alpha_3 F_m$$

Where,

D_m dummy variable that indicates if shippers have a preference for routing *m*

C_m shipping cost of routing *m* (incl. freight rate, handling charges, hinterland transport)

T_m transit time for routing *m*

F_m frequency of service of routing *m*

The generalized costs of a certain alternative are derived by dividing the utility with the coefficient of the shipping cost of that routing (α_1). Eventually these equations result in a formula for the market share of a port, which is described by the following.

$$\frac{P_m}{P_n} = \frac{e^{U_m}}{e^{U_n}} = e^{U_m - U_n}$$

So, the market share ratio is a function of the differences in utilities, determined by characteristics of the port in terms of shipping cost, transit time and frequency of service. The dummy variable D_m stands for all other factors that affect port choice, like the unquantifiable variables mentioned above. The ratio of two different coefficients is calculated to find the trade-off between two variables. For example α_1/α_2 illustrates the trade-off between shipping costs and transit time. The values of these coefficients depend on several factors like geographical setting, model specification, choice situation and the level of aggregation (Veldman & Bückmann, 2003).

The variable transport costs stands for the costs of transporting a container from the port of choice to the destination in the hinterland. The costs for sea transport are deliberately omitted, because carriers charge the same tariff to each port. Transport time is the time measured between the departure of the container from the port and the arrival at the final destination in the hinterland. For the frequency of service variable the reciprocal of the frequency is used. This is expressed by the average inter-arrival time (IAT) of two sequential calls of a liner vessel or another hinterland transport mode. This displays the fact that a port that offers many services is more attractive. This concept can also be applied to the different hinterland modes. The busier the mode, the more attractive it is assumed to be. It is also noted by Veldman and Bückmann (2003) that the model is improved by adding dummy variables for the three different inland transport modes: road, rail and inland waterway.

3.4 Qualitative assessment of the results

Many forecasting approaches use scenarios as input for their quantitative models. These scenarios are often developed by public agencies and include assumptions on future developments. Although scenarios account for uncertainty and give some direction for the future, they do not account for external shocks in the economy, because they are mainly based on historical data.

For quantitative models in freight forecasting, usually publicly available data is used. These data are globally focused and thus not very detailed. No wonder these models cannot account for strategic developments in production networks or technological innovation, for example. Therefore, it can be useful to analyse breaks in industry specific trends and assess their effect on the projections made by these models. This will improve forecasts in general (De Langen et al., 2012a).

The qualitative assessment should be based on interviews conducted with several players in the field. This can be experts working at the port authority, but also people from businesses that operate in the port. Also, experts from public organizations can be useful to complete the qualitative analysis of industry specific trends and possible breaks herein.

A way to incorporate expert knowledge in the complete modelling approach can be derived from the approach used for the port of Rotterdam, which is described in more detail in the next chapter. There, existing scenarios developed by the European Commission and Dutch planning agencies are used as input in a quantitative model (Transtools). The output is analysed by experts like the ones described above. Based on these analyses, the input (scenarios and I/O tables) or the output of the model (forecast) were modified (De Langen et al., 2012a). This process is illustrated in the figure below.

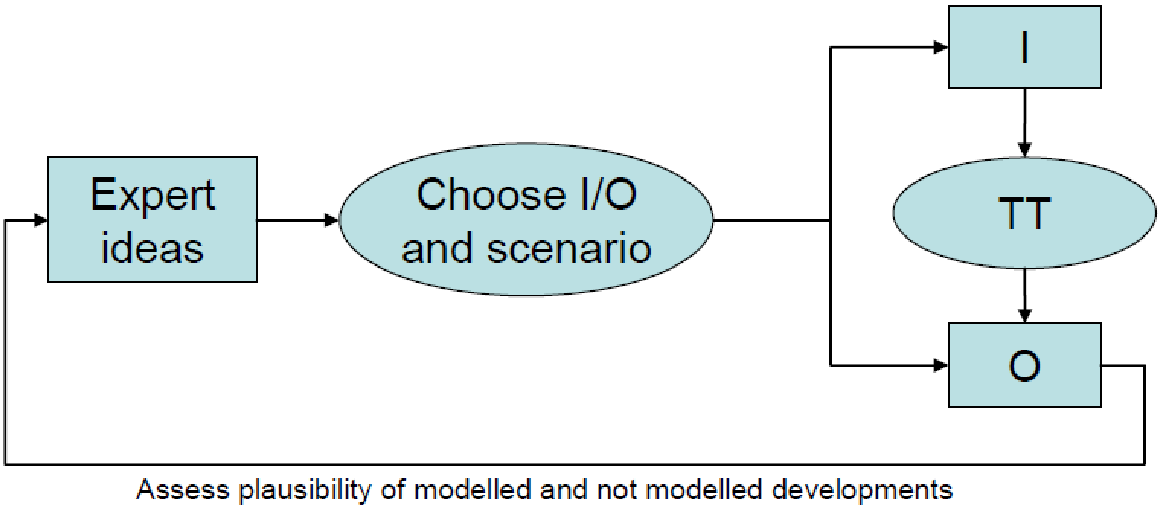

Figure 5: Combination of model results with expert judgment. Source: De Langen et al. (2012a).

4 Container cargo forecasting in three port categories

Demand for port services depends on international trade and economic activity. This makes that market conditions, international and in the residence country and hinterland of the specific port, are important factors. It can be assumed that these conditions, especially for the base country and the hinterland, differ from port to port, even as the port characteristics. It is therefore important to gain more knowledge about these conditions and port characteristics before making port throughput forecasts. Three main port categories are distinguished in this chapter, these include whether the port is mainly a transhipment hub or an import-export port, the level of economic development and if the port is a Greenfield or a Brownfield port. For each category, the situation is explained and two opposing ports are compared with each other in terms of container throughput and indicators that explain this throughput. Historical data on throughput volumes in these ports will be used for this analysis. Except for the analysis of the Greenfield and Brownfield port, where the forecasting method of each port is described and analysed. The purpose of these analyses is to identify whether the characteristics of a port ask for different modelling approaches and variables to be included in the models. A short overview of this is given in the last paragraph of this chapter.

4.1 Transhipment vs. import-export

Ports can offer import, export and transhipment services. Import and export services are the obvious ones and have a clear relation with economic growth as they directly explain trade and thus port throughput. Transhipment is a more complicated issue in forecasting container throughput as it has less to do with the economic situation of the hinterland of the transhipment port and thus the destination of the transported product. As a matter of fact, port throughput can be explained based on macroeconomic factors or regional port competition variables. Transhipment has something to do with the latter category. The success of a port that focuses on transhipment services mainly, heavily depends on the main routes of the biggest liner vessels. This is because the costs of diverting from these routes have to be low in order to make transhipment profitable, so the port should not be situated too far from the main routes. Also, the distance from the port to the market served and the handling costs in the transhipment port are of importance for the competitive position of the port as an important transhipment hub in the maritime network (Zohil & Prijon, 1999). To include this view into a model, Zohil and Prijon (1999) proposed a simple regression model with the estimated number of TEU transhipped in a port as dependent variable. This number is explained by the independent variables *hours,* the diversion distance in hours; *traffic*, total port traffic throughput; and *traffic/hours,* the quotient resulting from the first two variables. The authors claim that the results of their research suggest that the formula is "reliable enough".

Over the last decades transhipment operations have increased fast and will continue to do so in the future. Global trade has increased very fast in the past decades due to globalization and containerization. Around 28% of all container throughputs are transhipped. In figure 6 the increase in container transport and transhipment is illustrated. An incremental increase of 328.8% for total throughput volumes is identified between 1990 and 2012. For transhipment volumes the increase was 459.6% for the same period (Notteboom, Parola & Satta, 2014).

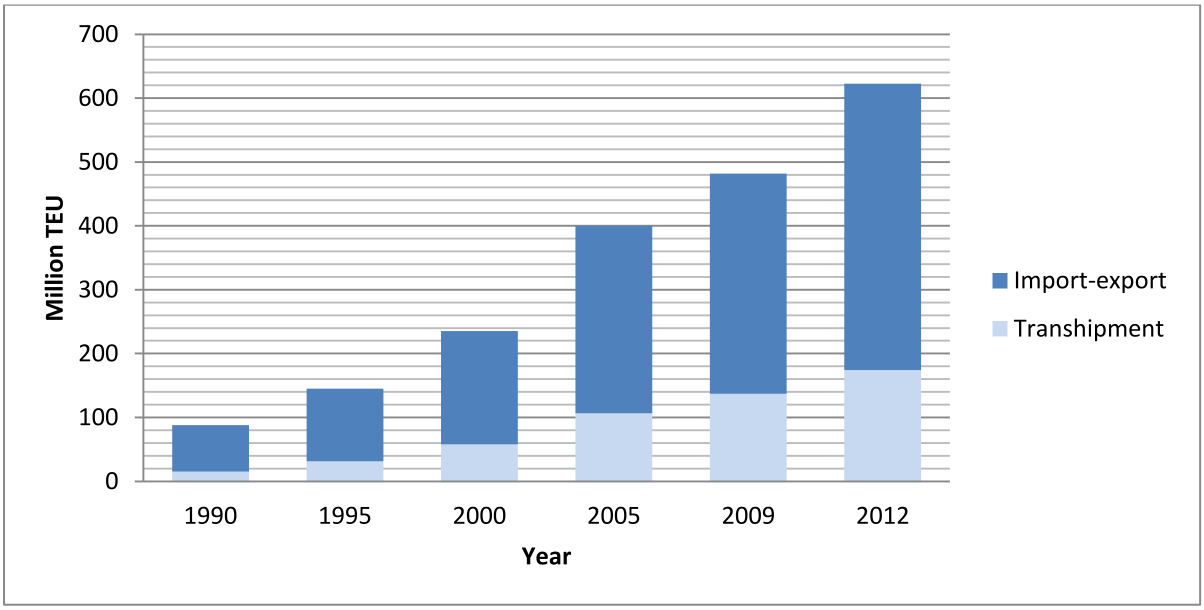

Figure 6: World container throughput volumes and the share of transhipment (million TEU). Based on: Tavasszy et al. (2014).

This increase is also due to the fact that ships will get bigger, so that shipping lines can profit from economies of scale (Baird, 2006; De Langen et al., 2012a; De Langen et al., 2012b). By making use of hub-and-spoke systems, the main liner vessels only have to enter big transhipment ports that can handle big ships with a deep draft. From there, feeder ships can take the freight to smaller regional ports (Lam & Iskounen, 2010). In general, the main focus is on the cost savings shipping lines achieve by making use of hub-and-spoke systems. However, Lam and Iskounen (2010) and Baird (2006) mention the downside of the emergence of transhipment ports. Baird (2006) states that the benefits of transhipment are a trade-off between extra feeder and handling costs against the extra costs of calling at an additional port. He argues that for none of the transhipment ports in Nothern Europe the feeder and handling costs of transhipment services are lower than the costs of calling an extra port. Thus, a dedicated transhipment port should be established to serve the whole Northern

European market from one point (Baird, 2006). However, it should be noted that this article dates from 2006 and ships are now approximately twice as big as they were then (9,600 vs. 18,270), which means the draft is deeper. This causes that the biggest vessels cannot call every port in Europe anymore, and thus transhipment has become more relevant.

No information is found in the literature about the effect of transhipment on the total throughput in a port. However, it is clear that transhipment increases the throughput volumes in numbers. From the information above, it can be reasoned that a port is suitable for transhipment if it is not located too far from the main liner shipping routes, as long as the feeder costs do not surpass the costs of just making another stop and minimal port depth is required for large vessels. So, the extra throughput generated by transhipment services has nothing to do with the economic conditions of the base country. However, it can be the case that in more developed economies, port services are more efficient which lowers the costs of transhipment in terms of time. In the light of this research, it is assumed that transhipment volumes partly explain the competitive position of a port. Hence, transhipment is a factor that should be accounted for in the port choice part of the forecasting rather than the trade generation and distribution part.

In the remainder of this section, the development of throughput volumes for containers over the years are discussed for a typical transhipment and import-export port. First, the transhipment port of Salalah in Oman is discussed. The annual reports for the years 2004 until 2013 are used to give an overview of the growth or decline in container volumes of the port. Next, the typical import-export port of Auckland in New Zealand is discussed. Throughput volumes are presented for the years 2004-2013 and the drivers of change in volumes between 2008 and 2013 are described. The section ends with a short comparison of the two types of ports.

4.1.1 The port of Salalah, Oman

The port of Salalah in Oman is a port that is specialized in the handling of containers and general cargo. Because of its strategic location on route to the Suez Canal, where much of the transport from Asia to Europe and vice versa passes, the port is perfectly suitable for transhipment services. This is reflected in the fact that the share of transhipment cargo handled in the port was 99.5% in 2004. The market share of the port of Salalah in the Middle East region is 10% regarding total container transport and 20% for total transhipment of containers (Port of Salalah, 2006). It should be noted that transhipment at this point of the Asia-Europe route is not required because not all ships can pass the Suez canal. The draft of the Suez canal is deepened to 20.1 meters in 2010 (Thomson Reuters, 2010). This deepness does not give problems for the biggest container vessels as those with the deepest drafts lay 16 meters under sea-level. There is even a decline in draft going on, as ships get wider and therefor require less depth, i.e. 15 meters (Rodrigue, Comtois & Slack, 2013). Therefore it is assumed that the main driver to use transhipment in the Middle East is economies of scale.

Figure 7: The location of the port of Salalah. Source: www.vacationstogo.com.

From 2004 until 2013, container throughput in the port of Salalah increased by somewhat over 1 million TEU, see figure 8. In the first years global container trade was growing fast, resulting in an increase in container throughput volumes in the port. The drop in volumes in 2006 is caused by the shipping line Maersk, that decided to realign its vessel schedule. This illustrates that seaports are very dependent on the strategic decisions of other players in the market, like shipping lines. Unfortunately, strategic decisions of shipping lines cannot be predicted, which makes it hard to account for its effects.

The fast growth between 2006 and 2008 is caused by the high oil price in that period. This increased investments in infrastructure in the Middle East, causing container throughput to grow by 12%. The position of the port of Salalah at this point is stable, having the certainty of long term contracts with the three major shipping lines Maersk, Mediterranean Shipping Company and American President Line. However, in 2008 the credit crisis started, causing freight rates to fall by 15%. Together with the high oil prices, shipping lines went to a hard time. On top of that, shipping lines started to invest in Ultra Large Container Ships (ULCS) to profit from economies of scale, triggered by the high levels of uncertainty in the market. This resulted in over capacity. This all led do a decrease in container throughput between 2010 and 2011. However, economies of scale is one of the biggest drivers of transhipment, causing the port of Salalah to recover quickly.

As can be seen in the last few years, the container transhipment market is volatile. More than import-export markets, because shipping lines have more options to tranship their cargo somewhere on the route. The final destination of the freight does not matter in that context. Hence, transhipment ports highly dependent on shipping lines. These shipping lines have more and more power as many of the pure transhipment terminals are (partly) owned by carriers and international terminal operators. Therefore, fluctuations and transhipment volumes are hard to predict (Notteboom et al., 2014).

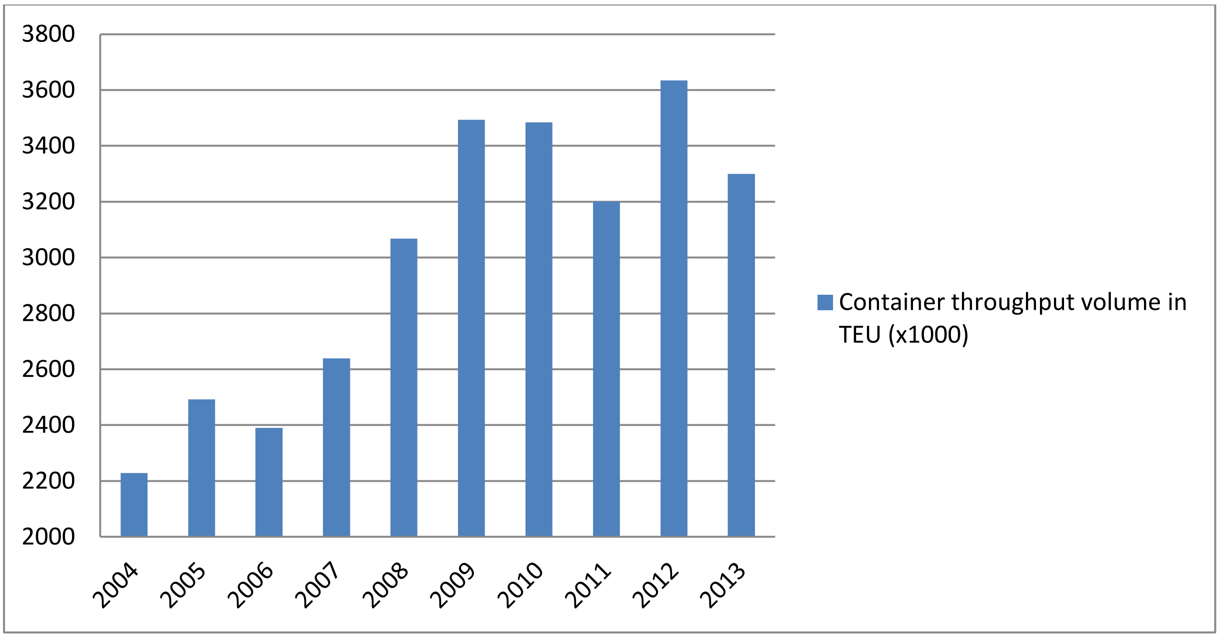

Figure 8: Development of container throughput in the port of Salalah between 2004 and 2013.

Based on this analysis it can be concluded that it is very hard to model what drives the container throughput in transhipment ports, because this heavily depends on the strategic decisions of shipping lines and these cannot be predicted. It is true that transhipment ports compete in service levels and location.

4.1.2 The port of Auckland, New Zealand

As transhipment is gaining ground in the container shipping industry, core import-export ports are hard to find. The port of Auckland is chosen for this analysis, because its location is not attractive for transhipment services. Many of the transhipment ports are located on the East-West route and thus in South East Asia, Latin America, Southern Europe and the Middle East. Transhipment volumes in Oceania in 2012 were 0.3% (Notteboom et al., 2014), which makes the port of Auckland a typical import-export port.

Figure 9: Location of the port of Auckland. Based on: www.worldatlas.com.

Container throughput volumes from 2004 until 2013 were also found for this port. However, the review on the development of the throughput volumes starts from 2008 onwards.

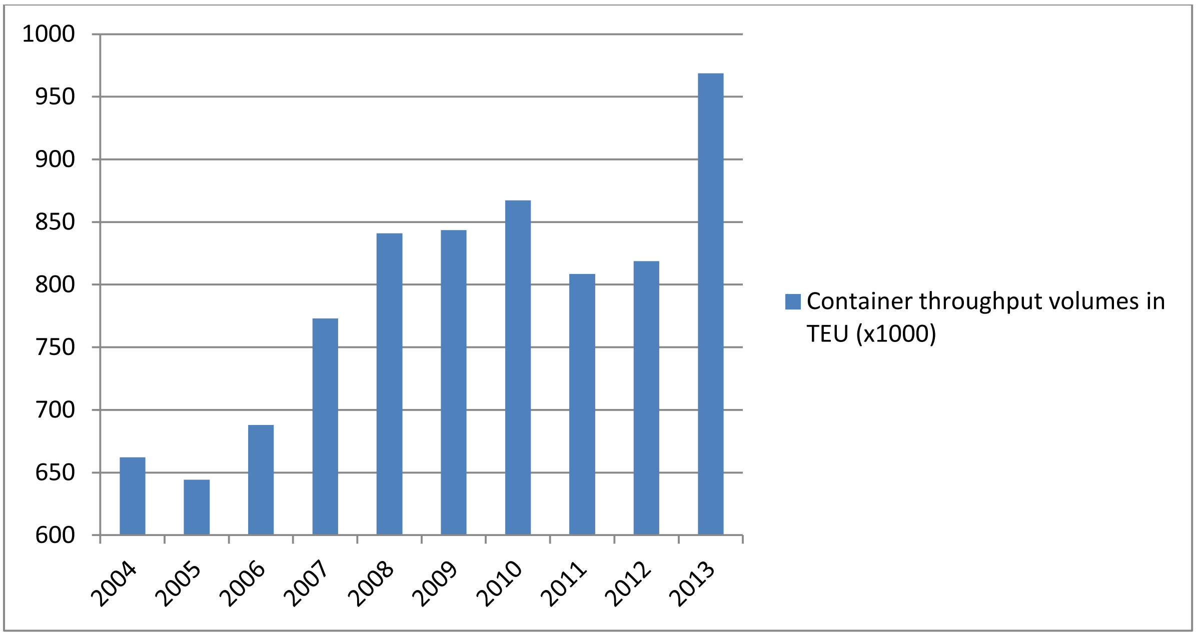

Figure 10: Development of container throughput in the port of Auckland between 2004 and 2013.

In the years 2008-2010 container volumes were small because of volatility in the shipping industry. Shipping lines were rationalizing their schedules to save costs and achieve greater efficiency. This effect is also visible in the port of Salalah. Transhipment volumes all over the world increase caused by liner shipper's search for cost savings and greater efficiency, investing in bigger vessels to benefit from economies of scale. This also effects the port of Auckland, where transhipment volumes increase by 55% in 2008 (Port of Auckland, 2009). Because the port of Auckland is located far away from the main transhipment routes, it is assumed these transhipment volumes are destined for smaller ports in New Zealand. This means the same country is served and thus the same macroeconomic variables are of importance to predict the freight flows. This is exactly the difference between transhipment in the port of Salalah and in Auckland, as for the first macroeconomic variables do not predict transhipment volumes. In 2009, transhipment volumes increased even more by 10%. Although this is a big increase, import and export seem the core services of the port. This indicates that transhipment rates before 2008 were probably extremely low, causing an increase in percentages to have little impact on entire throughput volumes.

The abrupt fall in throughput volumes in 2011 is a result of an industrial dispute, causing strikes and actions on the terminal. Container volumes fell down by 9.6% because of loss of significant services due to the strikes on the terminal (Port of Auckland, 2011). No agreement was reached for the beginning of 2012, which explains the small growth in container volumes. Meanwhile, the port started with re-organizing its business into a more customer focussed organisation. On top of that, operations became more efficient and flexible. This made it possible to attract Maersk back, one of the shipping lines that were lost during the industrial dispute, and attract new customers. Also, the deepening of water lanes in the port makes it accessible to larger container ships. These factors enabled the enormous growth of container cargo in 2013.

From 2008-2013 the proportion of imports and exports are presented. Although slow, there is a clear shift from exports to imports. This perfectly illustrates the fact that the size of the economy, determined by population and income, is a driver of imports. As can be seen in the figure below, the hinterland region of the port of Auckland, New Zealand is growing. The entire country of New Zealand is considered the hinterland, because the port of Auckland is the most important import port in New Zealand (Port of Auckland, 2011).

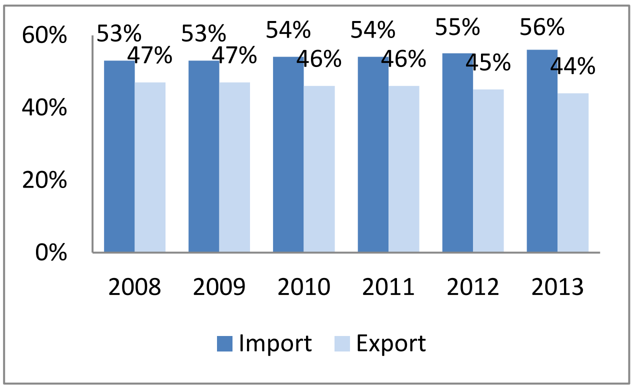

Figure 11: The proportion of import and exports in the port of Auckland.

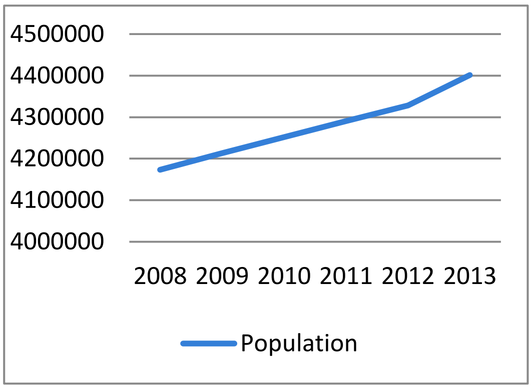

Figure 12: Population growth of New Zealand. Based on: www.census.gov.

4.1.3 Comparison Salalah with Auckland

What stands out is that both ports felt the effects of the worldwide financial crisis. The effects were generally caused by the changing behaviour of shipping lines, that changed their schedules to lower their number of port calls causing container throughput to decline. However, the effects were short lived. The uncertainty of shipping lines resulted also in the investments of bigger ships, causing an increase in transhipment volumes. This benefited the port of Salalah and, more surprisingly, also the port of Auckland. Although the latter is a typical import-export port, transhipment volumes increased by 55% in 2008. This is caused by the fact that differences between transhipment ports and import export, or gateway, ports become smaller. Main import-export ports are now also offering transhipment services, causing core gateway ports to disappear (Gouvernal, Debrie & Slack, 2005). Transhipped cargo was destined for the same hinterland in case of the Auckland port, however, so the same macroeconomic variables explain the trade flows. These effects show that both ports are dependent on the strategic decisions of shipping lines. This is also reflected in the decrease in throughput in the port of Auckland when strikes are going on and the fast increase as services are improved. So, for both types of ports competition is very important. Hence competition should be modelled for both ports, where quality of service in terms of (cost) efficiency and flexibility is equally important in both Salalah and Auckland.

A difference between the two types of ports is that developments in throughput volumes in Auckland are partly determined by the hinterland. As the population of New Zealand is growing, the percentage imports in relation to exports becomes higher. For the port of Salalah, such relations are not present. This difference can result in another formulation of the gravity model, where for the port of Salalah only distance is important as indicator of trade resistance and for the port of Auckland as many macroeconomic hinterland characteristics as possible should be added to the model, like population, GDP per capita, etc.

4.2 Greenfield vs. Brownfield

Port expansion projects for container terminals are based on a container cargo throughput forecast. This is very important because the required capacity of the new terminals is determined based on this forecast. Port expansion projects can either concern new ports (Greenfield) or expansion of existing ports (Brownfield). Unsurprisingly, historical data on throughput volumes is not available in case of Greenfield projects. This rises the assumption that when it comes to forecasts, Greenfield and Brownfield projects require different approaches based on the amount and nature of the data available. In case of a Greenfield port it can be that another port is already situated in the country. So there is already a flow of goods, and the question is how this new port will attract cargo and so competes with the existing port. This relies heavily on competitive characteristics like if the port is suitable for transhipment (see 4.1), or the hinterland connections of the port.

Most of the above described models raise the assumption that they focus on existing ports. For these ports, historical data on throughput volumes is available and can be extrapolated into the future. However, extrapolation of trends is not the desired approach for forecasting future throughput. So the question is, is there a substantial difference between Greenfield and Brownfield projects in choosing the best models? The difference in this situation is that for Brownfield ports, this competitive position can be expressed quantitatively by making use of, for example, the LSCI (see 2.1.2.1) (Zohil & Prijon, 1999). In case of Greenfield ports, no numbers are available and so the analysis of its competitiveness should be a qualitative comparison with the biggest competitors or based on goals. This makes the forecast less reliable, because it is based on predictions on the competitiveness of the port rather than on current situations.

In the following two sections the forecasting approaches for a Greenfield master plan project and a Brownfield port are discussed. The Greenfield port is Filyos in Turkey, a port for which a forecast is made but the construction has not started yet. This means no data is available on cargo throughputs and thus the forecasting approach will be described and analysed, using the CFCU Study Report dated March 2009. The Brownfield port is the port of Rotterdam, for which the Port Compass 2030 is used. This is not a forecast made for a particular expansion project, but it provides guidelines for the entire port future. After the two ports are described separately, a comparison is made.

4.2.1 The port of Filyos, Turkey

Maritime transport nowadays mainly goes over Istanbul, which serves as a gateway for the largest part of the inland of Turkey. As a result, Istanbul and the Bosporus strait are suffering from congestion. Therefore, the government of Turkey initiated the construction of a new container port at Filyos in the Zonguldak region to shorten inland routes as well as routes over the Black Sea to Russia and Ukraine. Because other ports in this region are not suitable for expansion, due to limited space, other core activities than containers and lack of good hinterland connections, it is chosen to construct a whole new port. Filyos already has a road connection to Ankara, which will serve as the main hinterland route. The other advantage is that it can serve as a hub for containers and general cargo that are destined for other countries (and thus ports) that are located at the Black Sea.

Figure 13: An overview of the transport routes in Turkey. Based on Donders (2010).

The throughput forecast made for the port of Filyos, that is supposed to handle general cargo and containers predominantly, is structured as follows. The method starts with the construction of a database on transport flows. This is done by using a transport chain structure and linking locations of production and consumption to assign imports and exports of the country to regions within Turkey. The transport flows are also assigned to the road and rail network of Turkey. This approach is based on the four step framework described in chapter 2. Two data categories serve as input for this model, trade data and transport data. The former contains data on the main market segments by product type and country of origin and destination, the so called O/D matrix. More specifically, data on regional chain structure, UN trade data of international flows, other flows based on socio-economic data and the location of the biggest industry and transport generators is used. For transport data, port statistics on demand at specific points in the Turkish transport network are used. To fill in gaps in the transport and trade data, a comprehensive qualitative analysis is made based on interviews with stakeholders. These interviews are important because they shed light on strategic issues that concern stakeholders of the port. It was intended to gain information from as many as possible differing parties of interest in the port, from governors to shipping agents.

For the forecast of trade flows the global trade model developed by NEA is used. The model uses historical trade data and socio-economic variables such as GDP and population to forecast imports and exports of a country to construct O/D-matrices. This model is also called an "agent-based simulation model", which means countries are seen as autonomous agents which have their own variables and behaviour. So, first import and export forecasts are made per individual country before they interact with each other to determine trade flows. When the trade flows are determined based on historical data, a trend model is used to forecast flows into the future. After that, as a final step, the future trade flows are assigned to the network based on WORLDNET data (CFCU, 2009).

Like many forecast studies, this study for the port of Filyos also makes use of different scenarios. Because a Greenfield port is considered, there are many varying scenarios possible. Four scenarios are developed: reference, sensitivity low, sensitivity high and do-nothing. Firstly, these outcome scenarios are based on two economic scenarios, middle and low economic growth. It should be noted by times of writing of the report, the world was at its lowest point in the economic crisis. So, the middle growth scenario predicts recovery of the economy in 2010 while the low growth sets 2012 as the starting year of recovery. Second, investments made in the hinterland of the port are also of importance for the scenarios. The writers of the forecast report on Filyos have made a distinction between types of investments. The first are certain investments of influence that are officially agreed and clear, for example investments in a second power plant in Catalagzi (nearby area). Second, uncertain investments of influence are mentioned which are investments that are discussed and for which the finances are not yet complete, for example a steel factory near the port. Third, there are required investments that are not certain but are essential for a good functioning of the port, for example connections between the port and the road and rail network. These three kinds of investments are assigned to the scenarios to be calculated. Following this process, four scenarios were developed that are clarified in the table below. The Middle and High parameters that are mentioned refer to the level of investments.

	Filyos port		No Filyos port
	Certain investments, Middle parameters	**Certain and uncertain investments, High parameters**	**Certain investments, Middle parameters**
Middle economic growth	Reference scenario	Sensitivity high scenario	Do nothing scenario
Low economic growth	Sensitivity low scenario	-	-

Table 7: Scenarios for the port of Filyos. Source: CFCU (2009).

To not fully ignore competition between ports, the writers of the report made an analysis of the cost structures. All possible routes per OD pair are compared in terms of generalized costs and the trade flow is assigned to the cheapest route. This is approach is similar to the multiregional I/O model described in chapter 2. It is especially important for container traffic to take into account generalized costs, because these ships are more flexible compared to bulk carriers when it comes to route choice (CFCU, 2009). A disadvantage to this approach is that it does not include trade-offs between costs and quality of service (Veldman & Bückmann, 2003). The forecast for Filyos would therefore benefit from including a port competition/choice model. However, many port characteristics are not yet known for sure and thus such a model would depend too much on speculation. That makes its reliability questionable.

Overall, the forecasting method used for the port of Filyos seems appropriate. Use is made of a regional model to determine trade flows and assign these flows to the possible routes according to generalized costs. Hence, macro-economic factors (GDP and population) are taken into account for determining trade flows and competition factors are included by assigning to routes based on generalized costs. Also, the essential qualitative analysis is used to fill in the gaps of the quantitative variables. Unfortunately, because the port of Filyos is not yet constructed, no analysis can be made on the fit of the forecasts with the realized throughput volumes.

4.2.2 Port of Rotterdam

For Brownfield ports it does not necessarily have to be the case that a forecast is made for a specific expansion project. The port of Rotterdam made a so called Port Compass for 2030 in the year 2011. Herein, forecasts are made for the future by use of different scenarios. Each scenario has another effect on cargo throughput in the port. Based on the forecasts that are made on cargo throughput volumes, policy goals for the future are formulated.

The starting point for making the forecasts was developing the scenarios. The port of Rotterdam identifies four main drivers that affect commodity flows: economic growth, global trade, fuel price and environmental policy. These four drivers serve as the basis for the development of the four scenarios that are used for making a forecast for 2030:

- **Low Growth**
  Low economic growth and a low fuel price. Fossil fuels stay dominant and environmental policy is moderate.
- **European Trend**
  Continuing existing policy and a moderate growth of the economy.
- **Global Economy**
  Further globalisation combined with low fuel price causing high economic growth and moderate environmental policy.
- **High Oil Price**
  High oil price, a strict environmental policy, moderate economic growth and a relatively fast move to sustainable industry and logistics.

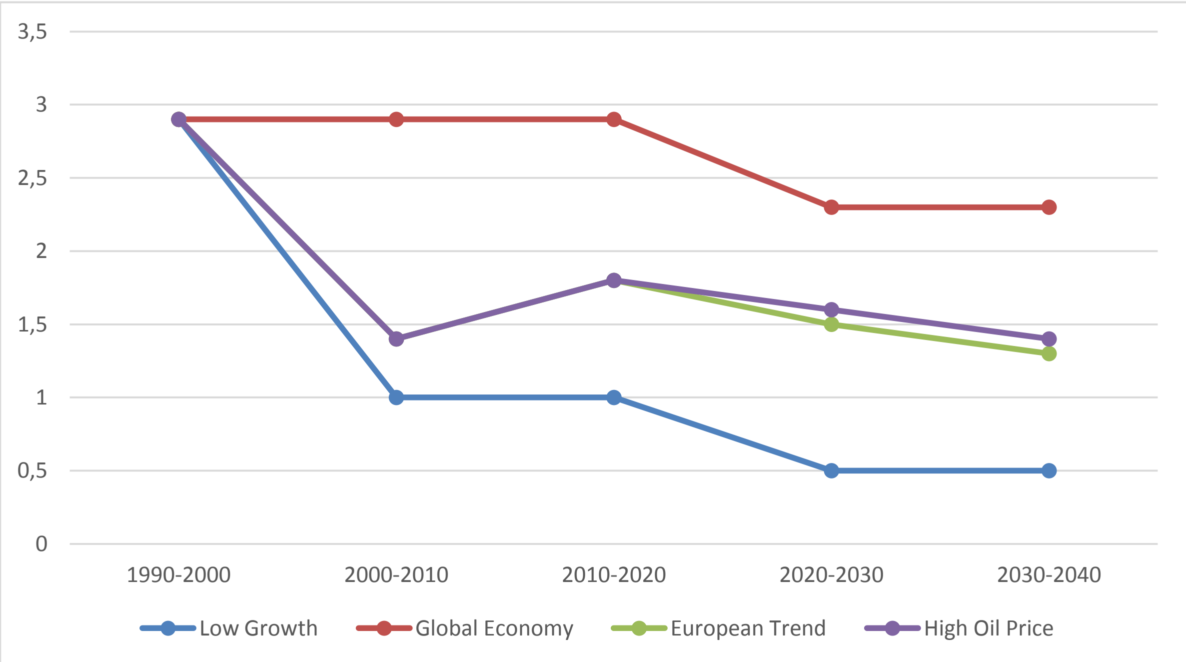

Figure 14: GDP growth per 10 year period, per scenario. Based on HBR (2011).

These scenarios are not developed by the Port of Rotterdam itself, but by Dutch and European governmental agencies to serve European policy studies. For all scenarios GDP growth for 10 year periods are determined, resulting in figure 14.

The GDP growth rates are used to calculate the forecasting outcomes by use of the Transtools model. A model that translates socio-economic developments in the transport industry, like the scenarios described above, into commodity flows. Based on the steps described in chapter 2, the first step that the model uses is a gravity model to determine the distribution of trade flows over the world. In the second step, mode choice is taken into account using an aggregate multinomial logit model based on transport times and tariffs. Thirdly, network assignment takes place. This is done by translating freight flows in tonnes to number of transport units and assigning those units to the infrastructure network. This results in information about the intensities of traffic, vehicle type and fuel use (De Langen et al., 2012a).

The results of the Transtools model were tested against segment specific knowledge of experts in the field. The eventual outcomes are thus a mix between quantitative outcomes of the Transtools model and expert knowledge. For the specific case of container transport, the Transtools model considers transhipment markets (UK, Ireland & ScanBaltic) and containerisation rates. However, the model insufficiently handles industry specific developments and cannot make a proper forecast for individual ports. This makes it necessary to modify the model results with expert judgment.

Port specific characteristics were also reviewed, be it qualitatively. The method would benefit from quantifying the competitive position of the Port of Rotterdam.

Because the port of Rotterdam is a Brownfield port, it is possible to compare the forecasts with the actual throughput volumes. The port of Rotterdam is constantly doing this to see which of the four trends based on the scenarios is followed. Now, it appears the port is following almost exactly the Low Growth scenario in terms of cargo throughput volumes (A. Willeumier, personal communication, June 25, 2014; HBR, 2013).

4.2.3 Comparison Filyos with Rotterdam

Fundamentally, the two forecasting approaches of the port of Rotterdam and Filyos do not differ. They both make use of different scenarios, a gravity approach to determine trade flows and a regional network approach to determine route choice. There are however some points in which the approaches differ from each other.

The ways in which the scenarios are built up are different. The drivers behind the scenarios for the port of Rotterdam have a global scope and depend on policy issues and growth of the economy and global trade. For the port of Filyos the scenarios depend also on economic growth rates, but also on investments in the port. The latter is a driver with a more narrow view (regional vs. global) for throughput volumes in the future. Maybe this is the difference between forecasts for Greenfield and Brownfield ports. For a Greenfield port, investments are essential to start up and expand while a Brownfield port is less dependent because it is running already. However, it should be noted that it is possible this difference is only present for this particular case, because the port of Filyos has a more regional focus than the port of Rotterdam. More important, the objective of the reports are different, as that for the port of Filyos concerns a port master plan study while that for the port of Rotterdam is just a forecast to give an idea of what the future may bring.

The second difference between the approaches of the two ports is the way they deal with port specific factors and competition. In the approach used for the port of Filyos, port choice is included in the model by assigning the trade flows to the routes with the lowest generalized costs. This is an overall accepted way to deal with competition between ports (Tavasszy et al., 2009). Thus, for the port of Filyos the competition between ports is quantified, using the location of the competing ports, and can be verified. For the port of Rotterdam on the other hand, the competitive position of the port is assessed qualitatively and the market share of the port in the Hamburg-Le Havre range is estimated using the throughput volumes for the whole range. Transtools predicts a total throughput for the Hamburg-Le Havre range of 1101 million TEU in 2030. The results are modified to 815 million TEU, using expert judgement. This is a change of -35% (De Langen et al., 2012a). No exact numbers are given for the predicted market share of Rotterdam, because the analysis is qualitative. But based on the figures for 2008 and the construction of Maasvlakte II it is assumed to be around 30%. Which means Transtools predicted 330.3 million TEU and the modified result is 244.5 million TEU (HBR, 2011). Although the rationale behind the market share is clear, it should be preferred to incorporate the competitive position of the port in a model, because this it is possible for Brownfield ports to use the WCM or PCM.

4.3 Developing vs. developed economy

A country in the third world is not as well developed economically as for example North-Western countries. This makes that trade flows to and from these countries, and thus their transport facilities like seaports, are also underdeveloped. This can have the effect that administration in such countries is less up to date. This makes it more difficult to prepare an O/D-table, because proper data is missing. So it should be kept in mind that models with less data requirements are preferred when a forecast is made for a port in an underdeveloped economy.

Also, the fact that a country is underdeveloped can affect the way things are handled in a port. In general, operations are less efficient in developing countries than in developed economies. This has to do with the fact that services in developing economies are in an earlier stage of evolvement and should get more efficient on the way. In general, cargo spends more time in a port in a developing country than elsewhere, due to long ship turnaround and waiting times, a high berth occupancy rate, long working times at berths, long cargo dwell time and a small number of crane moves per hour (UNCTAD, 2013). The fact that these ports are less efficient, for example in terms of waiting times, causes the transport cost to rise when such a port is called. This affects the competitive position of the port in a negative way. In order to improve the efficiency of ports in developing countries, it is important that these ports compare themselves with other ports in the region rather than with how they performed in the past. Many ports in developing economies are the main port of their country, which makes comparisons hard to execute (UNCTAD, 2013).

However, the fact that a less developed country is in an earlier stage of development than a well-developed country also means there is more growth potential in this country. In the literature, many researchers state that the expansion of exports can give a boost to economic growth in a country (Chou, Chu & Liang, 2007). This effect is assumed to be stronger for well-developed countries than for developing countries. However, Ram (1985) found evidence that later in time, when the world was further in its overall development, no difference between the positive effects of export expansion between developing and developed countries is present. It should be noted that research on this topic is very old, which could affect its relevance.

When comparing trade flows to and from developing countries with that of developed countries, it should be noted that in the past years the flows for developing countries show a much more positive growth figure. For the head-haul Asia-Europe volumes dropped by 2.6% in 2012, due to the decrease of European import volumes. This can be an effect of near-sourcing, but also of the low level of purchase power in Europe. On the other hand, North-South trade expanded by 3.9%. This difference is presumably caused by economic growth in the South and a shift of economic influence away from the traditional centres of economic growth. Also, in developed countries the population is ageing while the population in developing economies is growing fast, affecting global consumption patterns (UNCTAD, 2013). It is expected that Southern markets will continue to drive growth of containerized trade in the future (Clarkson Research Services, 2013). In line with this, the throughput volumes in ports situated in developing countries faced less negative effects of the crisis in 2008/2009 than ports in developed countries.

First, a port in a developing country is considered: the port of Dar es Salaam in Tanzania. Throughput volumes were found for the years 2007-2012. There was a lack of data on developments that caused volatilities in container throughput volumes, so the analysis is made based on economic logic. Next, Copenhagen Malmö port is described as a port based in a country with a well-developed economy. To make a fair comparison between the ports, data for 2007-2012 was analysed only. The section ends with a short comparison between the development of throughput volumes in both ports. For the comparison between a port in a developed economy and a port in a developing economy, economic growth is considered the main factor to be considered, because it is the main factor in which they are distinguished. Therefore, GDP growth percentages are also considered in paragraph 3.3.3.

4.3.1 The port of Dar es Salaam, Tanzania

The port of Dar es Salaam is the principal port of Tanzania. It handles around 90-95% of all cargo handled in the country. The market share deviates from year to year. The port predominantly handles bulk cargo, 35% of all cargo handled are containers. There are 2 terminals in the port where containers are handled: a container terminal that is operated by a private organisation and a general cargo terminal.

Figure 15: The port of Dar es Salaam and its hinterland. Source: dlca.logcluster.org.

When no distinction is made between types of cargo, throughput volumes are dominated by imports, which indicates that Tanzania is a less developed country (Chou et al., 2007). The high import rates are partly due to the fact that Tanzania serves as a gateway for several land-locked countries, i.e. Zambia, Democratic Republic of Congo, Malawi, Rwanda and Burundi. Transits of freight to these countries account for approximately 35% of total traffic. This share should be greater, given the strategic location of the port to serve as a transport hub. However, hinterland connections are poor and most of the countries in the hinterland are politically unstable. This is considered typical for a developing country and although freight for the hinterland does not make up the greatest part of total port throughputs it significantly affects the performance of the port of Dar es Salaam.

Another line of reasoning to explain why imports are higher than exports, especially in combination with the high amount of bulk cargo, is that Tanzania does not have a lot of resources, both human and natural. This can be one of the reasons why the country is still developing economically.

For containers imports and exports do not differ as much as for all cargo together, although imports are higher than exports (except for 2008). Only direct calls of container vessels are handled in the port of Dar es Salaam. In figure 14 the development of container throughput in the period 2007-2012 is given.

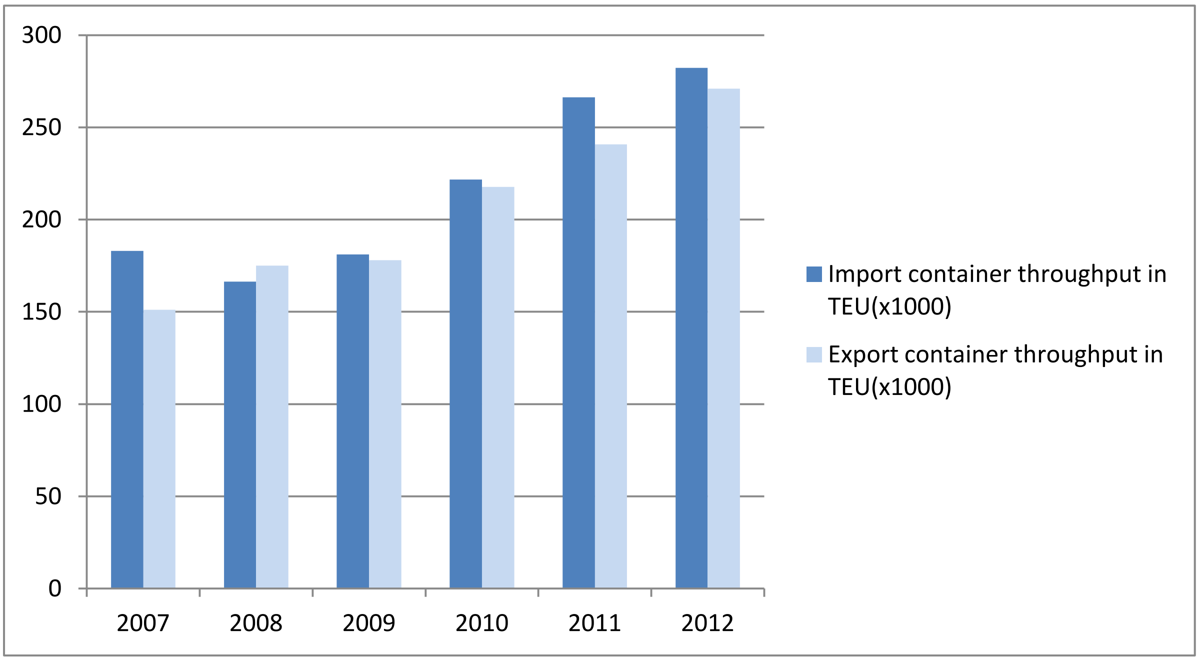

Figure 16: Development of container throughput in the port of Dar es Salaam. Source: Tanzania Port Authority.

Over the years, the proportion of imports and exports is more or less stable. Only in 2008 the exports exceed the imports. In this year, also a small decrease in throughput volumes is observed. This is due to the fact that less dry cargo was containerized in 2008 compared to the previous year. Which is remarkable as the main trend is that more and more cargo is containerized.

Throughput volumes are incessantly increasing in the years under consideration, indicating the port is able to attract more container traffic. Relative to 2010, there was an increase of 21% in container traffic calling the port in 2011. Also, the market trend of bigger vessels is affecting the port of Dar es Salaam in a positive way. At first, the port had to adjust to this movement, as stated in the annual report for 2009. It seems like this is done well, according to the continuing growth in container throughput volumes.

Although the port of Dar es Salaam faces stiff competition with ports in the region, especially for the markets of Zambia and Malawi in terms of distance and costs, this does not affect throughput volumes of containers as transit products are mainly bulk commodities.

4.3.2 Copenhagen Malmö Port

The port situated in a well-developed country that is analysed is the Copenhagen Malmö Port (CMP). This port is situated partially in Copenhagen, Denmark and partially in Malmö, Sweden. Denmark and Sweden are both well developed countries, which makes this port suitable to compare it with its opposite: the port of Dar es Salaam.

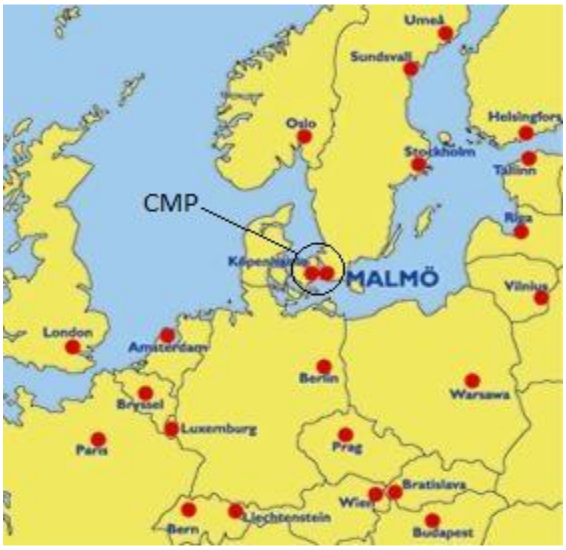

Figure 17: The location of the Copenhagen Malmö Port. Based on: www.greenroofs.com.

Container turnover in CMP mainly consists of imported consumer goods destined for the region. This is one of the reasons why throughput volumes fell heavily with 24% after 2008, the start of the credit crisis. Consumer goods are less cyclical than industrial products and consumer confidence declined, causing them to purchase less products. CMP responded to the decline by changing its opening hours to become more flexible and lowering operating costs in order to become more efficient and attract new shipping lines. These new customers, CMA CMG and Teamlines, provided the port a marginal increase in volumes in 2010. However, the volumes were disappointing for CMP because the economy was on its return. Unfortunately, growing demand for consumer goods did not yet pick up sufficient momentum at that time, forcing the port to cut its staff.

In 2011, throughput volumes stayed more or less the same. Although there was an increase of activity in Copenhagen by 6%, caused by higher productivity during vessel operations and the return of some of the greatest shipping lines in the world, the volumes for Malmö are unsatisfactory. This is probably due to the relocation of the container terminal to the northern harbour. This new container terminal resulted in overcapacity, responsible for the small drop in container volumes in 2012 as developments in Copenhagen remained stable (CMP, 2012).

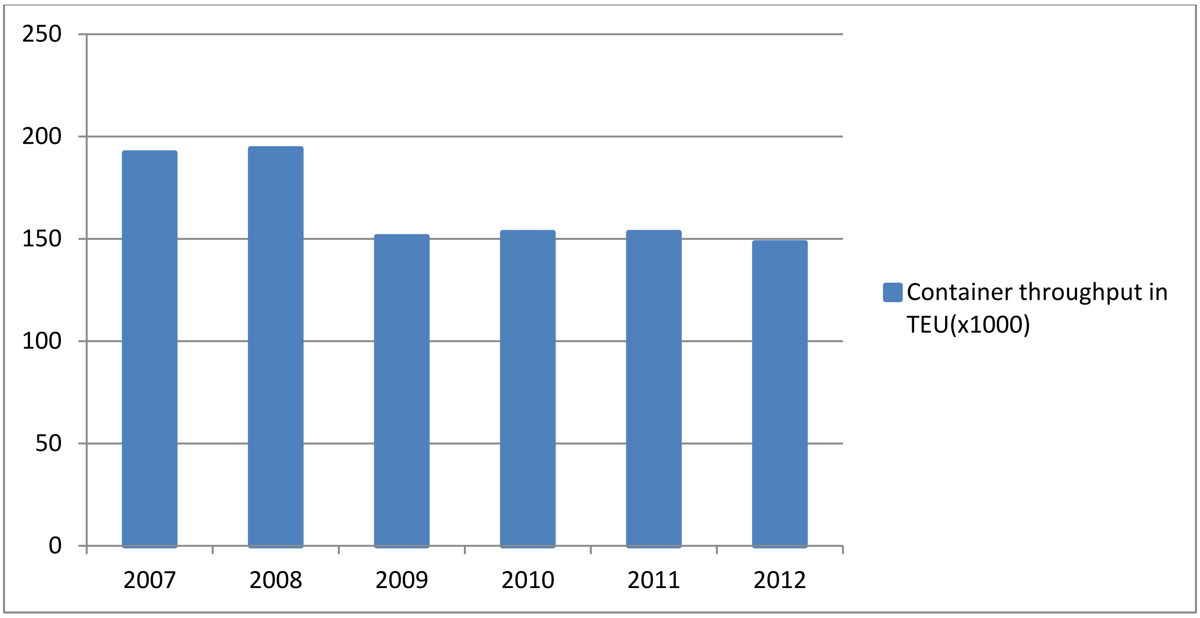

Figure 18: Container throughput volumes in Copenhagen Malmö Port. Source: CMP.

4.3.3 Comparison Dar es Salaam with Copenhagen-Malmö

In figure 19 the economic growth in Tanzania, Denmark and Sweden over the period 2007-2012 is given. When this figure is compared with figure 16 and 18, one can see that the effect of the financial crisis that started in 2008 seems to be absent in the port of Dar es Salaam. At most, a small drop of economic growth by 1.5% is caused by the credit crisis. This is reflected in the stagnating growth of throughput volumes in the port, which perfectly demonstrates the relation between GDP and port throughput. For CMP, however, the effects were evident as throughput growth dropped with 24% in 2009 and were not recovering before 2013. Economic growth also drops heavily in Sweden and Denmark after 2008, even below zero and does not recover within the analysed period. Although there was a fast increase in 2010, growth dropped again after this year. So for these countries there is also a demonstrable relationship between GDP growth and port throughputs.

The difference between the effects for Dar es Salaam and CMP are probably caused by the fact that in CMP, the content of containers handled mainly consists of consumer goods. Consumer confidence was low during the worldwide economic crisis, causing purchasing power to decline. Return to higher consumer confidence to boost the economy moves slowly.

For the port of Dar es Salaam it is not perfectly clear why throughput volumes did not decline as an effect of the crisis. One explanation that is derived from the annual reports of the port, is the appreciation of the Rand of South Africa, the major competitor of Dar es Salaam when it comes to serving Zambia. This causes the route via the Durban port of South Africa to be more expensive, resulting in more demand for transits via Dar es Salaam.

Figure 19: Percentage GDP growth for Sweden, Denmark and Tanzania in 2007-2012. Based on: IMF (2012).

Another explanation is that the content of the containers differ between the ports. It is not clear what is the content of the containers handled in Dar es Salaam, but the containers in CMP contain consumer products destined for the regional market. The containers handled in Tanzania can also be destined for their expansive hinterland. Also, it can be reasoned that the economies of African countries mainly depend on natural resources while the economies of European countries depend more on the service industry and consumer goods. The latter markets are more volatile and have an effect on the economy quicker. For example, one of Zambia's core businesses is the exportation of copper. Demand for copper increased in the past decade, causing the price to rise. This is a boost for the Zambian economy and transport over Dar es Salaam is intensified. Accompanied with the development of the Zambian economy, demand for commodities transported in containers can go up, causing throughput volumes in the port of Dar es Salaam to rise.

When it comes to import and port of call, CMP probably faces heavier competition than the port of Dar es Salaam. Containers destined for Denmark or Sweden can also be transported via another port in the region, because hinterland connections in that part of Europe are well developed in general. Hence, costs and time will not necessarily increase by choosing a port that is further away from the region of destination. For the containers that have to be transported to Tanzania there are not as much options as for Denmark and Sweden. First of all, the country is a lot bigger and thus the ports are located further away from each other. Hence, differences in speed and costs will be bigger for Tanzania than for Northern Europe. Second, the port of Dar es Salaam has the biggest market share in Tanzania (approximately 90%) and it is therefore assumed that it has the best hinterland connections, because hinterland connections in developing countries are very poor in general. This means that Dar es Salaam is the optimal port of choice in terms of costs and time. Therefore, shipping lines do not have the same amount of options in developing regions as in well-developed regions causing them to change their port of call less often.

4.4 Findings and implications for each port type

In this concluding paragraph of the chapter, the findings of the case studies per port category are discussed and the implications for the form of the models are mentioned. An abstract and holistic approach is used to describe these implications, leaving the exact formulation of the models to further research.

4.4.1 Transhipment port

For a pure transhipment port, the gravity model is less important than the model wherein port characteristics are used to predict the distribution of trade flows. Of course, it is essential to have an idea of how container flows are distributed over the world and therefore the gravity model should be used. It is however not important to zoom in on a specific OD-pair.

For a transhipment port it should be determined how far the port is located from the busiest trading routes over sea and how many of the vessels using the route it can attract. Thus, a port choice/competition model should be used where throughput volumes depend on location relative to the trade route the port wants to capture (Zohil & Prijon, 1999). Also, the level of service at a transhipment port is very important as it determines whether a shipping line makes the decision to call at the port (Schinas & Papadimitriou, 2001). This is especially a driver of competition for ports that are located approximately at the same distance from the trade route.

The model does not have to focus on the hinterland connections of the port as these are not of importance for its competitive position. Therefore, the WCM is the best port choice model to use in this case. In case the WCM is not available for use, the generalized cost approach can be used, for which not specific model is required. Trade flows are then assigned to the cheapest route. The disadvantage hereof is that it does not account for the trade-off between quality and costs.

4.4.2 Import-export port

An import-export port is very dependent on its hinterland, which is in contrast with the transhipment port. Therefore, the emphasis should be on the gravity model. This can be done by including as many variables as possible into the gravity model. For the specific case of Auckland, it was noticed the imports depended on the growth of the population in the hinterland. In general, this is just one of the many macroeconomic variables that affect imports and exports, like the industrial production index, GDP, GNP, etc. (Chou et al., 2008).

For import-export ports, the strategic decisions of shipping lines are also very important even as the contestability of the hinterland, so competition on service levels and connectivity should definitely be taken into account (Polo & Guittérez, 2006). The PCM should be used here, because this model focusses more on the hinterland than on the maritime side of the transport route. If it is not possible to make use of the PCM the generalized cost approach can be used.

4.4.3 Greenfield port

A forecast for a greenfield port should start with a gravity model, like all forecasts. The input for this gravity model should be arrived from several scenarios that predict GDP values. These scenarios should be based on predicted economic growth, policy measures or local investments, according to the scope of the port.

Port competition cannot be left out in the analysis of future volumes, but it is hard to include in a detailed model. It is therefore proposed that, although it does not account for the trade-off between quality and cost, generalized costs should be used to allocate routes via certain ports. This is a generally accepted way to deal with competition in case of greenfield ports (Cullinane, Teng & Wang, 2005). It is also possible to obtain O/D tables that are developed by public organizations. These matrices provide information on good flows between origins and destinations. The gravity model is used to compute relations within the tables. Overall, the functional form of the gravity model can be very simple making it easy to use.

4.4.4 Brownfield port

The forecasting for a brownfield port should not differ much from the method of a Greenfield port. The difference is that more detailed information can be used for a Brownfield port, which increases the explanatory power of the models. This is especially the case for the variables in the port competition or choice model. Data on the market share of the port is (in general) available and should be used in the forecast method. The same applies to variables as specialization of the port, the amount and quality of hinterland connections, service level, etc. Both the WCM and the PCM can be used to make a forecast for Brownfield ports.

4.4.5 Port in a less developed country

For ports in a less developed country the economic situation and political stability in the hinterland is very important (Hoyle & Charlier, 1995; TPA, 2011). Therefore, the gravity model should be complemented with variables that account for these characteristics to increase explanatory power.

As it comes to port competition in the region, there are many (standard) variables that are of importance to determine the market share of the port of interest. These are hinterland connections, level of service, congestion at the port, turn-around times, etc. A first principle on which the level of development of a seaport is based, is the interconnection with its hinterland (Charlier, 1983). It is assumed hinterland connections in developing countries are poor in general, which gives the few ports that do have proper connections a high market share that is not lost to other ports in the country easily. Hence, location is also very important referring to ports that have a natural inland connection in the form of a river. Because the focus for this kind of port is on the hinterland, the PCM is preferred, though the WCM is also suitable in this case.

4.4.6 Port in a well-developed country

It is assumed that where in developing countries hinterland connections are very poor in general, in developed countries there are more options for bringing the freight from A to B without losing time, money or service (Hoyle & Charlier, 1995). This makes ports in developed countries more dependent on the strategic decisions of shipping lines as they can deviate more easily. Unfortunately, strategic decisions of shipping lines cannot be included in the model. The focus should be on the PCM part for ports in developed countries, because regional competition is very fierce.

5 Conclusions and recommendations for further research

In this last chapter first an overview is given on the findings presented in this paper. Next, the implications of these findings for Witteveen+Bos are given and a recommendation is made for a suitable approach. The chapter concludes with listing the shortcomings of this research and recommendations for further research on the subject.

5.1 Overview of the results and conclusions

Forecasting methods nowadays build on the four step approach derived from the passenger transport literature. These are production & attraction, distribution, modal split and assignment to the network. For the first step, I/O models or O/D tables are the most suitable solution. These matrices show per region (or sector in case of I/O) which products are produced and which products are attracted or demanded. Strongly related to this is the distribution step in which regions of supply and demand are linked. In this way, the distribution of trade flows arises. The model that is usually used for this is the gravity model, ascended from physics. Big economies attract many trade flows, hence there is many trading going on between the richest countries in the world. There is however also the resistance of trade, for example physical distance or transport cost, causing a reduction in potential trade flows. Often, assignment to a mode and a certain route are done simultaneously. The best way to do this is by use of a multimodal network model. These are mainly multinomial logit models that deal with discrete choice models.

According to this research the essential elements of a good container cargo forecast are threefold. First, a good forecast has to be made on the level of GDP in the future. This can be done based on the forecasts of the IMF or other public organizations that are committed to making good GDP forecasts. Often, scenarios are used to give an idea of the development of GDP based on factors like oil price, environmental policy, economic growth etc. It is useful to use scenarios when making long term forecasts, because in that way different trends are considered and assumed possible. This lowers uncertainty, because when the actual volumes deviate from the forecasted ones there is always another forecasted trend that is close to the real values. These GDP forecasts are used as input for the gravity model that estimates the distribution of trade flows. So in this first step towards a good cargo forecast, macroeconomic variables are taken into account. These together form the first essential element of container cargo forecasting.

Second, port competition should be taken into account to come to an assignment of the trade flows over the port network. This can be done by including port characteristics in a multinomial logit model, like the World Container Model (WCM) and the Port Competition Model (PCM). These are the models that are proposed to use in order to account for port competition. The WCM focusses more on the maritime network and used the value of time to incorporate the trade-off between costs and transport time. The PCM focussed more on the hinterland connections of the port and leaves out the maritime side of the transport route. Here also the trade-off between several important variables are used to assign freight to a certain route.

Third, there are always factors that cannot be quantified or databases that are incomplete. Qualitative assessments are needed to account for effects that cannot be quantified or to fill in gaps in databases. It is proposed to use expert knowledge to analyse the results from the quantitative models. According to this knowledge the results should be modified or the input variables can be modified to come to a better forecasting result.

Unfortunately, from the analysis of the throughput volume development in different kinds of ports, it becomes clear that all ports are very dependent on the strategic decisions of shipping lines. These decisions cannot be predicted and therefore cannot be accounted for in the forecasting approach. This means there is always a certain level of uncertainty that cannot be removed.

The analysis of the different ports, namely transhipment, import-export, Greenfield, Brownfield, less-developed and well-developed economy, also made clear that for each type of port the focus should be on different variables. For the transhipment port, the hinterland is not important and thus the focus should be on competitive characteristics like the location relative to the main liner routes and the level of service compared to other ports in the region. In contrast, for an import-export port, the hinterland is very important and the focus should be on macroeconomic variables like GDP, population and income per capita. For a port in a well-developed economy competition is assumed to be fiercer than for a port in a less-developed economy. Therefore, the PCM is very important for a port in a well-developed country as hinterland connections are very important for the competitive position of a port. For a port in a developing economy the hinterland is also very important, but the focus is more on macroeconomic variables like economic growth and political stability. These can be included in the gravity model.

For the analysis of the case studies for a Greenfield and a Brownfield port, the forecasting approaches of two ports are used. Therefore, the analysis was somewhat different than for the other port categories. The approaches did not differ fundamentally. However, the scope of the scenarios used were different: for the Greenfield port more regional and for the Brownfield port more global. This result cannot be generalized as this can also be a result of the different scopes of the ports.

The purpose of this thesis was to shed light on the different factors that determine port throughput. Models used in practice are inseparably linked to these factors as the identified variables serve as input for these forecasting approaches. An overview of the models that can be used is given and the essential elements of a good forecasting method are described. This should all serve as a solid basis to build on for further research. The main objective for further research should be to develop a model that accounts for all the essential elements mentioned in this thesis. Based on this a software package can be developed to give people the possibility to make use of an integrated forecasting model that is ready to use and does not require specific modelling knowledge. Software and models are available now, like Transtools, but these are not straight forward and often extra modules are necessary to run the model properly. Also, it is not perfectly clear what factors are included in the model, making the result useless. Hopefully, this thesis contributes to the freight transport modelling in that it reveals which factors are necessary to include.

5.2 Advice for Witteveen+Bos

Whilst the models recommended in this thesis seem to work in theory, it is necessary to come up with a more practical advice for Witteveen+Bos. It is possible for the consultancy firm to use the gravity model, the WCM and the PCM. However, datasets are not always available and the gathering of data can be labour-intensive. If the formulation of the gravity model is simplified, it should be feasible to use it for any organisation or person. That is, when trade flows depend on the GDP in the origin and destination and trade resistance is formulated as the distance between countries. A disadvantage of this is that the explanatory power of the model decreases. One of the developers of the WCM, L. Tavasszy, claims that the WCM is useable for everyone and that datasets are publically available (L. Tavasszy, personal communication, July 11, 2014). The issue here is whether the datasets are up to date and ready to use, which should be verified before use. In this thesis, an old version of the PCM is presented that is publically available. The model is developed by the company Ecorys and is adjusted several times after 2003. The improved version of the model is not available for free and should be bought at Ecorys. A method that requires less data, but does not account for the trade-off between costs and quality is to assign flows to the network directly according to generalized cost. This is also done for the port of Filyos and requires less data.

It is also possible to use a software package like Transtools, which is the most discussed software in the literature. This model runs in ArcGis which is available at Witteveen+Bos. However, several modules have to be adjusted to the model that is available for free online. Also, it is not clear which aspects the model covers which makes it a "blackbox". Hence, specific knowledge of the software is needed, which makes this not the preferable option.

When it is preferred to make a forecast without a model, it has to be based on the factors that seem to be important in predicting throughput volumes. From the case studies in chapter 4, several relations between drivers and cargo growth become clear.

- GDP growth is a very solid indicator for cargo growth and they move in the same direction.
- Population and income growth are important indicators for the development of import volumes. They also move in the same direction.
- Level of service, costs and connectivity are important variables for transhipment ports. The higher the level of service, the lower the costs and the better the connectivity (or the shorter the time of deviation from the main routes), the higher transhipment volumes.
- Hinterland connections are important for every port except the transhipment port. Generalized costs for the route from port to final destination can be used to qualify the connections.

It is also possible to come up with an advice that is based on the developments in these drivers of changes in trade flows and throughput volumes in ports. The results are however not verifiable, because they are not tested statistically.

5.3 Recommendations for further research

There are several shortcomings of this research. One is that case studies of only two opposing ports are considered for each port category. Therefore, findings cannot be generalized as there is always the probability that they correspond with port specific element instead of the whole category. Further and more extensive research on several kind of ports and their implications on the forecasting method would contribute to the freight modelling literature.

Second, because developing a model that is ready to use does not fit into the scope of this thesis, it cannot be tested if the recommendations made on the elements of the model are grounded. Further research is necessary to test whether these elements are indeed essential.

References

Baird, A. J. (2006). Optimising the container transhipment hub location in northern Europe. *Journal of transport geography*, *14*(3), 195-214.

Bankes, S. (1993). Exploratory modeling for policy analysis. *Operations Research, 41*(3), 435-449.

Central Finance and Contracts Unit (CFCU) (2009). *Technical assistance for construction of new port in Filyos.*

Charlier (1983). Ports et regions françaises: line analyse macro-géographique. *Acta Geographica Lovaniensia. 24*, 1-198.

Chou, C. C., Chu, C. W., & Liang, G. S. (2008). A modified regression model for forecasting the volumes of Taiwan's import containers. *Mathematical and Computer Modelling*, *47*(9), 797-807.

Clarkson Research Services (2013). *Container Intelligence Monthly. 15*(6).

Copenhagen Malmö Port (CMP) (2007,2008,2009,2010,2011,2012). *Annual report.*

Coto-Millán, P., Baños-Pino, J., & Castro, J. V. (2005). Determinants of the demand for maritime imports and exports. *Transportation Research Part E: Logistics and Transportation Review, 41*(4), 357-372.

Cullinane, K., Teng, Y., & Wang, T. F. (2005). Port competition between Shanghai and Ningbo. *Maritime Policy & Management*, *32*(4), 331-346.

De Groot, H. L., Linders, G. J., Rietveld, P., & Subramanian, U. (2003). The institutional determinants of bilateral trade patterns.

De Jong, G., Gunn, H., & Walker, W. (2004). National and international freight transport models: An overview and ideas for future development. *Transport Reviews, 24*(1), 103-124.

De Langen, P.W., Nijdam, M.H. & Van der Lugt, L.M. (2012a). *Port Economics, Policy and Management* [Syllabus]. Erasmus Universiteit Rotterdam, Erasmus School of Economics.

De Langen, P. W., Van Meijeren, J., & Tavasszy, L. A. (2012b). Combining Models and Commodity Chain Research for Making Long-Term Projections of Port Throughput: an Application to the Hamburg-Le Havre Range. *European Journal of Transport & Infrastructure Research, 12*(3).

Dikos, G., Marcus, H. S., Papadatos, M. P., & Papakonstantinou, V. (2006). Niver lines: a system-dynamics approach to tanker freight modeling. *Interfaces, 36*(4), 326-341.

Donders, L. B. (2010). Layout design for greenfield port Filyos (Master's thesis, Delft University of Technology, The Netherlands).

Europe, R., & Brinckerhoff, P. (2002). Review of Freight Modelling. *Report B2–Review of Models in Continental Europe and Elsewhere, DfT*.

Garatt, M. (2006). *Forecasting for long term investment in the container shipping industry* [Powerpoint-presentation]. Retrieved from www.mdst.co.uk

Gosasang, V., Chandraprakaikul, W., & Kiattisin, S. (2010, June). An Application of Neural Networks for Forecasting Container Throughput at Bangkok Port. In *Proceedings of the World Congress on Engineering* (Vol. 1).

Gouvernal, E., Debrie, J., & Slack, B. (2005). Dynamics of change in the port system of the western Mediterranean. *Maritime Policy & Management*, *32*(2), 107-121.

Hausman, W. H., Lee, H. L., & Subramanian, U. (2005). Global logistics indicators, supply chain metrics, and bilateral trade patterns.

Havenga, J. H., & Van Eeden, J. (2011). Forecasting South African containers for international trade: A commodity-based approach. *Journal of Transport and Supply Chain Management, 5*(1), 170-185.

Havenbedrijf Rotterdam (HBR) (2011). Havenvisie 2030 Port Compass. *Direct the future. Start today. Rotterdam: HBR*.

Havenbedrijf Rotterdam (HBR) (2013). Progress report 2013 Port Compass 2030. *Direct the future. Start today. Rotterdam: HBR*

Hoffmann, J. (2012). Corridors of the sea: An investigation into liner shipping connectivity. *Les Corridors de Transport. Les océanides.*

Hoyle, B., & Charlier, J. (1995). Inter-port competition in developing countries: an East African case study. *Journal of Transport Geography*, *3*(2), 87-103.

Hui, E. C., Seabrooke, W., & Wong, G. K. (2004). Forecasting cargo throughput for the port of Hong Kong: error correction model approach. *Journal of urban planning and development, 130*(4), 195-203.

Hummels, D., & Schaur, G. (2012). *Time as a trade barrier* (No. w17758). National Bureau of Economic Research.

Lam, W. H., Ng, P. L., Seabrooke, W., & Hui, E. C. (2004). Forecasts and reliability analysis of port cargo throughput in Hong Kong. *Journal of Urban Planning and Development, 130*(3), 133-144.

Lam, L. and Iskounen, A. (2010). *Feeder ports, Inland ports and Corridors - Time for a closer look.* Retrieved on 8th of Augustus 2014 from www.portek.com.

Notteboom, T., Parola, F., Satta, G. (2014) *State of the European port system - market trends and structure update. Partim transshipment volumes* (Report No. 605176). Retrieved from European Commission website: ec.europa.eu.

Ortúzar, J. D., & Willumsen, L. G. (1994). *Modelling transport* (4th ed.). London, England: Wiley.

Polder, M. (2000). Forecasting international trade flows: A gravity-based approach. *Econometrische Toepassingen*, 24-28.

Polo, G., & Gutiérrez, D. D. (2006). A new generation of containerships: cause or effect of the economic development?. *Journal of maritime research: JMR,3*(2), 3-18.

Port of Salalah (2004,2006,2008,2010,2011,2012). *Annual report.*

Ports of Auckland (2008,2009,2010,2011,2012,2013). *Annual report.*

Ram, R. (1985). Exports and economic growth: Some additional evidence. *Economic Development and Cultural Change*, 415-425.

Rodrigue, J. P., Comtois, C., & Slack, B. (2013). *The geography of transport systems*. Routledge.

Schinas, O., & Papadimitriou, S. (2001). The Mediterranean ports in the era of mega-carriers: a strategic approach. *NAVIGATION-TOKYO-*, 45-64.

Seabrooke, W., Hui, E., Lam, W. H., & Wong, G. K. (2003). Forecasting cargo growth and regional role of the port of Hong Kong. *Cities, 20*(1), 51-64.

Tanzania Port Authority (TPA) (2006,2009,2011). *Annual report.*

Tavasszy, L. (2006). Freight Modeling. *Freight Demand Modeling*, 47.

Tavasszy, L., Minderhoud, M., Perrin, J. F., & Notteboom, T. (2011). A strategic network choice model for global container flows: specification, estimation and application. *Journal of Transport Geography*, *19*(6), 1163-1172.

Tavasszy, L., Davydenko, I., & Ruijgrok, K. (2009, December). The extended generalized cost concept and its application in freight transport and general equilibrium modeling. In *Integration of Spatial Computable General Equilibrium and Transport Modelling Bilateral Joint Seminar Under Agreement Between NWO and JSPS*.

Tavasszy, L. A., Ruijgrok, K., & Davydenko, I. (2012). Incorporating logistics in freight transport demand models: state-of-the-art and research opportunities. *Transport Reviews*, *32*(2), 203-219.

Egypt's Suez canal H1 revenue, traffic up; upgrade helps (2010, July 26). *Thomson Reuters.* Retrieved from af.reuters.com

UNCTAD (2013). *Review of maritime transport 2013*.

UNESCAP (2007). Regional Shipping and Port Development. Container Traffic Forecast 2007 Update. *Y.: UN ESCAP*.

Van Dorsser, C., Wolters, M., & van Wee, B. (2011). A Very Long Term Forecast of the Port Throughput in the Le Havre-Hamburg Range up to 2100.*European Journal of Transport & Infrastructure Research, 12*(1).

Veldman, S. J., & Bückmann, E. H. (2003). A model on container port competition: an application for the West European container hub-ports. *Maritime Economics & Logistics*, *5*(1), 3-22.

Zohil, J., & Prijon, M. (1999). The MED rule: the interdependence of container throughput and transhipment volumes in the Mediterranean ports. *Maritime Policy & Management*, *26*(2), 175-193