A freight traffic modelling approach for the waterway network in the Netherlands

Dr. M. Khoshyaran-Lebacque
August 2000

TU Delft

Faculty of Civil Engineering and Geosciences
Transportation Research Laboratory
**Abstract**

This study deals with the Dutch waterway transport system with emphasis on the efficiency and equity impacts of current and future infrastructure and policy measures. Efforts have focused on looking at building a model system for inland waterways. This allowed a more detailed insight into the long-term development possibilities of inland waterways traffic in and around the Netherlands. The result of this research is a model system for the inland waterway freight traffic, keeping in mind the objectives of the Dutch ministry of inland waterway transport. These objectives are to develop a model that would allow for stimulus-response analysis of the waterway system traffic flow. Stimulus refers to autonomous developments in infrastructure and policy measures that are in the process of being implemented or will be realized in the future. Response refers to the anticipated reaction of freight traffic flow to the stimulus. The concern is how would shippers and carriers react to changes in the infrastructure and what are the equity impacts of various policies. The model can analyze the current state of the waterway system and make systematic predictions on the future impacts of changes in the infrastructure and policy measures.
A FREIGHT TRAFFIC MODELING APPROACH FOR
THE WATERWAY NETWORK IN THE NETHERLANDS

Dr. Megan Khoshyaran-Lebacque, Ph.D.

June, 2000

Technische Universiteit Delft
Faculteit CiTG
Bibliotheek Civiele Techniek
Stevinweg 1
2628 CN Delft

Delft University of Technology
Faculty of Civil Engineering and Geo Sciences
Transportation Research Laboratory
A freight traffic modeling approach for the waterway network in the Netherlands
A freight traffic modeling approach for the waterway network in the Netherlands

1. INTRODUCTION

1.1 Models and Their Roles

1.2 Characteristics of Waterway Freight Transport Problems
  1.2.1 Characteristics of Waterways Freight Transport Demand
  1.2.2 Characteristics of Waterways Freight Transport Supply
  1.2.3 Equilibrium of Supply and Demand

1.3 The Structure of the Classic Transport Model

1.4 Analysis of the 'Data Flow Diagram Rekenmodules BVMS'

2. A Modeling Approach to Waterway Freight Traffic

2.1 A Modeling Approach to the Freight Traffic on the Canals in Netherlands
  2.2 The Forecast Year Prediction

2.3 Simulation

3. A Discrete Production Transportation Approach

3.1 Introduction

3.2 Preliminary Requirements
  3.2.1 Shippers
  3.2.2 Carriers

3.3 A Theoretical Discussion
  3.3.1 Hypothesis 1
  3.3.2 Hypothesis 2
  3.3.3 Hypothesis 3
  3.3.4 Hypothesis 4
  3.3.5 Hypothesis 5

3.4 Discussion

3.5 A Linear Optimization Problem — Manufacturer's Case

3.6 Carrier's Case — Transporting Cargo to Destinations

3.7 Carrier's Case — Picking-Up Cargo From Origins

3.8 Transforming Tons Carried to Number of Ships

3.9 Determination of Unit Costs of Shipping Cargo to Destinations — A Multinomial Logit Approach

3.9.1 Calculation of Unit Cost of Carriers

3.10 Overall Algorithmic Approach of the Above Methodology

3.11 Case Study
  3.11.1 Carrier's Case — Transporting Cargo to Destinations
  3.11.2 Carrier's Case — Picking-Up Cargo From Origins
  3.11.3 Transforming Tons Carried to Number of Ships
  3.11.4 Calculation of Unit Cost of Carriers
  3.11.5 Conclusion

4. Data Specifications

4.1 Introduction
  4.1.1 Measurement Errors
  4.1.2 Specification Errors
  4.1.3 Transfer Errors
  4.1.4 Aggregation Errors
  4.1.5 Data Specification
  4.1.6 Network Data
  4.1.7 Fleet Data

4.2 Data Use in the Modeling
A freight traffic modeling approach for the waterway network in the Netherlands

Technische Universiteit Delft
Faculteit CiTG
Bibliotheek Civiele Techniek
Stevinweg 1
2628 CN Delft
Preface

This report represents the culmination of the author's six months of research on the freight transport on the canals network system in the Netherlands. I began studying the Dutch waterway transport system in January 2000, with emphasis on the efficiency and equity impacts of current and future infrastructure and policy measures. My efforts have focused on looking at building a model system for inland waterways. This allowed a more detailed insight into the long-term development possibilities of inland waterways traffic in and around the Netherlands. The result of my research is a model system for the inland waterway freight traffic, keeping in mind the objectives of the Dutch ministry of inland waterway transport. The objectives of the ministry as I have understood are to develop a model that would allow for stimulus-response analysis of the waterway system traffic flow. Stimulus refers to autonomous developments in infrastructure and policy measures that are in the process of being implemented or will be realized in the future. Response refers to the anticipated reaction of freight traffic flow to the stimulus. The concern is how would shippers and carriers react to changes in the infrastructure and what are the equity impacts of various policies. The model I have developed can analyze the current state of the waterway system and make systematic predictions on the future impacts of changes in the infrastructure and policy measures.

The author is deeply indebted to a great number of individuals and to Delft University of Technology for their assistance in obtaining information essential to writing this report and for their support during my stay in the Netherlands. The author would particularly like to thank Professor Piet Bovy for his valuable advice. Moreover, I would like to thank my colleague at the Technical University of Delft, (Dr. Nanne van der Zijpp) who greatly facilitated the research on the subject by offering data, analysis, and friendship.

MEGAN KHOSHYARAN
Summary

With 438 people per sq. km, Netherlands is one of the highly dense populated countries in the world. It is situated at the crossroads of Northern and Western Europe, and therefore, accommodates a heavy flow of traffic of goods. Its geographical position is an important economic resource and access to transport nodes (seaports, airports, and large cities) is vital. However, the essential incompatibility of high population density and heavy traffic means that the Netherlands is currently facing dangerous levels of congestion. Although it is essential that people and goods be able to move easily across and within the country, the Dutch strongly resent the incursion of heavy traffic on their daily environment; thus environmental protection is a particularly sensitive issue.

If this is true at national level, it is even more so within urban areas. Dutch cities have long been famous for their ‘taming’ of the car and for being environmentally friendly, the result of long-term policies in favor of alternative modes. The need for an alternative mode of transporting goods seems to be especially significant in freight traffic. Trucks constitute a significant percentage of urban traffic. The availability of canals that connect all major ports in major cities in Netherlands makes freight ships a viable mode of alternative transport of goods as compared to trucks. To use this alternative mode efficiently, the first step is to look at the system of waterways, bridges and locks that are the particulars of such a network. The second step is to observe the current state of the system, given the existing infrastructure, and the existing policies that are implemented by the centralized policy oriented Dutch government. The major concern is how does the existing infrastructure and the existing policies affect the main participants of the waterway system, mainly the shippers (manufacturers) and the carriers (ship owners). Finally the objective is to have forecasting abilities. The idea is to be able to simulate changes in the waterway system infrastructure and implement new policies or modify the existing ones and observe the impact such changes have on the behavior of shippers and carriers. It is to achieve this end that the current report is written.

Chapter one is concerned with some of the basic ideas of transportation planning that directly applies to the context of waterway freight transport. A description of models and their roles are given. The characteristics of transport problems such as transport demand and supply are discussed. Issues in transportation modeling are explored. The structure of the classic transport model is described. And finally it looks at the current proposal to model freight shipping and forecasting. The shortcomings of the proposed model are discussed and alternative ways of improving the existing approach are explored.

The second chapter of this report describes a generalized model for freight shipping, with particular attention to the existing limitations and requirements of the problem in the Netherlands.
The chapter explores alternative models for estimating a ship matrix for the base year. The chapter ends by suggesting two forecasting models.

The third Chapter is exclusively devoted to describing a procedure to find a ship matrix. A ship matrix is a matrix of origins and destinations with cells indicating the number of ships of different sizes that carry freight between a particular origin-destination pair. The approach is economic based. The idea is to simulate the mechanics of a free market in determining the interaction between shippers and carriers and consequently determine the flow of cargo ships by type and size in the inter-city waterway system.

The model uses a probabilistic approach (*Logit probability*) in choosing carriers, and their unit cost of transport. The model then uses a deterministic approach (*an optimization*) to replicate the decision making process of a carrier who through minimizing his total costs determines how to satisfy all demands from current inventory of ships. The constraints in the optimization are dependent on either total demand or total supply and ship's maximum and minimum capacity. *A 3D iterative proportional fitting technique* is used to adjust the result of the optimization in order to make it consistent with total demand, total supply, and carrier's desired loading scheme. The last step is to translate tons into number of ships.

Chapter four is concerned with the specifics of data collection. The main problem in data collection techniques is inaccuracy. The chapter describes different types of errors associated with data collection such as: measurement errors, specification errors, sampling errors, transference errors, and finally aggregation errors. The chapter also identifies the data needed for the waterway freight traffic modeling and classifies and describes the data based on the report by van der Zijpp *et al*, 1999.

Chapter five will provide a design guidelines for network definition and suggests a dynamic assignment model. Two objectives are in mind: accuracy and cost. A waterway network can be represented as a system of nodes and links or in other words a system of directed graphs. Links are characterized by several attributes such as length, speed, capacity, depth, and water level. Waterway network is somewhat different from road networks. A waterway network has a set of bridges and locks as well as canals. Bridges and locks are represented as links. Aside from the characteristics mentioned above each bridge and lock is characterized by its storage capacity for queues. Canals are links that do not encounter queues. Therefore, delays on canals only occur before reaching bridges and locks. Generalized cost for route choice is a linear combination of time and distance. There may be other factors that affect route choice but for the sake of simplification they are not included here. The basic idea in assignment is the assumption of a rational traveler, i.e.; one choosing the route, which offers the least perceived (and anticipated) individual costs. A number of factors are thought to influence the choice of route when driving...
between two points; these include journey time, distance, monetary cost (fuel and operating costs, tariffs), congestion and queues at the bridges and locks, type of maneuvers required, reliability of travel time and habit.

The purpose of chapter six is to introduce a new model for estimating a dynamic or time-dependent trip table for a waterway freight transport system and to test the model using traffic count data obtained from detectors on a sample of canals in a waterway network.

So far the only reliable way to obtain an origin destination trip table has been through surveys at origins and destinations. These surveys are expensive and they do not reflect changes in traffic flow that occur over time.

There have been a considerable number of studies conducted to determine if an orgin-destination matrix could be derived from detectors, because these data can be collected and processed automatically. To develop more effective strategies for better management of canals it is necessary to improve the existing models for estimating O-D flows. In this chapter a reliable substitute to the traditional method of obtaining trip table is introduced.

The method proposed in this paper adds to existing literature first by introducing an alternative approach to estimating time dependent O-D volume which varies from the predecessor models in its functional relativity and sequential implementation. Second, the simultaneous inclusion of the possibility of multiple routes and allowing for trips with travel times exceeding one interval, makes this model different from earlier literature.

Chapter seven will expand more on the forecasting options given in chapter two. Two forecasting options were suggested in chapter two. The objective of both options is to forecast waterway freight traffic for the future. The difference between the two models is in the way that the endogenous variables are dealt with. As is a common knowledge there are two types of variables that are used in modeling: the endogenous variables and the exogenous variables. The endogenous variables are those variables that are to be forecasted as part of the modeling exercise like the ship matrix (that is the number of ships that go between origin-destinations by ship type). The exogenous variables are those that are required to run the model and are originated externally to the models themselves. Typical examples are the network data, and the ton matrix (the description of this matrix is provided in chapter 2). The values of these variables should be provided for the base year and for the forecast years.
The importance of these variables in influencing the accuracy of the whole modeling exercise is very high. The specification errors are less significant than the errors in exogenous variables. In many cases the errors are the result of errors in official forecasts. The difference between the two models of forecasting suggested in chapter two is in the level of detail and disaggregation.

In chapter eight the aim is to give general comments about the waterway freight transport and point out the requirements of future research in this area. Given road congestion, which in large part is caused by truck movement on the roads, facilitating the movement of freight on the waterway network system is likely to have a major impact on economic development. Considering the importance of waterway freight transport it is surprising that much less research has been undertaken on modeling this type of movement than the effort allocated to other types of freight movement.
1. INTRODUCTION

The purpose of this report is to explore the shortcomings of the general four-step model, which is adopted for the modeling of freight traffic on the canals in the Netherlands. The report will specifically focus on the flow diagram shown in the November 30, 1999 report entitled "Data flow diagram rekenmodules BVMS". To rectify the problems encountered in the four step modeling approach; a new modeling approach is provided. The idea is to maintain the general skeleton of the four step modeling but make it more compatible with the new age mentality of the people who work in the field. Nowadays, traffic engineers think of traffic or any type of transport as a dynamic or time dependent phenomenon. The aim these days is to simulate ideas and measure the consequences. To achieve this end it is important to know the transport system thoroughly. To keep up with the requirements of the 21st century, this report attempts to give an alternative approach to the four step modeling process, while still trying to integrate the general structure of the conventional four-step modeling.

1.1 Models And Their Roles

Waterway freight transport is a special case of freight transport. To model this form of transport attempt is made to define the system as a simplified representation of a part of the real world-the system of interest - which concentrates on certain elements considered important for its analysis from a particular point of view. The model should be therefore, problem and view specific. Therefore using such a broad definition allows for incorporating both physical and abstract models. A physical model is basically aimed at design. A physical model can provide a more realistic view of the waterway network system depicting the location of canals, the bridges and the locks. An abstract model is aimed at analytical representations of the theory about the nature of the waterway freight transport mainly how it works. The abstract models that will be used in waterway freight analysis are mathematical models. Mathematical models attempt to replicate the behavior of shippers and carriers and the interaction of carriers with the waterway network by means of mathematical equations based on certain theoretical statements about it. Although they are still simplified representations, these models may be very complex and often require large amounts of data to be used. However, they are invaluable in offering a 'common ground' for discussing policy and examining the inevitable compromises required in practice with a minimum of objectivity. Another important advantage of mathematical models is that during their formulation, calibration and use one can learn much, through experimentation, about the behavior and internal workings of the system under scrutiny.
A model is only realistic from a particular point of view. The same is true of analytical models; their value is limited to a range of problems under specific conditions. The appropriateness of a model is dependent on the context where it will be used. The ability to choose and adapt models for particular contexts is one of the most important elements in the complete planner's tool-kit. Transport modeling can make a significant contribution in decision making and planning in the transport field.

Although the use of models is inevitable and formal models are highly desirable, transport modeling is only one element in transport planning: administrative practices, an institutional framework, skilled professionals and good levels of communication with decision makers, the media and the public are some of the other requisites for an effective planning system. Moreover, transport modeling and decision making can be combined in different ways depending on local experience, traditions and expertise.

1.2 Characteristics Of Waterway Freight Transport Problems

Waterway freight transport problems have become more into focus as an alternative mode of freight transport. General increase in road traffic and transport demand has resulted in congestion, delays, accidents and environmental problems well beyond what has been considered acceptable so far. Economic growth seems to have generated levels of demand exceeding the capacity of most transport facilities. These problems are not likely to disappear in the near future. Given the resources are not unlimited, waterway freight transport can be considered as a new transport provision. It would be prudent to maximize the advantages of using this mode of freight transport while minimizing the money costs and undesirable side effects.

1.2.1 Characteristics Of Waterways Freight Transport Demand

There are two types of demand to be considered here. Type one demand is the demand for carriers to transport goods. Type two demand is the demand for transport services. Both demands should be analyzed. In type one demand the demand for shipping goods induces the shippers (manufacturers, producers) to seek carriers to transport their goods to desired destinations. This desire on the part of shippers to transport goods in turn attracts carriers into the transport market place. Each of the participants namely shippers and carriers have specific characteristics. Shippers are characterized by the type of goods they produce, the origin (warehouse) they want to ship the goods from, and the destinations they want to send the goods to, and the amount (tons) of product they want to ship, and finally the price they offer to transport their goods.
The carriers are characterized by the number of ships they own (fleet size), by the type of ships they own (barges, tankers, etc.), by the sizes of ships they own (small, medium, and large), and by their willingness to either accept shippers prices or offer their own prices to carry goods or their willingness to negotiate prices.

There is one common characteristic that is shared by both groups of carriers and shippers. This characteristic is that both groups aim at minimizing costs and maximizing profits. What induces the supply of carriers is the shipper’s willingness to pay the cost of transporting goods plus providing monetary benefits to the carriers. In essence what brings carriers into the market depends fundamentally on how attractive the market is.

The less attractive market means less carriers are willing to enter the market therefore reducing the supply of means of transport which affects the supply of transportation services mainly the infrastructure, its evolution, its management, and its maintenance.

The second type demand is the demand for transportation services, namely infrastructure. A good transport system widens the opportunities to satisfy the industrial needs that are distributed over space. A poorly connected system restricts options and limits economic and social development. Waterway freight transport system like any other transport system takes place over space. A space is characterized by zones, which generate either supply or demand for goods. Zones are connected through a network of canals (waterways). Canals are characterized by their width; depth, water level and super imposed structures such as bridges and locks. The waterway system demand and supply have very strong dynamic elements. A good deal of demand for the waterway network system is concentrated on a few hours a day, and is affected by seasonal variations. This time variable character of transport demand makes it more difficult and interesting to analyze and forecast. It may very well be that a transport system could cope well with the average demand for travel in an area but that it breaks down during peak periods.

1.2.2 Characteristics Of Waterways Freight Transport Supply

The first distinctive characteristic of transport supply is that it is a service and not a good. Therefore, it is not possible to stock it for use in times of higher demand. A transport service must be consumed when and where it is produced, otherwise its benefit is lost. Many of the characteristics of transport systems derive from their nature as a service. In very broad terms a transport system requires a number of fixed assets, the infrastructure, and a number of mobile units, the ships. The supply of ships depends on the carriers. Carriers offer a service. They offer to transport goods to destinations.
It is the equilibrium between the supply of carrier service and the transport service that brings about a desirable environment in which both the shippers and the carriers are satisfied. It is the combination the transport supply and the carrier supply, together with a set of rules for their operation that makes possible the movement of goods.

1.2.3 Equilibrium Of Supply And Demand

There are two levels of equilibrium between supply and demand. Level one occurs in the market place where shippers and carriers (demand and supply) have to reach equilibrium; by which it is meant an agreement on price of transporting goods. A market equilibrium is achieved when all shipper and all carriers who participate in the market are satisfied. Consider a set of shippers \( SH \) with a set of characteristics \( Ul \). Demand \( D1 \) for shipping goods is a function of the characteristics of the shipper and a set of prices \( P \) he is willing to offer for transporting goods.

\[
D1 = f(U1, P) \tag{1.1}
\]

Consider also a set of carriers \( C \) with a set of characteristics \( U2 \). The supply of carriers is then a function of

\[
S1 = f(U2, P) \tag{1.2}
\]

Combining equation (1.1) and (1.2) for a short period and a fixed set of shippers and carriers, given that the their respective characteristics stay unchanged for the duration of the market activity, a set of equilibrium points between supply and demand can be found. The long term forecasting of the equilibrium points is difficult. The forecasting of the shippers and carriers activities in the market place strongly depends on how well the costs of transporting goods as well the level of shipments are forecasted.

In general terms the role of transport planning in waterway freight transport is to ensure the satisfaction of certain demand \( D2 \) for goods movements with one trip purpose, delivery of goods to certain destinations from certain origins. The trips are made at different times of day and the year, using various modes (ship types, like barges and tankers), given a transport system (a waterway network system) with a certain operating capacity. The waterway transport system itself can be made up of:

- An infrastructure (e.g. a waterway network system);
- A management system (i.e. a set of rules, for example it is first come first serve at the bridges and locks, and control strategies, for example at bridges and locks);
- A set of transport modes (ship types, like barges and tankers), an their operators.
Consider a set of volumes on a network $V$, a corresponding set of speeds $S$, and an operating capacity $Q$, under a management system $M$ specifically the management of bridges and locks. In very general terms the speed on the network can be presented by:

$$S = f(T, Q, V, M, C, P) \quad (1.3)$$

with $T$ being the aggregated demand function. The volumes $V$ are functions of $T$ through an assignment process.

Delay in the network is then a function of:

$$A = f(T, Q, V, M, C, P) \quad (1.4)$$

The delay can be taken as an initial proxy for a more general indicator of the level service (LOS) provided by the transport system. In more general terms a LOS would be specified by a combination of speed, delay or travel times, waiting times and loading times, and price effects.

The management system $M$ may include traffic management schemes, area control and regulations at the bridges and locks. The capacity $Q$ would depend on the management system $M$ and on the levels of investment $I$ over years, thus:

$$Q = f(I, M) \quad (1.5)$$

The management system may also be used to redistribute capacity among the infrastructure, producing $Q'$ and/or giving priority to certain types of users over others based on physical limitations (such as seasonal canal depths). As in the case of most goods and services, one would expect the level of demand $T$ to be dependent on the level of service provided by the transport system and also on the allocation of activities $A$ over space:

$$T = f(A, A, S, Q, SH, P) \quad (1.6)$$

Combining equations (1.3), (1.4) and (1.6) for a fixed system one would find the set of equilibrium points between supply and demand for transport. But then again, the activity system itself would probably change as levels of service change over space and time. Therefore, one would have two different sets of equilibrium points: short-term and long-term ones. The task of transport planning is to forecast and manage the evolution of these equilibrium points over time so that social welfare is maximized. This is, of course, not a simple task: modeling these equilibrium points should help to understand this evolution better and assist in the development and implementation of management strategies $M$ and investment programs $I$. 
Sometimes very simple cause-effect relationships can be depicted graphically to help understanding the nature of the waterway transport system.

Figure 1.2 Shipping and Infrastructure relationship

Economic growth provides the first impetus to increase ship ownership. More carriers means more shippers wanting to transfer from road freight transport to waterway freight transport.

A good waterway network facilitates the transport of goods which results in easier movement of ships. This in turn modifies the cost of transport and encourages more shippers to seek waterway freight transport as opposed to other means. This in turn means fewer road freight transport users. Many physical improvements can encourage the use of waterways by the carriers and result in relieving truck congestion on the roads by a some degree.

1.3 The Structure Of The Classic Transport Model

Years of experiment and development have led to a general structure, which has been called the classic transport model. This structure in effect has come about as a result of practice in the 1960s, but in large has remained unaltered despite major improvements in modeling technique during 1970s, 1980s, and 1990s.

The general form of the model is shown in figure 1.5. The method starts by considering a zoning and a network system, and the collection and coding and planning, calibration, and validation data. These data would include base year levels of various types of needed data such as fleet data, and the tons carried between origin-destination pairs. These data are then used to estimate a model of the number of trips generated and attracted by each zone of the study area (trip generation). The next step is the allocation of these trips to particular destinations, in other words their distribution over space, thus producing a trip matrix. The following stage normally involves modeling the choice of mode and thus results in modal split, i.e. the allocation of trips in the matrix to different modes. Here modes refers to ship types and sizes.
The classic model is a sequence of four sub-models: trip generation, distribution, modal split and assignment. It is generally recognized that travel decisions are not actually taken in this type of sequence; a recent view is that the 'location' of each sub-model depends on the utility function assumed to rule all these travel choices. However, the above structure provides a point of reference to alternative models.

1.4 Analysis Of The 'Data Flow Diagram Rekenmodules BVMS'

To better follow the arguments, the diagram of the November 30, 1999 report entitled "Figure 2. Data flow diagram rekenmodules BVMS" is copied here. The objective of the modeling approach is to setup a user-friendly system that analyses the existing state of the system and is capable of forecasting the consequences of either changes in the infrastructure or management policies. In this section some of the short-comings of the approach below is discussed.
In the base year section of the flow chart there are three sources of input data: the OD matrix in tons (all modes) for the base year, the fleet data for the base year, and the trip data for the base year. Each data comes from a different source. For example the OD matrix is derived based on assumptions made by economists, whereas the fleet data comes from the Census bureau. Therefore, there is no common procedure used to develop these data sets, which implies that the data sets do not necessarily correspond. This automatically dictates adopting an approach that would try to correlate the data sets together in order to get consistent results.

The second observation is to do with the time scale. The OD matrix by tons is a yearly data as well as the fleet data, (number of ships broken down by type and size), whereas the trip data, (number of ships that go between an origin-destination pair) is collected or can be collected daily and at different times of day. It seems that there is no consistency in time scale.

The third observation is about assignment and traffic production boxes. The traffic production box is not clear at all. What does it refer to? What does it do? Does it replace all trip generation, distribution, and modal split boxes in the four step classical model? And if so how?
Or may be by the term traffic production it is meant some sort of an aggregate choice model, then the utility function can not possibly be just the function of travel times. There are many other variables that are not transportation related but rather related to the interaction between the carrier companies, and the carriers, and these variables should be included in the choice model. Travel time seems to be estimated through an assignment model. But the other parameters are not as easily calibrated as the travel time parameters. Therefore, even though it is possible in the figure 2 flow chart approach to calibrate travel time parameters through an assignment model it is nearly impossible to calibrate the other variables that must be included since no approach is given for such estimation.

The assignment model of figure 2 is not specific at all as to what kind of assignment is intended for the analysis. It seems that the assignment is used mainly to find travel times which are then incorporated into the “traffic production” box to modify the “ship matrix”, and then compare it with the observed rip data. At this point it is important to be clear about what specific approaches need to be used. For example, a time dependent Logit assignment seems to be a far better approach than let’s say a static assignment model. Both the traffic production and the assignment are vague. In place of “traffic production”, an aggregate mode choice model based on the utilities of the carriers and the customers should be used. In place of an arbitrary assignment a time dependent Logit assignment and rip table estimation can be used. There are two levels of choice modeling here, one before the ships are loaded and one after the ships are loaded. Level one choice is related to the choice that is concerned with both shippers and carriers, and level two choice is when the carriers and shipper are matched and carriers have to choose a path between their origin-destination pairs that fits them best. More detail will be provided when describing the modeling approach.

The fourth observation has to do with the placement of the assignment in the flow chart. It seems that first through a “traffic production” procedure the trips are distributed per origin-destination by ship characteristics, and then attempt is made to find what route these ships have taken. Whereas in this sort of environment the carriers may not necessarily take the shortest route, and even so their utility function may not be just based on shortening travel times. Other elements such as time of day, specifications of the ship, carriers’ deadlines or lack of, the volume of load, the type of goods, and the type and size of the ship used in transport can affect the route choice. The path choice based on trip data that does not incorporate the above elements will not produce useful results. The trip data should be used to reduce error in an origin-destination trip matrix (trip table) estimation model which has an implicit Logit assignment imbedded in it.
In this case the ship matrix is used as a base matrix in the estimation in order to provide a reasonable starting point. The ship matrix should be aggregated by type and size for each origin and destination in order to to be compatible with the observed trip data, it then can be inputted into the origin-destination matrix estimation as a base matrix. The origin-destination matrix estimator model is time depended and the route choice is by time of day as well. This approach will produce results that are both externally (trips per origin-destination), and internally (freight traffic on canals) consistent.

In the second half of the flow chart that deals with forecasting, it seems that the trip data is forecasted. The forecast trip data can not possibly have any relation to the trip data of the base year; because in the base year the trip data is collected several times during a day and is therefore an observed data, whereas the forecast year trip data will be just an estimated aggregate data that is derived specifically for the forecast year.

The trip data of the forecast year is used in the assignment, which means that the output (travel times) can not possibly have the same meaning as the base year assignment output, and therefore can not be used with the corresponding parameters in the Logit assignment model that are calibrated using base year data. If a bi-level choice modeling is used, one that deals with the freight problem before loading ships and one that deals with the network related problem after loading ships one might be able to solve the problem more efficiently.

In forecast modeling, two options are suggested. Each one will produce satisfying results. The difference is that the second option is more responsive to changes in policy measures, such as tax increases, and modification of waiting time, and operational time at locks and bridges. Both models will be responsive to changes in the network, as the specifications are included in the Logit assignment used for path choice modeling.

“Figure 2. Data flow diagram rekenmodules BVMS” conforms with the structure of four step modeling without attempting to modify it in order to reflect the specifics of the waterway freight transport. It does not consider modern day capabilities of data collection, and ignores modern day requirements as far as analysis and management decision-makings are concerned. Based on the survey of the available data and modern data collection capabilities it is possible to implement a more detailed model. One that would take into account the effects of time variations on the behavior of freight transport and could estimate the effects of decisions made in the market place which directly affect the infrastructure of waterway network system and its management strategies.
A modeling approach to waterway freight traffic
2. A MODELING APPROACH TO WATERWAY FREIGHT TRAFFIC

2.1 A Modeling Approach To The Freight Traffic On The Canals In Netherlands

In this section two approaches in modeling the freight traffic both in the present and in the future are presented. Both approaches to modeling will enable the operators not only to have a better understanding of the system but also operate the bridges and the locks better. The second approach also helps the planners in setting up policies that could facilitate the freight flow on the waterways. This in turn will make inland waterway traffic a more competitive alternative as compared to other modes such as trucks.

2.1.1 General Approach- Base Year

As was mentioned earlier, the approach to modeling this problem is to have a bi-level choice process. A choice is made both on the part of the carriers and the shippers before loading cargo into ships. This choice has to do with the selection of ships and the optimal loading possibilities. There are two utility functions involved: one reflects the utility of the shipper, and the second reflects the utility of the carrier. Both parties attempt to maximize their profits or minimize their costs. The problem can be solved as an optimization problem with 2 linear/non-linear set of objective functions to be maximized/minimized, given a set of constraints. The linearity or non-linearity of the objective functions should be investigated. The constraints have to do with limitations in lets say rates that can be charged; and tons of cargo, etc. An aggregate Logit Mode Choice (ALMC) model can be used to determine unit price of transport both for carriers and shippers. The unit prices of transporting goods are used as coefficients in the objective functions. The obvious input data to this model: 1) OD matrix (tons) base year and 2) Fleet data base year. The output from implementing this model is a matrix with rows as ship types \{1,2,\ldots,K\} and columns as tons \{x_1\text{tons}, x_2\text{tons},\ldots, x_n\text{tons}\} by cargo type, as is shown below:
The second phase of the operation is to translate the matrix that is obtained above into a matrix of origins and destinations by ship types. Here the relevant operation is an IPF (Iterative Proportional Fitting) procedure, particularly, an Entropy IPF. The procedure is simple and can be defined mathematically. The input to the process is 1) total tons shipped per origin, and per destinations. 2) Total tons shipped by ship type/size. The output should be a matrix of origin-destination by ship type/size. This is an aggregated matrix, and is derived in order to give a ship matrix per year. The IPF procedure is a trivial step in the overall procedure, but nonetheless necessary. The aggregated ship matrix then goes into the trip table estimation procedure as an initial OD matrix by ship type/size. The observed counts of ships by type/size and by time of day are used for error estimation and minimization. The procedure described is depicted in Figure 2.
The third phase of the operation is to do a *Dynamic Logit Assignment and Trip Table Estimation*. At this stage the ships are loaded and the carrier has to make a choice of which path to take. The reason to follow this format is that even though there are always specific routes that are taken by the carriers, the probability of taking these routes varies considerably during different times of day depending on the utility of choosing a path for that time interval. Given the data that is collected by time interval per day from the data collection points (telpunktkaart met vaarwegnummers), the required input is already available. By the required input it is meant the number of ships by ship type that cross these points, and their load. Given the load it is possible to derive the actual speed on the waterway by time of day. Given network specifications such as distances, then it is possible to calculate travel times on all the waterways in the waterway network. At this point a set of possible paths given the network constraints(such as the water level, the canal depth, etc.) and ship constraints can be obtained through a static assignment model. Given travel times on the links of these paths, the probabilities of choosing a path during each interval are calculated.

The time dependent origin-destination, OD (trip table) matrix estimation follows. But before going through the procedure, one should address the question “why do we need a time dependent OD estimation?” The answer is this: The trip data is given by time of day per day, for several days during a month and for several months; typically one month or two during each season. If we try to aggregate the values per day up to a year, and then try to match the ship matrix to the aggregated trip data, chances are that we will not get a match.

By matching the aggregate data we just get an external consistency, without bothering to look into the internal consistency of the system, that is matching data with the observed data that is collected at collection points or in other words the time dependent observed counts data. Even if there is a match at an aggregate level, chances are that there is not a match with the observed counts data that are obtained within time intervals during a day. Therefore, the estimated ship matrix will not give any insight into the movement of these ships within the network. If one of the objectives of the study is to analyze the network system for efficiency, then that aim will not be achieved. The estimation process has to give results that are consistent both per origin-destination and per observed link counts in the network system.

Based on the above reasoning, it is more efficient to estimate ship types per origin – destination by time of day, and try to match these matrices with the observed trip data. The resulting matrices will match the observed trip data across time intervals and will match the observed trip data at an aggregate level too. The idea is to achieve both external and internal consistency.
By external consistency it is meant consistency on an aggregate level, matching estimated trips per origin-destination with the observed trip table. Internal consistency refers to matching estimated freight traffic on the canals with the observed freight traffic. The process suggested here is a modified version of the Kalman filtering algorithm, with iterative Least Square Estimation. The overall procedure is depicted in Figure 3 below in a flow chart.
A freight traffic modeling approach for the waterway network in the Netherlands

Figure 3. Flow Chart Depicting Time Dependent Logit Assignment and Dynamic Trip Table Estimation Process
For each combination of ships and tons derived from the mode choice model, run the assignment and trip distribution process until it cannot reduce any further the difference between the observed and estimated ship matrices. The parameters are calibrated based on the combination of the mode choice, assignment and trip distribution that gives the best match. Once the parameters both in the mode choice and the Logit assignment and trip distribution are calibrated, they can be used in the forecasting process. The flow chart below combines figures 2 and 3 in order to give an overall view of the tasks involved in the base year. The general process is necessary both for estimating and calibrating the required variables and parameters by attempting to achieve to reach both external and internal consistency.

Figure 4. The Overall View of Parameter Calibration, Base Year
2.2 The Forecast Year Prediction

The main idea is to apply the same procedure in the forecast year. Therefore, the first step in the forecasting process is to come up with different cargo/ship loading alternatives based on both the shippers and the carriers maximizing/minimizing objectives. The parameters used in this process are calibrated in the base year and are used unchanged for forecasting purposes. The resulting ship matrix is then input into the Dynamic Logit Assignment in Option 1, and Time Dependent Trip Table Estimation model in Option 2.

OPTION 1

The methodology is best described in the following Flow Chart.

Figure 5. Recommended Procedure for Forecasting Freight Traffic
OPTION 2

The second option is to implement Dynamic Logit Assignment and Trip Table Estimation (DLATTE). With respect to forecasting, the following is proposed: The trip data should be forecasted through an statistical method; for the sake of argument let's say a time series method or a regression method in order to derive forecast numbers. The counts of ships by ship type should also be forecasted through a statistical approach, a likely candidate would be, a Dynamic Systems approach to derive the forecast numbers for the counting stations. The idea is to match data that is derived from entirely separate forecasting methods. This is necessary in order to make sure that our forecasting procedure is reliable, and still achieves both external and internal consistency in the forecast.

The parameters in the DLATTE are calibrated in the base year and are used in the forecast year. The process is shown in a Flow Chart Format as follow:
2.3 Simulation

The above methodology can be simulated graphically. Events can be incorporated into the simulation as the level of complexity increases. There are package programs such as Geographic Information System (GIS) that can be used for this purpose. GIS has many advantages, among which is its capability to incorporate both the canal network and the road network. It accumulates data in layers. It then can superimpose the layers. Therefore, it can superimpose the two networks. The GIS database can include canal specifications, as well as road specifications. Both traffic on canals and traffic on the roads can also be included in the GIS maps. Another source of database is the shortest paths chosen by ship types, and travel times on the shortest paths. This data is the result of modeling canals freight traffic as well as road traffic.

Given all the above information GIS can then show the interaction between the road network and the canal network, i.e. the analyst can analyze the efficiency of the operation of the bridges where there is a high level of road traffic, as well as high level of ship traffic. Therefore, he/she will be able to make policy measures that can affect both the canal network flow and the road network flow. GIS has the capability of including events and their consequences as well.
3. A DISCRETE PRODUCTION TRANSPORTATION APPROACH

3.1 Introduction

The purpose of this chapter is to introduce a method that could simulate the mechanics of a free market in determining the interaction between shippers and carriers, and consequently determine the flow of cargo ships by type and size in an inter-city waterway system. The model uses a probabilistic approach (Logit probability) in 1) choosing carriers, and 2) their unit cost of transport. The model then uses a deterministic approach (An optimization) to replicate the decision making process of a carrier who through minimizing his total costs determines how to satisfy all demands from current inventory of ships. The constraints in the optimization are dependent on either total demand or total supply and ship's maximum and minimum capacity. A 3D iterative proportional fitting technique is used to adjust the result of the optimization in order to make it consistent with total demand, total supply, and carrier's desired loading scheme. The last step is to translate tons into number of ships.

The idea is to find how many ships of certain type and size carry cargo among origin-destination pairs that are connected through a network of waterways. To find such a matrix, the methodology should explain the interaction between shippers and carriers. Shippers look for carriers that offer them the best price. Carriers on their part have to decide on the price they would like to charge, and the number of ships of various sizes they want to engage in transporting cargo. The elements that affect the carrier's decision are, the location of origins, the location of destinations, the level of supply and demand, the type of cargo, and the transportation costs. Two cases are looked at here, case one is when the carrier considers the cost of carrying cargo to destinations. Case two is when the carrier has to consider the cost of picking-up cargo from origins.

Both for the carriers and the shippers the decision process consists of minimizing total costs given a set of constraints. What relates to cost in shipper's case is the unit cost of transporting products to destinations. Carriers are concerned with the unit cost of picking up and dropping off cargo given a particular ship type and size. The underlying assumption is that the two parties are active in a market environment. In such environment shippers have to choose among different carriers. The carriers are active in as much as they provide shippers with a price of shipping cargo to destinations. The shippers on their part provide information on prices and the type and amount of cargo that is to be carried to different destinations. In short, both shippers and carriers have perfect knowledge of prices, as well as demand, and supply. The pairs that are highly compatible will be the ones that use the waterway transport system. There is always a match, and the proof is in the free market mechanism of supply and demand, where pricing plays an
important role. The result of this interaction is a matrix of origins and destinations with cell entries representing number of ships by given sizes and tons they carry.

### 3.2 Preliminary Requirements

To set up the model some initial information is needed. The following outlines the required data:

#### 3.2.1 Shippers

- Information on total demand and supply should be available.
- Shippers should provide a fixed unit cost of transporting goods to their desired destinations.
- All details relating to the characteristics of the products to be shipped should be made available to carriers.

#### 3.2.2 Carriers

- A table of origin-destination matrix with each entry indicating tons carried to a particular origin-destination of a certain cargo type. Cargo can be divided into three basic types: Solids (perishable / non-perishable), Liquids(hazardous/non-hazardous), Gases(hazardous/non-hazardous).

The type of cargo usually dictates what type of ship should be used, for example, gases are usually carried by tankers, which in turn determine the price elasticity of demand for the tanker, and therefore, affect setting-up the unit cost of transportation.

The specifications of the cargo type usually dictates a change in the unit cost of transportation, since, for example it costs more to deliver oxygen cylinders that can easily explode than let’s say helium gas. In this case, some safety measures should be taken in order to avoid explosion. This is usually translated into a rise in the cost of transport.

- Total cost of shipping cargo from an origin and to a destination. This information is necessary in order to calculate the unit cost of transporting cargo from the carrier’s point of view. A table with rows being origin-destination pairs, and columns being different
cost variables relating to cost of shipping cargo is required. The point will be cleared up later on in the paper.

- For each ship type, maximum and minimum capacities for different ship sizes should be specified. The concept is rather clear here; the assumption is that ships should travel with at least minimum capacity loaded, in other words, it is not worth while for a carrier to accept transport of cargo if his ship is not at least loaded up to minimum it’s capacity. It is assumed that the only empty trips are the return trips to base, and pick-up trips to the origin.

- Information on the number of available ships of different type and size per carrier is necessary.

3.3 A Theoretical Discussion

In order to build up the theory, it is necessary to state a few hypothesis.

3.3.1 Hypotheses 1

There are two sets A and B of Non-empty, Bounded, Discrete, and mutually disjoint elements.

In this case let’s assume that Set A is a collection or union of shippers (manufacturers). This set consists of elements $A_{ik}$, where $i = 1, \ldots, m$ and $m$ is the number of shippers in the set, and $k$ is an specific characteristic of $A_i$. $k$ is contained in a larger set $K$, i.e. $k \in K$. For example, $k$ is a shipper who produces a particular type of cargo; and $K$ consists of: producer of a particular type of product, who has a pre-set origin-destination, and offers a specific unit cost of transport, etc. Please note that shippers of a particular type of cargo have conflicting interests, since they compete with each other in the market in order to find suitable carriers. We could demonstrate this as follows:

$$A = \{A_{1k}, A_{2k}, \ldots, A_{mk}\} \quad i = 1, \ldots, m$$

where

$$A_{ik} \cap A_{i'k} = 0 \quad \forall A_{ik}, A_{i'k} \quad i \neq i'$$

By the same token, set B is a collection or union of carriers. This set consists of elements $B_{jk}$, where $j = 1, \ldots, n$ and $n$ represents the number of carriers, and $k'$ is an specific characteristic of B. $k'$ is contained in a larger set $K'$, i.e. $k' \in K'$. For example, $k'$ is a carrier who has a particular
fleet size. $K'$ would be a set consisting of: a particular fleet size, partial to carrying a particular type of cargo, etc. Please note that carriers of a particular type of cargo have conflicting interests, since they compete with each other in the market in order to attract suitable shippers. We could demonstrate this as follows:

$$B = (B_{1k'}, B_{2k'}, \ldots, B_{nk'}) \quad j = 1, \ldots, n$$

where

$$B_{jk'} \cap B_{j'k'} = 0 \quad \forall B_{jk'}, B_{j'k'} \quad j \neq j'$$

3.3.2 Hypothesis 2

Both shippers and carriers are in a Market. A market is the place or context in which buyers and sellers buy and sell goods, services, and resources. Though it might take time, eventually, the market condition reaches equilibrium.

Equilibrium refers to the market condition, which once achieved, tends to persist. Equilibrium results from the balancing of the market forces. The equilibrium is determined exclusively by the interaction of the forces of demand and supply only in a perfectly competitive market.

A market is said to be perfectly competitive when the number of buyers and sellers of the identical commodity are so numerous that no individual buyer or seller is able to affect the price of the commodity. In addition, in a perfectly competitive market, entry into and exit from the industry are "easy", there is a perfect knowledge of prices and quantities, and there are no interferences with the operation of the market mechanism.

3.3.3 Hypothesis 3

Both for carriers and shippers the decision process consists of minimizing total costs of transport given a set of constraints. What relates to cost in shipper's case is the unit cost of shipping products to destinations. Carriers are concerned with the unit cost of transporting goods to destinations given a particular ship type and size, knowing the cargo type.

The two parties are active in a market environment. In such environment shippers have to choose among different carriers to ship their cargo to desired destinations. The carriers are active in as much as they provide shippers with a price of shipping cargo to destinations.
The shippers on their part provide information on the type and amount of cargo that is to be carried to different destinations, from origins. In short, both shippers and carriers have *perfect knowledge* of prices, as well as demand, and supply, and any other relevant information.

### 3.3.4 Hypothesis 4

Two elements are important here, location of warehouses or factories where carriers have to pick-up cargo, and the location of the drop-off points or destinations. Two events are looked at here, event one is when the carrier considers the destinations only, i.e., the determining factor is the cost of carrying goods to destinations. Event two is when the carrier considers the cost of picking-up cargo from origins, and this is the determining factor in his decision whether to accept or reject transporting cargo. These events are considered as disjoint events.

By disjoint it is meant that if the carrier has to make a trip to an origin to pick-up cargo, he would mainly consider the utility of picking-up cargo at the origin. This is a determining factor in whether he accepts to transport goods or not irrespective of the cost of transport to destinations.

### 3.3.5 Hypothesis 5

The level of supply and demand clearly not only affects the shippers but the carriers as well. Subjectively carriers are in the supply and demand chain. The effect is in the unit cost of transporting goods.

### 3.4 Discussion

In a perfectly competitive market there is always a probability that a shipper will come to an agreement with a carrier as to the amount of cargo, and the price of transport.

Let's define a market as a union of the two sets A and B (see Appendix D).

\[
A \cup B = M
\]

where

\[
M = \text{market place}
\]

In terms of probabilities, the probability of a match between a shipper and a carrier exists and is greater than zero. In this case an event is an agreement between one shipper and one carrier. Implying that 2 person teams are selected at random from a group of m shippers and n carriers.
The total outcome is when every shipper has found a carrier. Therefore, the total outcome is given by:

\[ n(A_{ik}, B_{jk}) = \binom{m + n}{2} \]

It is assumed that "random selection" means that each of the outcomes is equally likely. Let \((A_{ik}, B_{jk})\) be the event that a shipper chooses a carrier. Then the number of outcomes belonging to an event is given by:

\[ n(A_{ik}) = \binom{m}{1}, \quad n(B_{jk}) = \binom{n}{1} \]

The probability that a match between a shipper and a carrier occurs is:

\[ P(A_{ik}, B_{jk}) = \frac{n(A_{ik}) n(B_{jk})}{n(A_{ik}, B_{jk})} = \frac{\binom{m}{1} \binom{n}{1}}{\binom{m + n}{2}} \leq 1 \]

The probability of all the shippers being matched with all the carriers is the probability of the sample space \((A, B)\).

\[ p(A, B) = \binom{m}{m} \binom{n}{n} = 1 \]

The objective here was to prove that the probability \(P(A_{ik}, B_{jk})\) exists, and the sum of probabilities is equal to 1.

The above description of probability is only appropriate if there is no bargaining involved. In this case both shippers and carriers have a bargaining power. This implies that they both can make a decision as to whether accept or reject an offer. In other words, they have a choice. Therefore, instead of the probabilities being simply defined as above they acquire a Logit form.

So, let's define the probability that \(A_{ik}\) and \(B_{jk}\) agree on a deal as:

\[ P(A_{ik}, B_{jk}) \]
Then, three situations are possible. Situation one, is when a shipper accepts the unit costs of a carrier. Let's denote this as:

\[ P(A_{ik} \subseteq B_{jk}) \]

The assumption is that the unit costs are the determining factors in the bargain. These unit costs are a function of many variables that will be discussed later.

The next two cases involve adjustment in the unit costs of transport of the carrier. Situation two is when a carrier accepts the unit costs of a shipper. Let's denote this as:

\[ P(B_{jk} \subseteq A_{ik}) \]

And the last situation arises, when both parties come to an agreement. Let's denote this as:

\[ P(A_{ik} \cap B_{jr}) \]

In each case the probabilities can be defined in terms of a Logit probabilities (see

\[
p(A_{ik} \subseteq B_{jk}) = \frac{\exp (\beta_j U_{B_{jk}})}{\sum_{j' \in B} \exp (\beta_{j'} U_{B_{j'k}})}
\]

where

\[ U_{B_{j'k}} = \text{the utility of a carrier} \]
Appendix A). Case one can be expressed as:

\[ \exp (a_j U_A) \]

Case two is similar to case one except that the utility function used is that of the shipper.

\[ p(A_{ik} \in B_{jk'}) = \frac{\exp (\alpha_j U_{A_{ik}})}{\sum_{A_{i'k} \in A} \exp (\alpha_j U_{A_{i'k}})} \]

where

\[ U_A \] is the utility of a shipper

In case three the assumption is that the utility of a carrier is a function of the utility of the shipper. By this, it is meant that it is a fraction of the utility of the shipper.

\[ \exp (\theta_j f(U_{A_{ik}})) \]

\[ p(A_{ik} \in B_{jk'}) = \frac{\exp (\theta_j f(U_{A_{ik}}))}{\sum_{f(U_{A_{i'k}}) \in B} \exp (\theta_j f(U_{A_{i'k}}))} \]

where

\[ f(U_{A_{ik}}) \] is the utility of a carrier as a function of the utility of the shipper

The outcome is three matrices with rows being the carriers and columns being the shippers. The choice of rows and columns is arbitrary. The cells contain the probabilities of 50% or higher. Once the matrices are set up, the procedure SELECT can be followed (see Appendix E):

**STEP 1:** Form a new matrix by multiplying term by term the probability matrices. In the matrix locate the highest probability. The cell containing this probability is designated as a cell indicating a match between a carrier and a shipper. If there are two probabilities that are exactly the same then choose one. The choice is arbitrary.

**STEP 2:** In the matrix, exclude the row and the column that contains the highest match probability.

**STEP 3:** In the remaining rows and columns designate the cell that yields the next highest probability as the next match.

**STEP 4:** Continue in this manner till all possible matches are obtained.

**STEP 5:** Construct a new matrix with cells either zero or one. One if there was a
match and Zero elsewhere.

**STEP 6:** Find the final most likely set of matches.

It is shown that given a free market, there is always a probability of a match between a shipper, and a carrier. These probabilities are tangible, and can be calculated. Usually the probabilities of at least 50% or higher should be considered. The chosen carriers are the ones using the inter-city waterway system. Once the carriers are identified, the next step is to show how the market mechanism actually works, and how through the market process we can construct a table of ship frequency by type and size per given cargo type.

First, let’s start by looking at the transport problem of goods from the manufacturer’s point of view. The problem is a basic transport problem with an objective function that should be minimized given certain constraints. The reason for looking at the manufacturer’s case is to show that both shippers and carriers consider similar information when making decisions about transporting goods. The second part of the paper will deal with the transport problem from the carrier’s point of view. In this case there are two situations that need to be considered. Situation one is when the carrier considers carrying goods from an origin to a destination. The implication is that in this case the ships are normally loaded with cargo. The second situation is when the carrier picks up the goods at the origin starting from some place other than the origin. The implication in this case is that the ships will be empty from the starting point of the carrier to the origin (where he has to pick up the cargo). Of course in either case after the delivery of cargo, ships return to base empty. I include the cost of return trips in the unit cost of transporting goods to destinations are included.

In the case of the carrier the problem is treated the same as the basic transport problem. The approach will be the same as in the case of the manufacturer. Again, there is an objective function that will be minimized given certain constraints using integer programming. To bridge the gap between the shippers optimize delivery choice and the carriers optimal delivery choice, an Iterative Proportional Fitting approach (IPF) is used. The outcome of this will be an origin-destination matrix with the cells indicating the tons that carriers are willing to carry in different ship sizes given shippers desired level of demand and supply. Knowing maximum and minimum capacities of the ships and total number of ships available, it should be easy to convert this table to a table with cells indicating the frequency of ships by ship type. This is the final outcome of our modeling activities.

One piece of information that is necessary is the unit shipping cost. In the case of the manufacturer it is assumed given. In case of the carrier the unit costs would need to be estimated.
Again two cases can be considered, one is determining the unit costs of shipping goods to the destinations, and the second is the unit cost of picking up the goods from the origins. In case one, a matrix should be set up with rows indicating ships by size and columns representing destinations. In such matrix each cell represents the probability of a carrier wanting to go to a destination, given a ship type/size. Each cell would represent the probability of a carrier wanting to go to an origin to pick up cargo, given a ship type/size. In each case, the probabilities have multinomial Logit form. Once the probabilities are determined, they are then multiplied by general costs in order to produce unit costs of transport. This is justified if the Logit probabilities are considered as restricted domain coefficients of price elasticities.

3.5 A Linear Optimization Problem – Manufacturer’s Case

Let’s assume the following case: A manufacturer (Ajt) of gas cylinders wants to ship them to retailers. He has \( s_1 \) cylinders of transport in stock in one warehouse and another \( s_2 \) cylinders at its second warehouse. The manufacturer has orders for this product from three different retailers, in quantities of \( d_1 \), \( d_2 \), and \( d_3 \) cylinders, respectively. The unit shipping cost \( c_{ij} \) (in cents per cylinder) is the cost of sending cargo from the factories to the retailers. The objective is to determine a minimum shipping cost schedule for satisfying all demands from warehouse \( i = 1,2 \) to retailer \( j = 1,2,3 \). Let \( x_{ij} \) be the number of cylinders (in tons) to be shipped from \( i \) to \( j \). The objective can be written as:

\[
\text{Minimize } Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}
\]

The constraints are:

\[
\begin{align*}
\sum_{j=1}^{n} x_{ij} &\leq s_i \quad (i = 1,\ldots,m) \\
\sum_{i=1}^{m} x_{ij} &\geq d_j \quad (j = 1,\ldots,n)
\end{align*}
\]

Since the total supply \( s_1 + s_2 \) is equal to total demand, \( d_1 + d_2 + d_3 \), each inequality constraint can be tightened to an equality. Doing so, and including the hidden conditions that \( x_{ij} > 0 \), we obtain the mathematical program

\[
\text{Minimize } Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}
\]
Subject to:

\[ \sum_{j} x_{ij} = s_i \quad (i = 1, \ldots, m) \]
\[ \sum_{i} x_{ij} = d_j \quad (j = 1, \ldots, n) \]

With: all variables nonnegative and integral

The above problem is an integer program. The \( x^*_ij \) are the optimal solutions to the integer programming problem. The \( x^*_ij \)s are the cylinders (in tons) that the manufacturer has decided to sent to the retailers. The decision is based on minimizing his cost.

### 3.6 Carrier's Case - Transporting Cargo To Destinations

**Case I:** Carrier is located at the same general area of the warehouses. Therefore, there is no cost associated with picking-up cargo from the warehouse. Graph 1 below shows the general form of

![Graph 1: Depicting the travel pattern of a carrier](image)

a trip made by a carrier.

Graph 1. Depicting the travel pattern of a carrier

The intention is to use the same methodology as above. Here, the carrier (\( B_{kj} \)) has to minimize the cost of carrying cylinders to destination \( j \). For the sake of demonstrating the approach it is assumed that the unit costs of transporting goods to destinations are given.
Later on a method for estimating the unit cost of transporting goods (carrier’s case) will be given. Assume that the unit costs (cents) of carrying goods to destinations is denoted by $c_{kj}^*$. Here $k$ is the number of ship sizes; $j$ still represents destinations. The minimum and maximum, $(L_{k,\text{min}}, L_{k,\text{max}})$ capacities of different tanker sizes are also given. It is assumed that the carrier would not consider using a particular tanker size unless it is at least filled by a minimum capacity. It is assumed that carriers do not wish to make trips to destinations empty. The maximum carrying capacity respectively is the maximum load that the carrier would consider per tanker size, and would not exceed it no matter what price the manufacturer would offer. The objective is to minimize the cost of carrying cylinders to destinations $j$. Let $x_{kj}^*$ be the cylinders (tons) carried by tanker size $k$ to destination $j$ that satisfies the minimum cost requirement.

$$\text{Minimize } Z = \sum_{k} \sum_{j} c_{kj}^* x_{kj}^*$$

Subject to:

$$\sum_{j} x_{kj}^* \geq (\sum_{j} n_j) L_{k,\text{min}} \quad (k = 1, \ldots, l)$$

$$\sum_{j} x_{kj}^* \leq (\sum_{j} n_j) L_{k,\text{max}} \quad (k = 1, \ldots, l)$$

$$\sum_{k} x_{kj}^* \geq d_j \quad (i = 1, \ldots, m)$$

The constraints are subject to minimum and maximum capacities, and demand levels. All variables are nonnegative. The dual simplex method can be used to solve the above integer-programming problem.
3.7 Carrier’s Case- Picking-Up Cargo From Origins

In this case the carrier has to pick up cargo from the origin. The carrier is located in a place other than the origin. He then has to travel with an empty tanker/tankers from where he is located (his base) to the origin to pick up cargo. In this case the cost is associated with traveling with an empty tanker to the origin, as well as transporting cargo to a destination. In this case the cost of transporting cargo to destinations is of secondary concern. If we choose a weighting factor, then the trip to origin in this case weighs higher than the trip to the destination. The determining factor is whether the carrier would pick-up the cargo at the origin. The unit costs in this case are similarly calculated as the previous case, except for travel times and associated costs. Graph 2 below illustrates this concept. The objective is to minimize the cost of picking-up cargo from the origin. Assume that the minimum and maximum capacities are the same as before. Graph 2 below shows the general form of a trip made by a carrier.

![Graph 2](image)

Graph 2. The routing pattern of a carrier- Case where the pick-up location is different from the origin.

The unit costs in cents are given $c_{kj}^o$. The objective is to minimize the cost of picking-up cargo from the origin (warehouse). Let $x_{ki}^o$ be the cylinders (tons) picked up by tanker size $k$ from origin $i$ that satisfies the minimum cost requirement.

\[
\begin{align*}
\text{Minimize} & \quad Z = \sum_{k} \sum_{j} c_{kj}^o x_{ki}^o \\
\text{Subject to:} & \quad \sum_{k} x_{ki}^o \geq (\sum_{i} n_i) L_{k_{min}} \quad (k = 1, \ldots, l)
\end{align*}
\]
A freight traffic modeling approach for the waterway network in the Netherlands

\[ \sum_{k=1}^{m} x_{ki} \leq \left( \sum_{i=1}^{m} n_i \right) L_{k,\text{max}} \quad (k = 1, \ldots, l) \]

\[ \sum_{k=1}^{l} x_{ki} \leq s_i \quad (i = 1, \ldots, m) \]

The dual simplex method can be used to solve the above integer-programming problem.

3.8 Transforming Tons Carried To Number Of Ships

The objective here is to adjust the tons carried by different ship sizes as is calculated by the carrier with total supply and demand provided by the shipper. To achieve this aim, a three dimensional Iterative Proportional Fitting (IPF) procedure can be used. IPF is a straight-forward and self-checking procedure that makes adjustments in data obtained through sampling in order to achieve consistency with data obtained from other sources or with deductions from established theory. IPF can be used to estimate the cross-tabulation of the individuals by certain characteristics where a cross-tabulation of only a sample of the individuals by each characteristic is available and the distribution of the total number of individuals by each characteristic (marginal totals) is known. The IPF method allows adjustments in the cell frequencies of the sample data to estimate cell frequencies of the population data by matching the marginal totals. Therefore, the minimum cost shipping schedule obtained as a result of performing optimization can be adjusted in order to achieve consistency with total demand and supply provided by the shipper. The final step is to transform the adjusted matrix of tons to number of ships using minimum and maximum capacities and total number of ships available (see Appendix G).
3.9 Determination Of Unit Costs Of Shipping Cargo To Destinations – A Multinomial Logit Approach

3.9.1 Calculation Of Unit Cost Of Carriers

The utility function \( V_j \) of transporting goods to destinations is equal to the total cost. This cost depends on costs due to the fleet’s combination of ship types and sizes, and operating costs. For example in the case of inland shipping the potential cost variables are:

- Depreciation costs (usage costs) (euro/year)
- Personnel costs (euro/hr)
- Base tariff (euro)
- Cost of waiting at the docks to be loaded (euro/hr)
- Cost of waiting idle at the docks
- Gasoline cost (euro/km)

Travel time = Travel time from base to the origin (hrs) + Total travel time per given path (hrs) + total travel time of the return trip (hrs)

Travel times

Note: The variable travel time should be broken down into several elements: 1) trip from base to an origin (that is if base differs from the origin), 2) trip from origin to destination, 3) and the return trip. Each segment of the trip should be multiplied by an attractiveness factor. This factor determines the significance of each portion of the trip in the overall path assignment.

The utility function per ship size \( k \) going to destination \( j \) \( (V_{kj}) \) can be written as a linear combination of the above cost elements. Let \( P_{kj} \) be the probability of using ship size \( k \) going to destination \( j \). Then \( P_{kj} \) is a Logit probability of choosing ship size \( k \) to carry cargo to destination \( j \) (see Appendix A).

\[
P_{kj} = \frac{\exp(\beta V_{kj})}{\sum_h \exp(\beta V_{hj})}
\]

It is possible to make an analogy between the Logit probabilities above and the coefficient of price elasticity in economics. In the economic sense a price elasticity of a product \( e \) measures the percentage change in the quantity of a commodity demanded per unit of time resulting from a given percentage change in the price of the commodity.
\[ e = - \frac{\Delta Q}{Q} \times \frac{\Delta P}{P} \]

Q represents the quantity of a commodity. P represents the price. The sign \( \Delta \) represents a change. Demand is said to be elastic if \( e > 1 \), inelastic if \( e < 1 \), and unitary elastic if \( e = 1 \). The Logit probabilities are quantities between zero and one. By analogy they could represent the price elasticity in the region where the elasticities are either absolutely inelastic \( P_{kj} = 0 \), or inelastic \( 0 < P_{kj} < 1 \), or unitary elastic \( P_{kj} = 1 \). That means the Logit probabilities show that portion of the price charged by the carrier that should be acceptable by the shipper; anything outside of this price would cause a reduction in demand for the carrier. Therefore, the cost of transporting goods multiplied by the Logit probability gives that portion of the cost for which the carrier can charge the shipper for, which would not induce a reverse reaction on the side of the shipper.

where

\[ C_{kj} = V_{kj} \]

\( c_{kj} \) is the quantity that the carrier would use as a unit cost or unit price charged to the shipper. The reason for this modification also stems from the idea that the Logit probabilities indicate the degree of attractiveness of using an option. The idea implies that the Logit probabilities show the extent to which the unit costs should be modified to reflect that portion of the unit costs that will stay inelastic with respect to changes in total costs.

3.10 Overall Algorithmic Approach Of The Above Methodology

The overall algorithmic approach to the problem is as follows:

3.11 Case Study

A Linear Optimization Problem – Manufacturer's Case

Let's assume the following case: A manufacturer of gas cylinders wants to ship them to retailers. He has 1200 cylinders of transport wrap in stock in one warehouse and another 1000 cylinders at its second warehouse. The manufacturer has orders for this product from three different retailers, in quantities of 1000, 700, and 500 cylinders, respectively.
The unit

Shipping cost (in cents per cylinder) from the factories to the retailers is as follows:

<table>
<thead>
<tr>
<th></th>
<th>Retailer 1</th>
<th>Retailer 2</th>
<th>Retailer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factory 1</td>
<td>14</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>Factory 2</td>
<td>13</td>
<td>13</td>
<td>12</td>
</tr>
</tbody>
</table>

The objective is to determine a minimum shipping cost schedule for satisfying all demands from warehouse \( i = 1,2 \) to retailer \( j = 1,2,3 \). Let \( x_{ij} \) be the number of cylinders to be shipped from \( i \) to \( j \). The objective is:

\[
\text{Minimize } Z = 14x_{11} + 13x_{12} + 11x_{13} + 13x_{21} + 13x_{22} + 12x_{23}
\]

The constraints are:

\[
\begin{align*}
x_{11} + x_{12} + x_{13} &< 1200 \\
x_{21} + x_{22} + x_{23} &< 1000
\end{align*}
\]

also:

\[
\begin{align*}
x_{11} + x_{21} &> 1000 \\
x_{12} + x_{22} &> 700 \\
x_{13} + x_{23} &> 500
\end{align*}
\]

Since the total supply 1200 +1000 is equal to total demand, 1000 + 700 + 500, each inequality constraint can be tightened to an equality. Doing so, and including the hidden conditions that \( x_{ij} > 0 \), we obtain the mathematical program

\[
\text{Minimize: } Z = 14x_{11} + 13x_{12} + 11x_{13} + 13x_{21} + 13x_{22} + 12x_{23}
\]

Subject to:

\[
\begin{align*}
x_{11} + x_{12} + x_{13} &= 1200 \\
x_{21} + x_{22} + x_{23} &= 1000 \\
x_{11} + x_{21} &= 1000 \\
x_{12} + x_{22} &= 700 \\
x_{13} + x_{23} &= 500
\end{align*}
\]

With: all variables nonnegative and integral
The above problem is an integer program. For the sake of simplicity, we can rewrite the above problem as follows:

Minimize: \[ Z = 14x_1 + 13x_2 + 11x_3 + 13x_4 + 13x_5 + 12x_6 \]

Subject to:

\[ x_1 + x_2 + x_3 = 1200 \]
\[ x_4 + x_5 + x_6 = 1000 \]
\[ x_1 + x_4 = 1000 \]
\[ x_2 + x_5 = 700 \]
\[ x_3 + x_6 = 500 \]

The outcome of integer programming approach is shown below:

<table>
<thead>
<tr>
<th></th>
<th>Retailer 1</th>
<th>Retailer 2</th>
<th>Retailer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factory 1</td>
<td>0</td>
<td>700</td>
<td>500</td>
</tr>
<tr>
<td>Factory 2</td>
<td>1000</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

As you can see the row totals are 1200 and 1000 respectively. The column totals adds up to 1000, 700, 500 respectively. The sum total is 2200.

In the canal flow problem this matrix is given. It is the OD matrix by tons.

3.11.1 Carrier's Case- Transporting Cargo To Destinations

Case 1: Carrier is located at the same general area of the warehouses. Therefore, there is not cost associated with picking-up cargo from the warehouse. The graph below shows the general form of a trip made by a carrier.
the intention is to use the same methodology as above. Here, the carrier has to minimize the cost
of carrying cylinders to destination j. For the sake of demonstrating the approach let’s assume
that the unit cost of transporting goods to destinations is given. Later on a method for estimating
the unit cost of transporting goods (carrier’s case) will be provided. Normally, this information
should be estimated.

Assume that the unit costs (cents) of carrying goods to destinations is given in a table below:

<table>
<thead>
<tr>
<th>Shipsize/destinations</th>
<th>Retailer 1</th>
<th>Retailer 2</th>
<th>Retailer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small tankers</td>
<td>8</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Medium tankers</td>
<td>15</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>Large tankers</td>
<td>23</td>
<td>25</td>
<td>35</td>
</tr>
</tbody>
</table>

Assume that the capacities of different tanker sizes are given: note that the minimum carrying
capacity signifies that the carrier would not consider using a particular tanker size unless it is at
least filled by a minimum load. Assume that carriers do not wish to make trips to destinations
empty. The maximum carrying capacity respectively is the maximum load that the carrier would
consider per tanker size, and would not exceed it no matter what the manufacturer would offer.

<table>
<thead>
<tr>
<th></th>
<th>Minimum capacity(tons)</th>
<th>Maximum capacity(tons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small tankers</td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td>Medium tankers</td>
<td>501</td>
<td>700</td>
</tr>
<tr>
<td>Large tankers</td>
<td>701</td>
<td>1000</td>
</tr>
</tbody>
</table>

The objective is to minimize the cost of carrying cylinders to destination j. Let \( x_{kj} \) be the
cylinders (tons) carried by tanker size k to destination j.

Minimize \( z = 8x_{11} + 15 x_{12} + 7 x_{13} + 15 x_{21} + 18 x_{22} + 17 x_{23} + 23 x_{31} + 25 x_{32} + 35 x_{33} \)

Subject to:

\[
\begin{align*}
x_{11} + x_{12} + x_{13} & \geq 3 \times 100 \\
x_{11} + x_{12} + x_{13} & \leq 3 \times 500 \\
x_{21} + x_{22} + x_{23} & \geq 3 \times 501 \\
x_{21} + x_{22} + x_{23} & \leq 3 \times 700 \\
x_{31} + x_{32} + x_{33} & \geq 3 \times 701 \\
\end{align*}
\]
A freight traffic modeling approach for the waterway network in the Netherlands

\[ x_{31} + x_{32} + x_{33} \leq 3 \times 1000 \]
\[ x_{11} + x_{21} + x_{31} \geq 1000 \]
\[ x_{12} + x_{22} + x_{32} \geq 700 \]
\[ x_{13} + x_{23} + x_{33} \geq 500 \]

Constraints are subject to minimum and maximum capacity, and demand levels. Here the capacity is not per trip but are considered as intrinsic characteristic of a ship.

For the sake of simplicity, we can rewrite the above problem as follows:

Minimize:  
\[ z = 8x_1 + 15x_2 + 7x_3 + 15x_4 + 18x_5 + 17x_6 + 23x_7 + 25x_8 + 35x_9 \]

Subject to:
\[ x_1 + x_2 + x_3 \geq 3 \times 100 \]
\[ x_1 + x_2 + x_3 \leq 3 \times 500 \]
\[ x_4 + x_5 + x_6 \geq 3 \times 501 \]
\[ x_4 + x_5 + x_6 \leq 3 \times 700 \]
\[ x_7 + x_8 + x_9 \geq 3 \times 701 \]
\[ x_7 + x_8 + x_9 \leq 3 \times 1000 \]
\[ x_1 + x_2 + x_3 \geq 1000 \]
\[ x_4 + x_5 + x_6 \geq 700 \]
\[ x_7 + x_8 + x_9 \geq 500 \]

Expressing all the constraints in the \( \leq \) form and adding the slack variables, the problem becomes:

Minimize:  
\[ z = 8x_1 + 15x_2 + 7x_3 + 15x_4 + 18x_5 + 17x_6 + 23x_7 + 25x_8 + 35x_9 + 0x_{10} + 0x_{11} + 0x_{12} + 0x_{13} + 0x_{14} + 0x_{15} + 0x_{16} + 0x_{17} + 0x_{18} \]
Subject to:  
\[-x_1 - x_2 - x_3 + x_{10} = -300\]
\[x_1 + x_2 + x_3 + x_{11} = 1500\]
\[-x_4 - x_5 - x_6 + x_{12} = -1503\]
\[x_4 + x_5 + x_6 + x_{13} = 2100\]
\[-x_7 - x_8 - x_9 + x_{14} = -2103\]
\[x_7 + x_8 + x_9 + x_{15} = 3000\]
\[x_1 + x_2 + x_3 + x_{16} = -1000\]
\[x_4 + x_5 + x_6 + x_{17} = -700\]
\[x_7 + x_8 + x_9 + x_{18} = -500\]

With: all variables nonnegative

Use the dual simplex method to solve the above integer-programming problem. Let’s assume that the following table is the feasible optimal solution to the above problem.

<table>
<thead>
<tr>
<th></th>
<th>Retailer 1</th>
<th>Retailer 2</th>
<th>Retailer 3</th>
<th>Row sums</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small tanker</td>
<td>200 (tons)</td>
<td>50</td>
<td>100</td>
<td>350</td>
</tr>
<tr>
<td>Medium tanker</td>
<td>300</td>
<td>100</td>
<td>150</td>
<td>550</td>
</tr>
<tr>
<td>Large tanker</td>
<td>500</td>
<td>550</td>
<td>250</td>
<td>1300</td>
</tr>
<tr>
<td>Column sums</td>
<td>1000</td>
<td>700</td>
<td>500</td>
<td>2200</td>
</tr>
</tbody>
</table>

3.11.2 Carrier’s Case- Picking-Up Cargo From Origins

In this case the carrier has to pick up cylinders from the origin. There are two situations to consider. One is that the carrier is located at the origin. The second case is when the carrier is located in a place other than the origin. He then has to travel with an empty tanker from where he is located (base) to the origin to pick up cargo. In this case the cost is associated
A freight traffic modeling approach for the waterway network in the Netherlands

Graph 1. The routing pattern of a carrier-case where the pick-up location is different from the origin.

with traveling with an empty tanker. In both cases after the carrier drops the cargo at the destination, he has to return home. Graph 1 above illustrates this concept. There is a cost associated with the journey back home. Again let’s go through an example. The objective is to minimize cost of picking-up cargo from the origin. Assume that the minimum and maximum capacities are the same as before.

The unit costs in cents are given in a table below:

<table>
<thead>
<tr>
<th></th>
<th>Small tanker</th>
<th>Medium tanker</th>
<th>Large tanker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warehouse 1</td>
<td>6</td>
<td>12</td>
<td>28</td>
</tr>
<tr>
<td>Warehouse 2</td>
<td>8</td>
<td>15</td>
<td>32</td>
</tr>
</tbody>
</table>

The objective is to minimize the cost of picking-up cargo from the origin (warehouse). Let $x_{ik}$ be the cylinders (tons) picked up by tanker size $k$ from origin $i$.

Minimize $z = 6x_{i1} + 12x_{i2} + 28x_{i3} + 8x_{i21} + 15x_{i22} + 32x_{i23} + 0x_{i21} + 0x_{i22} + 0x_{i23}$

Subject to:

$x_{i1} + x_{i2} + x_{i3} \leq 1200$

$x_{i21} + x_{i22} + x_{i23} \geq 1000$

$x_{i1} + x_{i21} \geq 3 \times 100$

$x_{i1} + x_{i21} \leq 3 \times 500$

$x_{i2} + x_{i22} \geq 3 \times 501$

$x_{i2} + x_{i22} \leq 3 \times 700$

$x_{i3} + x_{i23} \geq 3 \times 701$

$x_{i3} + x_{i23} \leq 3 \times 1000$
A freight traffic modeling approach for the waterway network in the Netherlands

For the sake of simplicity, we can rewrite the above problem as follows:

Minimize:  \[ z = 6x_1 + 12x_2 + 28x_3 + 8x_4 + 15x_5 + 32x_6 \]

Subject to:
\[
\begin{align*}
x_1 + x_4 & \geq 3 \times 100 \\
x_1 + x_4 & \leq 3 \times 500 \\
x_3 + x_5 & \geq 3 \times 501 \\
x_2 + x_5 & \leq 3 \times 700 \\
x_3 + x_6 & \geq 3 \times 701 \\
x_3 + x_6 & \leq 3 \times 1000 \\
x_1 + x_2 + x_3 & \leq 1200 \\
x_4 + x_5 + x_6 & \geq 1000
\end{align*}
\]

Expressing all the constraints in the \( \leq \) form and adding the slack variables, the problem becomes:

Minimize:  \[ z = 6x_1 + 12x_2 + 28x_3 + 8x_4 + 15x_5 + 32x_6 + 0x_7 + 0x_8 + 0x_9 \\
0x_{10} + 0x_{11} + 0x_{12} + 0x_{13} + 0x_{14} \]

Subject to:
\[
\begin{align*}
x_1 - x_4 + x_7 & = -300 \\
x_1 + x_4 + x_8 & = 1500 \\
x_2 - x_3 + x_9 & = -1503 \\
x_2 + x_5 + x_{10} & = 2100 \\
x_3 - x_6 + x_{11} & = -2103 \\
x_3 + x_6 + x_{12} & = 3000 \\
x_1 + x_2 + x_3 + x_{13} & = 1000 \\
x_4 + x_5 + x_6 + x_{14} & = 700
\end{align*}
\]
Use the dual simplex method to solve the above integer-programming problem. Let’s assume that the following table is the feasible optimal solution to the above problem.

<table>
<thead>
<tr>
<th></th>
<th>Small tanker</th>
<th>Medium tanker</th>
<th>Large tanker</th>
<th>Row sums</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warehouse 1</td>
<td>200 (tons)</td>
<td>50</td>
<td>950</td>
<td>1200</td>
</tr>
<tr>
<td>Warehouse 2</td>
<td>150</td>
<td>500</td>
<td>350</td>
<td>1000</td>
</tr>
<tr>
<td>Column sums</td>
<td>350</td>
<td>550</td>
<td>1300</td>
<td>2200</td>
</tr>
</tbody>
</table>

### 3.11.3 Transforming Tons Carried To Number Of Ships

The objective here is to use the information that I have derived previously in setting up an origin-destination matrix with the cell entries indicating the split of tons carried by different tanker sizes. In this case they’re only three sizes: small, medium and large. To achieve this aim, a three dimensional Iterative Proportional Fitting (IPF) procedure is used. The principle idea behind the IPF procedure is to adjust the data cell values such that the expected value of the sum of the squares of the differences between the adjusted and estimated values is minimized.

In this case assume that the cargo has the following characteristics. Each cargo has one of two origins (warehouse 1, warehouse 2), it has one of the three destinations (retailer 1, retailer 2, retailer 3), and it is carried in one of three tanker sizes (small, medium, large). The totals for each categories are given: the row totals are 1200, and 1000 respectively; The column totals are 1000, 700, 500 respectively. Depth totals signify the number of tanker sizes by their load totals. Depth totals are equal to totals of case one if the carrier is located around the warehouse, and is equal to the totals of second case only if he has to make a trip to the warehouse to pick-up cargo. Let’s consider case 1. The depth totals are 350, 550 and 1300 respectively.

The structure of the problem is as follows:

\[
\begin{align*}
X_{111} & \quad X_{121} & \quad X_{131} \\
X_{211} & \quad X_{221} & \quad X_{231} \\
X_{112} & \quad X_{122} & \quad X_{132} \\
X_{212} & \quad X_{222} & \quad X_{232} \\
X_{113} & \quad X_{123} & \quad X_{133} \\
X_{213} & \quad X_{223} & \quad X_{233}
\end{align*}
\]
A freight traffic modeling approach for the waterway network in the Netherlands

row totals:

<table>
<thead>
<tr>
<th></th>
<th>1200</th>
<th>1000</th>
</tr>
</thead>
</table>

Column totals:

<table>
<thead>
<tr>
<th></th>
<th>1000</th>
<th>700</th>
<th>500</th>
</tr>
</thead>
</table>

Depth totals:

<table>
<thead>
<tr>
<th></th>
<th>350</th>
<th>550</th>
<th>1300</th>
</tr>
</thead>
</table>

$x_{ik}$ is the cargo carried from $i =$ (warehouse 1, warehouse 2) to destination $j =$ (retailer 1, retailer 2, retailer 3) by tanker size $k =$ (small, medium, large). In this case the distribution of the total number of cargo (tons) by each characteristic (origin, destination, ship size), (marginal totals) is known. As initial values for the cells in the block matrix above, one possibility is to use some historical average loads carried by carriers to destinations in (tons) per ship size. Special case is when the shipper does not want to ship cargo to a particular destination; then those cells should be zero. So the initial matrix will look like this:

<table>
<thead>
<tr>
<th></th>
<th>100</th>
<th>25</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
<td>25</td>
<td>50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>150</th>
<th>50</th>
<th>75</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>150</td>
<td>50</td>
<td>75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>250</th>
<th>275</th>
<th>125</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>250</td>
<td>275</td>
<td>125</td>
</tr>
</tbody>
</table>
The cell values are then adjusted by the following totals:

1200
1000
1000
700
500
350
550
1300

Again for the sake of demonstration, let’s assume that the solution matrix is obtained given the above-described method. Let’s assume the matrix is as follows: The first block of the matrix shows the tons carried from warehouse 1 by different tanker sizes to the three retailers. The second block matrix shows the same information from warehouse 2.

\[
\begin{matrix}
0 & 0 & 700 \\
100 & 0 & 0 \\
50 & 250 & 0 \\
0 & 250 & 0 \\
850 & 0 & 0 \\
\end{matrix}
\]

From the above matrix we can deduce the following: origin destination matrix by ship type.

<table>
<thead>
<tr>
<th></th>
<th>Retailer1</th>
<th>Retailer2</th>
<th>Retailer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warehouse 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Warehouse 2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Here each cell represents the number of ships given tons carried. As can be noticed from this made up example that not all tankers will travel with some load. If the cell entry for the block matrix of tons above exceeds the maximum allowable load for a tanker, then assume that two or more tankers of the same size are used. The matrix that should go into an assignment and trip estimation is an aggregate of the above block matrix. The matrix should look like:

<table>
<thead>
<tr>
<th>Warehouse 1</th>
<th>Retailer 1</th>
<th>Retailer 2</th>
<th>Retailer 3</th>
<th>Row sums</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 (tankers)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Warehouse 2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Column sums</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>6 (total)</td>
</tr>
</tbody>
</table>

The matrix above is the final outcome of the procedure that I have suggested.

3.11.4 Calculation Of Unit Cost Of Carriers

The utility function in the case of transporting cargo to destination should be the function of the following attributes $x$: travel time (translated into $\$), tanker size (loaded)($), gasoline cost ($), cost of return trip to base empty ($). The utility function is a linear function that would look as follows:

$$V_j = \theta_1 x_{1j} + \theta_2 x_{2j} + \theta_3 x_{3j} + \theta_4 x_{4j}$$

$V_j$ is the utility of selecting $j$

$\theta_i$ are the coefficients

$x_{ij}$ are the travel times on the paths ($\$)$

$x_{2j}$ is the tanker size ($\$)$

$x_{3j}$ is the cost of return trip to base ($\$)$

In the case where carrier starts from base and has to travel to the origin, the utility function is the same linear function except that there is the added cost of traveling from base to the origin empty.

The probabilities represent the probability of choosing a tanker size $k$ to go to retailer $i$.

$$p_{ki} = \frac{\exp(\beta V_{ki})}{\sum_{h,j} \exp(\beta V_{hj})}$$
Once all the probabilities are calculated, the following table can be constructed:

<table>
<thead>
<tr>
<th></th>
<th>Retailer 1</th>
<th>Retailer 2</th>
<th>Retailer 3</th>
<th>Row Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small tanker</td>
<td>.30</td>
<td>.10</td>
<td>.10</td>
<td>.50</td>
</tr>
<tr>
<td>Medium tanker</td>
<td>.20</td>
<td>.10</td>
<td>.0</td>
<td>.30</td>
</tr>
<tr>
<td>Large tanker</td>
<td>.10</td>
<td>.10</td>
<td>.0</td>
<td>.20</td>
</tr>
<tr>
<td>Column Sums</td>
<td>.60</td>
<td>.30</td>
<td>.10</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Keep in mind that the numbers are made up for the sake of finishing the explanation. Let's assume that the total cost $C_{qj}$ of carrying cargo to destinations has been obtained from the carriers. $q$ represents carrier $q$, and $j$ represents destination $j$. The unit cost of carrying cargo to destination $j$ is obtained by:

$$c_{kj} = V_{kj} \times p_{kj}$$

$k$ is the tanker size (small, medium, large), and $j$ is the destination (retailer1, retailer2, retailer3). The table of unit costs will look as follow:

<table>
<thead>
<tr>
<th></th>
<th>Retailer 1</th>
<th>Retailer 2</th>
<th>Retailer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small tanker</td>
<td>8.0</td>
<td>1.8</td>
<td>1.5</td>
</tr>
<tr>
<td>Medium tanker</td>
<td>9.6</td>
<td>20.4</td>
<td>13.8</td>
</tr>
<tr>
<td>Large tanker</td>
<td>6.4</td>
<td>13.6</td>
<td>11.6</td>
</tr>
</tbody>
</table>

### 3.11.5 CONCLUSION

As was mentioned at the beginning of this chapter, the aim has been to introduce a procedure for calculating the number of ships of type/size per pre-designated origin-destination pairs. The general approach has two levels: a probabilistic level, and a deterministic level. Logit probabilities are used to make decisions on market behavior. They are also used to determine unit costs of transport on the carrier's side. Except that in this case the Logit probabilities can be compared to the coefficient of price elasticities. Desirable loads per ship type/size and origin-destination pairs are determined through an application of an optimization method. An IPF method is used to redistribute loads among ship sizes based on total demand, total supply, and total desirable load per ship type/size. Two cases have been considered: Case one is the case of carrying cargo to destinations from a warehouse where the only empty trip is the return trip.
The second case is when the carrier has to travel to the warehouse. In this case there are at least two empty trips, one going to the origin and the return trip.

The chapter does not consider chained trips where a destination could also be a place where ships reload for another destination. Trip assignment is considered to be exogenous to the model. Therefore, it is not covered in the scope of the chapter. This chapter presents a few new ideas. Both the shippers and the carriers are active participants in the market. They both can make decisions that can affect the market. A method is devised to model the behavior of shippers and carriers in the market place. The idea of equating Logit probabilities to the coefficients of elasticities in a limited way seems very useful in determining the unit costs of transportation. Future efforts will be in expanding the ideas presented in the paper.
4. DATA SPECIFICATIONS
A freight traffic modeling approach for the waterway network in the Netherlands
4. DATA SPECIFICATIONS

4.1 Introduction

This chapter is concerned with the specifics of data collection and data specification. The main problem in data collection techniques is inaccuracy. Many problems occur in modeling and forecasting due to errors in data collection procedures. The procedures normally used in modeling assume not only that the correct functional specification of the model is known a priori, but also that the data used to estimate the model parameters have no errors. In practice, seldom these conditions are satisfied; furthermore, even if they were satisfied, model forecasts are usually subject to errors due to inaccuracies in the values assumed for the explanatory variables in the design year. The ultimate goal of modeling is often forecasting; an important problem model designers face is to find which combination of model complexity and data accuracy fits best the required forecasting precision and study budget. To this end, it is important to distinguish between different types of errors, in particular:

- Those that could cause even correct models to yield incorrect forecasts, e.g. errors in the prediction variables of the explanatory variables, transference and aggregation errors; and

- Those that actually cause incorrect models to be estimated, e.g. measurement, sampling and specification errors.

4.1.1 Measurement errors

These occur due to the inaccuracies inherent in the process of actually measuring the data in the base year, such as network measurement errors, coding and digitizing errors, and so on. There are many aspects to the data that can be used in modeling the waterway freight traffic on the canals. Aside from network data, there is a need for collecting data from shippers as well as carriers. What relates to shippers data is explanatory variables such as:

- Supply levels at sources. This assuming that the producers have cargo storage houses at harbors (sources) from which they ship cargo;

- Demand levels at destinations. The destinations are usually places like retailers and any kind of place that uses the cargo;

- Any shipper has an estimate of the cost of transporting goods from his origin to his pre-assigned destinations. This cost is a function of path choice, and travel times on the path. Other relevant cost related explanatory variables are: Fuel consumption (expenses), cost of
providing containers for the cargo (packaging costs), tariffs that should be paid to government at different points on the way to destinations, insurance cost in case of damage to the unit during the transport period, the price shipper is willing to offer carrier($ per hour or per mile).

The measurement errors that occur here are mainly due to the bad interpretation of the questions by the interviewee, and bad registration of answers by the interviewer(s). Another point of consideration is that only those variables should be included that do not violate the privacy act of the country.

Variables related to carriers are:

The carriers incur two types of costs: the non-transportation costs, and the costs that relate to transportation. The non-transportation costs include but not limited to the following:

- Personnel costs; these costs relate to the crew required to load and unload cargo, and maintain the vessels (tankers, and barges);
- Cost relating to the up-keep of the vessels such as; parts and paints and other materials;
- Cost of living;
- Insurance costs; in case of accidents, and theft of material and parts.

The transportation costs are:

- Cost of choosing a path, This cost is a function of travel times on the chosen path;
- Cost of traveling empty to an origin to pick-up cargo;
- Cost of not travelling at full capacity;
- Cost of traveling empty back from a destination;
- Cost of waiting to be loaded;
- Cost of waiting at the harbor before loading starts; case there is a queue;
- Fuel consumption ($ expenses);
- Depreciation costs;
- Anticipated profit; this variable is difficult to measure; but it is possible to estimate an average number by looking at historical data; or just by analyzing the market. There are specific guidelines to make the task of estimating profits easy; Appendix C.

Measurement error, as defined here, should be distinguished from the difficulty of defining the variables that ought to be measured. Ideally modeling should be based on the information perceived by individual users of the system but whilst the reported data may give some insight into perception, its use raises the difficult question of how to forecast what the system users are
going to perceive in the future. So it appears inevitable that models will be endowed with perception errors which tend to be greater for non-chosen alternatives due to the existence of self-selectivity bias (i.e. the attributes of the chosen option are perceived as better and those of the rejected option as worse than they are, such as to reinforce the rationality of the choice made).

![Figure 4.1 Attribute of Measurement Error to Choice](image)

**4.7.2 Specification Errors**

These errors arise either because the phenomenon being modeled is not well understood or because it needs to be simplified in excess for whatever reason. Important subclasses of this type of error are the following:

- Inclusion of an irrelevant variable (i.e. one which does not affect the modeled choice process). This error will not bias the model (or its forecasts) if the parameters appear in linear form, but it will tend to increase sampling error, in a non-linear model, however, bias may be caused Tardiff 1979.
- Omission of a relevant variable; perhaps the most common specification error. Interestingly, models incorporating a random error term are designed to accommodate this error; however, problems can arise when the excluded variable is correlated with variables in the model or when its distribution in the relevant population is different from its distribution in the sample used for model estimation Horowitz 1981b.
Exclusion of taste variations on the part of the shippers and carriers; this will always produce biased models. Unfortunately, this is the case in most practical models of choice; a notable exception is the generally unyielding multinomial probit models Daganzo 1979.

Other specification errors, in particular the use of model forms that are not appropriate, such as linear functions to represent non-linear effects, compensatory models to represent behavior that might be non-compensatory, or the omission of effects such as habit or inertia Williams and Ortuzar 1982a.

All specification errors can be reduced in principle simply by increasing model complexity. However, the total costs of doing this are not easy to estimate as they relate to model operation, but may induce other types of errors which might be costly or impossible to eliminate (e.g. when forecasting more variables and at higher level of disaggregation). Moreover, removal of some specification errors may require extensive behavioral research and in some cases it must simply be conceded that such errors may be present in all feasible models.

4.7.3 Transfer errors

These occur when a model developed in one context (time and / or place) is applied in a different one. Although adjustments may be made to compensate for the transfer, ultimately the fact must be faced that behavior might just be different in different contexts. In the case of spatial transfers, the errors can be reduced or eliminated by partial or complete re-estimation of the model to the new context (although to do the latter would imply discarding the substantial cost savings obtainable from transfer). However, in the case of temporal transfer (i.e. forecasting), this re-estimation is not possible and any potential errors must just be accepted.

4.7.4 Aggregation Errors

These arise basically out of the need to make forecasts for groups of people while modeling often needs to be done at the level of individual in order to capture behavior better. The following are important sub-classes of aggregation error:

- Data aggregation. In most practical studies the data used to define the choice situation of individual carriers is aggregated in some form or another. Even when the carrier is asked to report the characteristics of his/her available options, they can only have based their choice on the expected values of these characteristics. Models estimated with aggregate data will suffer from some form of specification error Daly and Ortuzar 1989.
Aggregation of alternatives. Again due to practical considerations it may just not be feasible to attempt to consider the whole range of options available to each shipper or carrier.

Model aggregation. This can cause severe difficulties to the analyst except in the case of linear models where it is a trivial problem. Aggregate quantities such as flow on waterways links are a basic modeling result in freight transport planning, but methods to obtain them are subject to aggregation errors which are often impossible to eliminate.

4.7.5 Data Specification

The data available for this project has the following form: van der Zijpp et al., 1999

4.1.6 Network data

The network specification describes the waterway system as a set nodes and links that connect the nodes together. Each link is distinguished by: an a-node, b-node, length, width, depth, and type. Link types are broken into: waterways, bridges, and locks. The nodes represent either an origin, or a destination or both. They have geographical coordinates to indicate their position. Each link has additional characteristics such as:

- Operating period: these are times when the links are operational;
- Traffic volume on the links; this include all waterway traffic such as recreational vessels;
- Seasonal flow speed: is the speed on the links during the operational periods;
- Each link is classified by its Seasonal depth;
- Each link is specified by maximum allowed ship width;
- Maximum allowed kegels (cones); this has to do with vessels carrying hazardous material;
- Maximum allowed speed full;
- Maximum allowed speed empty;
- Maximum length of ship that can make u turn on the link;
- Capacity in ship equivalents per hour;
- Overtaking allowed or not;
- Maximum allowed ship equivalents;
- Minimum required ship class;
- Maximum allowed ship class;
- Seasonal recreational vessel load allowed;
- Fixed travel delay;
- Each link either has a counter or not.
Bridges have the following characteristics:

- operational period;
- maximum and minimum height of the bridge;
- time required to open the bridge;
- number of directions allowed (a-b, or b-a, or (a-b and b-a));
- time needed to pass the bridge;
- time required to reverse allowed direction under bridge (in case traffic is two sided);
- longest allowed time to block road traffic over bridge;
- minimal required period between operating the bridge.

Locks are specified by the following characteristics:

- operational period;
- width of the lock;
- length of the lock;
- time needed to enter or leave the lock;
- number of parallel chambers;
- number of sequential chambers;
- time needed to change the water level in the lock chamber.

4.1.7 **Fleet data**

Here data collection is designed to indicate which ship types have waterway accessibility, and are designed to carry what good types. Data is stratified by good type, ship types, ship classes, and ship depths.

Goods categories are:

- Code identifying the good type;
- Name of the good;
- Value per cubic meter;
- Volume per cubic meter;
- Average shipment size;
- Fraction of goods that are classified as hazardous.

Ship classes are classified by:

- Description of ship class;
- Length;
- Width;
- Height empty;
A freight traffic modeling approach for the waterway network in the Netherlands

- Height fully loaded;
- Maximum loading capacity;
- Maximum load volume;
- Maximum speed;
- Fuel usage;
- Power;
- Yearly depreciation;
- Personnel costs;
- Other costs;
- Fixed costs;
- Rate of waiting;
- Rate of staying;
- Cost of travel full;
- Cost of travel empty;
- Number of available ships in this class.

Ship type are categorized by:
- Ship class to which the ship type belongs;
- Containership (1: yes, 0: no);
- Tanker (1: yes, 0: no);
- Number of kegels (indicating hazardous materials);
- Number of ships available in this class.

The table 4.1 below extracted from the “Ministerie van Verkeer en Waterstaat”, December of 1997 identifies different ship classes.
### Table 4.1 Ship Classes

<table>
<thead>
<tr>
<th>Klasse</th>
<th>Type motorschip</th>
<th>tonnage</th>
<th>Samenstelling duwstel</th>
<th>tonnage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>recreativaart</td>
<td>&lt;250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>splits</td>
<td>250 - 400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>kempenaar</td>
<td>400 - 650</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>Dortmund-eemkanaalship</td>
<td>650 - 1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>Rijn-Hernekanaalschip</td>
<td>1000</td>
<td>1250</td>
<td>1450</td>
</tr>
<tr>
<td>Va</td>
<td>Groot rijnschip</td>
<td>1500 - 1600</td>
<td>3000</td>
<td>3200 - 6000</td>
</tr>
<tr>
<td>Vb</td>
<td></td>
<td>3200 - 6000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vla</td>
<td></td>
<td>3200 - 6000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vlb</td>
<td></td>
<td>6400 - 12000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vlc</td>
<td></td>
<td>9600 - 18000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9600 - 18000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4.2 shows the number of registered ships in the Netherlands in the period between 1994-1996

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Motorvrachtschip</td>
<td>00</td>
<td>3750</td>
<td>3665</td>
<td>3541</td>
</tr>
<tr>
<td>RO-RO schip(Personenwagens)</td>
<td>02</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>RO-RO schip(Opleggers)</td>
<td>03</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Beunsschip</td>
<td>04</td>
<td>484</td>
<td>477</td>
<td>473</td>
</tr>
<tr>
<td>Containerschip</td>
<td>05</td>
<td>14</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Cementtanker</td>
<td>06</td>
<td>54</td>
<td>55</td>
<td>58</td>
</tr>
<tr>
<td>Tankschip</td>
<td>10</td>
<td>764</td>
<td>785</td>
<td>812</td>
</tr>
<tr>
<td>Sleepvrachtschip</td>
<td>20</td>
<td>242</td>
<td>231</td>
<td>215</td>
</tr>
<tr>
<td>Vrachtduwbak</td>
<td>30</td>
<td>11</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Tankduwbak</td>
<td>41</td>
<td>68</td>
<td>67</td>
<td>68</td>
</tr>
<tr>
<td>Elevator</td>
<td>43</td>
<td>222</td>
<td>215</td>
<td>204</td>
</tr>
<tr>
<td>Sleepboot</td>
<td>50</td>
<td>649</td>
<td>596</td>
<td>575</td>
</tr>
<tr>
<td>Duwbout</td>
<td>51</td>
<td>145</td>
<td>152</td>
<td>150</td>
</tr>
<tr>
<td>Sleep-duwbout</td>
<td>52</td>
<td>347</td>
<td>392</td>
<td>411</td>
</tr>
<tr>
<td>Dek-zolderschuit</td>
<td>60</td>
<td>1306</td>
<td>1273</td>
<td>1205</td>
</tr>
<tr>
<td>Motordekschuit</td>
<td>61</td>
<td>130</td>
<td>130</td>
<td>126</td>
</tr>
<tr>
<td>Passagiersschip(Meerdags)</td>
<td>70</td>
<td>89</td>
<td>95</td>
<td>103</td>
</tr>
<tr>
<td>Passagiersschip(Dagtocht)</td>
<td>71</td>
<td>247</td>
<td>247</td>
<td>263</td>
</tr>
<tr>
<td>Drijvend werkvartuig</td>
<td>80</td>
<td>30</td>
<td>32</td>
<td>49</td>
</tr>
<tr>
<td>Overige vaartuigen</td>
<td>90</td>
<td>336</td>
<td>313</td>
<td>211</td>
</tr>
<tr>
<td>Duwbak RO-RO (Personenautos)</td>
<td>91</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Duwbak RO-RO (Opleggers)</td>
<td>92</td>
<td>3</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>Onbekend scheepstype</td>
<td>99</td>
<td>913</td>
<td>991</td>
<td>1105</td>
</tr>
</tbody>
</table>

Transportables are classified by:

- Code identifying good type;
- Number identifying ship type used to carry this good type.

There is another data structure that is used in the modeling. This data represents the tons carried between origins and destinations. Specifically, it represents demand at the destination, and assumes that supply at the origins is equal to demand at the destinations. This data is not exactly clear. For the sake of just introducing what tentatively can be used, it is included.
The tons matrix structure is as follows:

- Origin;
- Destination;
- Year;
- Category of goods;
- Tons of this good category that is demanded per year;
- Generalized travel costs for this good category.

The report written by van der Zijpp et al., 1999 provides a more detailed set of tables. The details are dropped here due to author's uncertainty as the meaning of some of the extra explanations.

4.2 Data Use In The Modeling

As can be seen the database is very extensive. But given the difficulties discussed above, it is reasonable to consider only those variables that are clearly defined and are easy to collect information on. To achieve a reasonable level of precision in modeling the criteria for including variables was chosen as follows:
5. A DYNAMIC ASSIGNMENT MODEL
5. A DYNAMIC ASSIGNMENT MODEL

5.1 Introduction

This chapter will provide a design guidelines for network definition. Two objectives are in mind: accuracy and cost. A waterway network can be represented as a system of nodes and links or in other words a system of directed graphs. Links are characterized by several attributes such as length, speed, capacity, depth, and water level. Waterway network is somewhat different from road networks. A waterway network has a set of bridges and locks as well as canals. Bridges and locks are represented as links aside from the characteristics mentioned above each bridge and lock is characterized by its storage capacity for queues. Canals are links that do not encounter queues. Therefore, delays on canals only occur before reaching bridges and locks. Generalized costs for route choice are a linear combination of time and distance. There may be other factors that affect route choice but for the sake of simplification they are not included here.

5.2 Networks

As was mentioned earlier a system of canals can be represented as a network of links, bridges, and locks. Canals can be represented as arcs with an A-node and a B-node. The arcs can be either one directional or two-directional as the case may be. Bridges can be represented as arcs. These are links with particular characteristics; there is change in the speed before entering into the link, there is therefore, delay before entering the link but no delay while in the link, and the link is always one directional. An analogy with the traffic network would be a bridge is similar to a traffic light at an intersection where there are no right or left turns. Locks can be represented as arcs. The A-node of the arc is the entry point to the lock; and the B-node of the arc is the exit point of the lock.

The waterway network if superimposed on a street network would give a more clear aspect to the general topology. Of course the only topology in common with roadway network are the bridges. In the case of a bridge, it only has virtual representation in the waterway network; since it is shown as a link with special characteristics. In the roadway network the same bridge has a quantitative representation, it is a link on the street network; the special quality of this link is that it has systematic discontinuity with the roadway network. While the bridge is in operation (either lifted up, or sways to the side) there is a discontinuity in the street network, while the waterway network will stay continuous. Actually, the bridges are the only points where the road network and waterway network interact. Otherwise, these two networks exist irrespective of each other. The point that is being made here is that any change in the operation of the bridges affects ships as well as vehicles. If the bridge is high enough and there is no need for the bridge either to be
A freight traffic modeling approach for the waterway network in the Netherlands

lifted or shifted, then the bridge has no effect on the operation of the ships on the canals. It can just be represented as a point on the arc. Otherwise, if the bridge is low, and it is required to move the bridge in order to let ships pass, then if the bridge is not operational, then there is a dead-end point in the waterway network. Both the links before and after the bridge are excluded from paths; unless the arcs just before and after the bridges are destinations where ships drop off cargo, or are origins where ships pick up cargo.

The origins and destinations in the waterway network are the harbors. These are designated places where ships either pick up or drop off goods. The home bases of ships are either the designated harbors, or traditional spots where ships usually anchor. All the origins and destinations, and home bases can be identified from various sources such as city maps, carrier companies or individual carriers, and shipper companies, or the census data.

With each link in the network there is associated a cost. The costs are the function of a number of attributes. In a road network, these attributes: distance, free-flow speed, capacity and a speed-flow relationship. The costs associated with the waterway network links include some of the costs associated with the roadway links, such as distance, and free-flow speed; capacity is defined in terms of variables such as the canal depth, and canal water level. The demand for the waterway network system can be defined the same way as the roadway network. The demand is made up of the number of trips per origin-destination pair and mode that would be given for a level of service. Level of service in this context would be travel time, but monetary costs such as fuel and fares are also relevant. If the actual level of service offered by the transport network turns out to be lower than estimated, then a reduction in the demand and perhaps a shift to other destinations, modes and/or times of day would be expected.

Equilibrium between transport supply (network) and transportation demand (users) is reached when the users from a fixed trip matrix seek routes to minimize their travel times. This results in their trying alternative routes, exploring new ones and perhaps settling into a relatively stable pattern after much trial and error. This allocation of trips to routes yields a pattern of path and link flows which could be said to be in equilibrium when users can no longer find better routes to their destinations: they are already travelling on the best routes possible. This is equilibrium. The equilibrium points are not static and within a time period several equilibrium points occur. The resulting flow pattern may affect the choices of mode, destination and time of day for travel. These shifts in demand will induce in turn changes in the corresponding equilibrium points. In modeling terms, the new flow pattern produces levels of service for routes and modes, which may or may not be consistent with those, assumed in estimating the presumed fixed trip matrix. This requires re-estimating the matrix and therefore, feeding back the new levels of service into the estimation process to obtain a new one. The process may need to be repeated in a systematic way.
5.3 Dynamic Assignment

5.3.1 Route Choice

The basic idea in assignment is the assumption of a rational traveler, i.e.; one choosing the route, which offers the least perceived (and anticipated) individual costs. A number of factors are thought to influence the choice of route when driving between two points; these include journey time, distance, monetary cost (fuel and operating costs, tariffs), congestion and queues at the bridges and locks, type of maneuvers required, reliability of travel time and habit. The production of generalized cost expression incorporating all these elements is a difficult task. Furthermore, it is not practical to try to model all of them in a traffic assignment model, and therefore approximations are inevitable.

The most common approximation is to consider only two factors in route choice: time and monetary cost; further, monetary cost is often deemed proportional to distance. The majority of assignment programs allow the user to allocate weights to travel time and distance in order to represent users perceptions of these two factors. The weighted sum of these two values then becomes a generalized cost used to estimate route choice. There is evidence to suggest that, at least for urban car traffic, time is the dominant factor in route choice. In waterway freight traffic a combination of travel time and delay at bridges and locks would make for a reasonable generalized cost function.

The idea these days is that due to many factors such as congestion, differences in perception, and imperfect information on route costs or simply errors, it is more appropriate to consider multiple paths. These are a set of reasonable paths that might be used by the carriers. The number of such reasonable paths is limited due to the topology of the network. The range should be between three and five alternate routes.

5.3.2 Graph – Example And Notation

To illustrate the concept of multiple path choice a graph is used. This graph has eight consecutive nodes (graph 5.1). The graph is denoted by $G = (X,A,W)$. $X$ represents the set of nodes, $A$ represents a set of links that connect the nodes, and $W$ represents link costs. There are $N$ nodes and $M$ links (arcs). The point of departure is denoted by $s$. The table that represents the paths between the nodes and their costs is denoted by $V$. $V[x]$ is the cheapest path cost between $s$.
and x. Usually $V[x]$ is initialized to infinity, with the exception of $V[s]$ which is equal to zero, $V[x] = \infty$, $V[s] = 0$. The first thing that is calculated is the probability of traversing between any node $s$, and $x$ given the costs. Those paths are chosen that have the highest probability given the cost $p(s,x)$. Two tables can be constructed; one that shows the probabilities; and the second the table $V$ that shows the costs (travel times, or distances).

Graph 5.1 A network with link costs indicated over each link

There are many algorithms that can be used in finding the shortest path. Among these algorithms there are the Dijkstra's algorithm (positive link costs). One of the constraints of this algorithm is that the cost of each link should be positive. This algorithm is very well suited to the problem of finding the shortest path between two nodes $s$ and $t$. A more general version of the problem above is given the departure node $s$, find shortest paths between $s$ and the all other nodes. To solve the more general problem of finding shortest paths, the procedure Dijkstra is suggested. Short version of Dijkstra's algorithms is given here:

(initialization)
Initialize table $V$ to $\infty$
Initialize table $P$ to 0
Initialize table Done to False
$V[s] := 0$
$P[s] := s$
Repeat
{search for the node in table $V$ that is not fixed}
$V_{\text{min}} := +\infty$
For
A freight traffic modeling approach for the waterway network in the Netherlands

\[ y := 1 \text{ to } N \]
If (not Done[y]) and (V[y] < Vmin) then
\[ x := y \]
\[ Vmin := V[y] \]
EndIf
EndFor
If Vmin < + \infty then (if x exist)
Done[x] := True
For k := 1 to \#Succ[x]
\[ y := \text{Succ}[k] \]
If V[x] + W[x,k] < V[y] then
\[ V[y] = V[x] + W[x,k] \]
\[ P[y] := x \]
EndIf
EndFor
EndIf
Repeat until Vmin = + \infty (Done[x] = True for all x)

The results of implementing shortest path algorithms on the graph above is the table below:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>-1</td>
<td>5</td>
<td>7</td>
<td>\infty</td>
<td>\infty</td>
<td>\infty</td>
<td>\infty</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>5</td>
<td>7</td>
<td>-4</td>
<td>\infty</td>
<td>\infty</td>
<td>\infty</td>
<td>5</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>-4</td>
<td>7</td>
<td>4</td>
<td>\infty</td>
<td>\infty</td>
<td>\infty</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>-5</td>
<td>13</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>11</td>
<td>11 end</td>
</tr>
</tbody>
</table>

The result of the shortest path algorithm is as follows:
A freight traffic modeling approach for the waterway network in the Netherlands

Path between node 1 and node 1: cost = 0. Connecting nodes = 1
Path between node 1 and node 2: cost = 1. Connecting nodes = 1, 2
Path between node 1 and node 3: cost = 4. Connecting nodes = 1, 2, 5, 3
Path between node 1 and node 4: cost = 7. Connecting nodes = 1, 4
Path between node 1 and node 5: cost = 4. Connecting nodes = 1, 2, 5
Path between node 1 and node 6: cost = 5. Connecting nodes = 1, 2, 5, 3, 6
Path between node 1 and node 7: cost = 9. Connecting nodes = 1, 2, 5, 3, 6, 7
Path between node 1 and node 8: cost = 11. Connecting nodes = 1, 2, 5, 3, 6, 7, 8

5.4 Dynamic Route Choice

The idea here is that the number of paths that users are willing to choose from are limited and usually fixed. What does vary is the probability of choosing any particular path given the departure time from the origin, and anticipated delays in the route. Therefore, if there are 3 reasonable paths with path one being the shortest distance path it may be that this path is considered to be the shortest path depending on when the user has left the origin. The idea in dynamic route choice modeling is to divide a day into a number of fixed time intervals and find the fraction of traffic that is on the path during each time interval. This can be done if the probabilities of choosing any path are known during each interval.

A survey of data collected from Dutch ministry of water management shows that there are pre-designated locations on the canals where there are counting devices installed in order to count the number of ships that cross the area. A sample of data collected by “Visuele Tellingen” the visual counters during 1995, “Tabel bij figuur 3.3 Visuele tellingen Periode 1995”. The time interval during which the data is collected are coarse. The locations are few and data collection days are not many. The idea here is that they capability to collect data be on real time basis exits and it can be expanded. Figuur 3.1 show that in 1996 the ministry of waterworks installed many more counter (telpunten) in more locations. These counters are exclusively located at locks or “sluifs”.

Another map “Bijlage Telpuntaart met vaarwegnummers” shows a more extensive data collection counters both visual and radar types. The problem is that the data from these counters is not collected regularly, and even if it is collected the data is not used in the analysis of waterway freight transport problem. To use the counter efficiently, the time intervals should be redefined in a way that there will be more time intervals during a day for example 30 minutes intervals seem to be reasonable. The choice of including all days of the week is good. The choice of period of year to collect data is good. The argument is that if the counters and radars are installed permanently why not collect data regularly every day 365 days a year by time intervals of 30 minutes. Data observed can be collected regularly and updated. This data will be
valuable in providing an insight into the structure of the freight traffic in the waterway system of the Netherlands.

A path is a directional route between an origin and destinations. The origins and destinations are determined through zoning. A zoning system is used to aggregate harbors and wharfs where cargo can be dropped off into manageable chunks for modeling purposes. The main two dimensions of a zoning system are the number of zones and their size. The waterway network zoning is shown in “Bijlage Regionale indeling verkeersgebieden in Netherland Situatie 1997”. Zones are represented in the computer models as if all their attributes and properties were concentrated in a single point called zone centroid. Centroids are arbitrary and do not represent any particular location on a map. Centroids are attached to the network through centroid connectors representing the average costs (time, distance) of joining the transport system for trips with origin or destination in that zone. Nearly as important as the cost associated to each centroid connector is the node in the network it connects to. These should be natural access/egress points for the zone itself. Based on the map it can be seen that there are 53 origin-destinations in the Netherlands. Each origin can also be a destination. A set of origin-destination pairs can be determined given the zoning system. A (53 x 53) table of origin-destinations can be constructed.

The next item of importance is departure times from origin. The following notation will be used:

- $t(d)$: departure time from origin (am/pm).
- $t(c,d)$: time (allocated by shipper) to deliver cargo (am/pm).
- $tt$: duration of trip (total travel time) (hrs.).
- $s(t)$: speed during interval $t$ (kmph).
- $d$: total distance from origin to destination (km).
- $w(t)$: waiting time to load cargo at harbor (hrs).
- $\delta(t)$: envisioned delay (hrs).
- $tt^*$: anticipated time available for the trip (hrs).
- $t^*(d)$: tentative departure time from origin (am/pm).
- $s(t)$: slack time (time needed to make the trip)

Then

$$tt = \frac{s(t)}{d}$$

$$tt^* = t(c,d) - t^*(d)$$

$$s(t) = tt^* - \{ tt + w(t) + \delta(t) \}$$

The following procedure is suggested in order to estimate departure time from origin.
If 
\[ s(t) < 0 \] then 
Add 
\[ t^*(d) = (t^*(d) + s(t)) \] 
Put 
\[ t(d) = t^*(d) \] 
end 
est if 
\[ s(t) > 0 \] then 
if \[ s(t) = \text{optimal} \] then 
Put 
\[ t(d) = t^*(d) \] 
elseif 
\[ s(t) \neq \text{optimal} \] then 
revise \( ttt & w(t) & \delta(t) \) 
set 
\[ t(d) = \text{new } t^*(d) \] 
end

The probability of choosing a path \( k \) given departure time from origin \( t(d) \), \( p(k / t(d)) \) can be estimated by a multinomial logit model. The multinomial logit form of the choice probabilities is expressed as:

\[
p(k / t(d), r) = \frac{\exp(-\mu(ttt(r,k,t(d))))}{\sum_{m} \exp(-\mu(ttt(r,m,t(d))))}
\]

where \( \mu \) is a scale parameter, and \( ttt(r,k,t(d)) \) is the total travel time on any path \( k \) for an O-D pair \( r \), given that departure time from the origin is during interval \( t(d) \). The parameter \( m \) represents any other reasonable path. The values for the parameter \( \mu \) is calculated based on real time link traffic data information and are not arbitrary.

In finding multiple paths, it is assumed that in any waterway network, due to topological limitations such as the number of available canals, canal depths, and water levels the number of reasonable routes one can take to go from an origin to a destination are limited, and are always
the same. What is dynamic or time-dependent in this case is the probability of choosing any of these paths at any particular point in time.

Therefore, a shortest path between an origin and a destination is found using either Dijkstra's or an equilibrium assignment algorithm. To find an alternate path, every single shortest path is traced and a link from each path is deleted. The choice of which link to delete strongly depends on familiarity with the network topology. One plausible way is to find the second shortest path is on each shortest path already found, delete the link with the highest cost. Run the program again given that now some of the links are eliminated from the pool of links. For example in graph 5.1, from node 1 to nodes 2 and 4, there is only one link that connects them together; therefore, there is one path between them. But among the paths that connect node 1 to nodes 3, 5, 6, 7, and 8 it is possible to find another shortest path. In the path between node 1 and node 5, if link that connects node 2 to node 3 is deleted; then the next reasonable shortest path is from node 1 to 4 and from node 4 to node 5. The total cost of this path is 6.

The third shortest path is found by deleting the link that connects node 1 to node 4. Then the shortest path between node 1 and node 5 is through node 6 to node 5; the cost is 9. Even though it is possible to find other reasonable paths; it is not necessary. Most users do not consider more than three reasonable paths; especially if they have to traverse the path regularly. This is due to individual perception, and habit, and particularly lack of sufficient information. The equilibrium assignment is run again to find the second shortest path. This process is repeated to find other alternate paths. Once all reasonable paths are found, links on the paths are loaded with travel times that correspond to a particular interval and the probabilities of choosing each path for that interval are calculated.

5.5 PRACTICAL CONSIDERATIONS

Correct implementation of the assignment model is critical to waterway freight transportation modeling. There are many sources of errors. One such source is network modeling. There are many potential errors in network coding such as miscoding distances, depths and water levels. There are many software packages that can flag errors out, use of graphical display of the network and graphic editing the links usually helps correcting errors. The other important source of error is costs. These relate to speed, delay and other relevant variables.
Tabel bij figuur 3.3 Visuele tellingen

Periode 1995

<table>
<thead>
<tr>
<th>Periode</th>
<th>Tijd</th>
<th>Maandag</th>
<th>Dinsdag</th>
<th>Woensdag</th>
<th>Donderdag</th>
<th>Vrijdag</th>
<th>Zaterdag</th>
<th>Zondag</th>
</tr>
</thead>
<tbody>
<tr>
<td>s'-Gravendeel</td>
<td>nacht</td>
<td>00.00 - 08.00</td>
<td>24 jun</td>
<td>23 apr</td>
<td>24 apr</td>
<td>21 nov</td>
<td>22 nov</td>
<td>22 jun</td>
</tr>
<tr>
<td></td>
<td>dag</td>
<td>08.00 - 16.00</td>
<td>22 apr</td>
<td>23 apr</td>
<td>24 apr</td>
<td>21 nov</td>
<td>22 nov</td>
<td>22 jun</td>
</tr>
<tr>
<td></td>
<td>avond</td>
<td>16.00 - 24.00</td>
<td>22 apr</td>
<td>23 apr</td>
<td>20 nov</td>
<td>21 nov</td>
<td>22 nov</td>
<td>22 jun</td>
</tr>
<tr>
<td></td>
<td>nacht</td>
<td>00.00 - 08.00</td>
<td>23 dec</td>
<td>23 jan</td>
<td>24 jan</td>
<td>22 feb</td>
<td>23 feb</td>
<td>23 mrt</td>
</tr>
<tr>
<td></td>
<td>dag</td>
<td>08.00 - 16.00</td>
<td>22 jan</td>
<td>23 jan</td>
<td>24 jan</td>
<td>22 feb</td>
<td>23 feb</td>
<td>23 mrt</td>
</tr>
<tr>
<td></td>
<td>avond</td>
<td>16.00 - 24.00</td>
<td>22 jan</td>
<td>23 jan</td>
<td>21 feb</td>
<td>22 feb</td>
<td>23 feb</td>
<td>23 mrt</td>
</tr>
<tr>
<td></td>
<td>nacht</td>
<td>00.00 - 08.00</td>
<td>25 mrt</td>
<td>23 apr</td>
<td>24 apr</td>
<td>23 mei</td>
<td>24 mei</td>
<td>22 jun</td>
</tr>
<tr>
<td></td>
<td>dag</td>
<td>08.00 - 16.00</td>
<td>22 apr</td>
<td>23 apr</td>
<td>24 apr</td>
<td>23 mei</td>
<td>24 mei</td>
<td>22 jun</td>
</tr>
<tr>
<td></td>
<td>avond</td>
<td>16.00 - 24.00</td>
<td>22 apr</td>
<td>23 apr</td>
<td>22 mei</td>
<td>24 mei</td>
<td>22 jun</td>
<td>23 jun</td>
</tr>
<tr>
<td></td>
<td>nacht</td>
<td>00.00 - 08.00</td>
<td>24 jun</td>
<td>23 jul</td>
<td>24 jul</td>
<td>22 aug</td>
<td>23 aug</td>
<td>23 mrt</td>
</tr>
<tr>
<td></td>
<td>dag</td>
<td>08.00 - 16.00</td>
<td>22 jul</td>
<td>23 jul</td>
<td>24 jul</td>
<td>22 aug</td>
<td>23 aug</td>
<td>23 mrt</td>
</tr>
<tr>
<td></td>
<td>avond</td>
<td>16.00 - 24.00</td>
<td>22 jul</td>
<td>23 jul</td>
<td>21 aug</td>
<td>22 aug</td>
<td>23 aug</td>
<td>23 mrt</td>
</tr>
<tr>
<td></td>
<td>nacht</td>
<td>00.00 - 08.00</td>
<td>23 sep</td>
<td>22 okt</td>
<td>23 okt</td>
<td>21 nov</td>
<td>22 nov</td>
<td>21 dec</td>
</tr>
<tr>
<td></td>
<td>dag</td>
<td>08.00 - 16.00</td>
<td>21 okt</td>
<td>22 okt</td>
<td>23 okt</td>
<td>21 nov</td>
<td>22 nov</td>
<td>21 dec</td>
</tr>
<tr>
<td></td>
<td>avond</td>
<td>16.00 - 24.00</td>
<td>21 okt</td>
<td>22 okt</td>
<td>20 nov</td>
<td>21 nov</td>
<td>22 nov</td>
<td>21 dec</td>
</tr>
<tr>
<td></td>
<td>nacht</td>
<td>00.00 - 08.00</td>
<td>12 feb</td>
<td>12 mrt</td>
<td>13 mrt</td>
<td>14 mrt</td>
<td>12 apr</td>
<td>13 apr</td>
</tr>
<tr>
<td></td>
<td>dag</td>
<td>08.00 - 16.00</td>
<td>12 feb</td>
<td>12 mrt</td>
<td>13 mrt</td>
<td>11 apr</td>
<td>12 apr</td>
<td>13 apr</td>
</tr>
<tr>
<td></td>
<td>avond</td>
<td>16.00 - 24.00</td>
<td>12 feb</td>
<td>12 mrt</td>
<td>12 mrt</td>
<td>11 apr</td>
<td>12 apr</td>
<td>11 mei</td>
</tr>
<tr>
<td></td>
<td>nacht</td>
<td>00.00 - 08.00</td>
<td>13 mei</td>
<td>11 jun</td>
<td>12 jun</td>
<td>13 jun</td>
<td>12 jul</td>
<td>13 jul</td>
</tr>
<tr>
<td></td>
<td>dag</td>
<td>08.00 - 16.00</td>
<td>13 mei</td>
<td>11 jun</td>
<td>12 jun</td>
<td>11 jul</td>
<td>12 jul</td>
<td>13 jul</td>
</tr>
<tr>
<td></td>
<td>avond</td>
<td>16.00 - 24.00</td>
<td>13 mei</td>
<td>11 jun</td>
<td>12 jun</td>
<td>11 jul</td>
<td>12 jul</td>
<td>10 aug</td>
</tr>
<tr>
<td></td>
<td>nacht</td>
<td>00.00 - 08.00</td>
<td>12 aug</td>
<td>10 sep</td>
<td>11 sep</td>
<td>12 sep</td>
<td>11 okt</td>
<td>12 okt</td>
</tr>
<tr>
<td></td>
<td>dag</td>
<td>08.00 - 16.00</td>
<td>12 aug</td>
<td>10 sep</td>
<td>11 sep</td>
<td>10 okt</td>
<td>11 okt</td>
<td>12 okt</td>
</tr>
<tr>
<td></td>
<td>avond</td>
<td>16.00 - 24.00</td>
<td>12 aug</td>
<td>10 sep</td>
<td>11 sep</td>
<td>10 okt</td>
<td>11 okt</td>
<td>9 nov</td>
</tr>
<tr>
<td></td>
<td>nacht</td>
<td>00.00 - 08.00</td>
<td>11 nov</td>
<td>10 dec</td>
<td>11 dec</td>
<td>12 dec</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>avond</td>
<td>16.00 - 24.00</td>
<td>11 nov</td>
<td>10 dec</td>
<td>11 dec</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A modeling approach to freight traffic on the canals in the Netherlands

Telpunten

Figuur 3.1 IVS-telpunten op het hoofdvaarwegennet
Situatie per 1 januari 1996

<table>
<thead>
<tr>
<th>Nr</th>
<th>IVS-’90-telpunt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Zeesluizen te Delfzijl</td>
</tr>
<tr>
<td>2</td>
<td>Oosterhuis te Groningen</td>
</tr>
<tr>
<td>3</td>
<td>Sluis te Gaarkeuken</td>
</tr>
<tr>
<td>4</td>
<td>Prinses Margrietsoor</td>
</tr>
<tr>
<td>5</td>
<td>Schutsluis te Eefde</td>
</tr>
<tr>
<td>6</td>
<td>Schutsluis te St. Andries</td>
</tr>
<tr>
<td>7</td>
<td>Schutsluis te Oudezanden</td>
</tr>
<tr>
<td>8</td>
<td>Schutsluis te Arumbeek</td>
</tr>
<tr>
<td>9</td>
<td>Schutsluis te Huigenbergen</td>
</tr>
<tr>
<td>10</td>
<td>Schutsluis te Würtemberg</td>
</tr>
<tr>
<td>11</td>
<td>Henneketelssluizen te Engelen</td>
</tr>
<tr>
<td>12</td>
<td>Sluiskomplex te Panheele</td>
</tr>
<tr>
<td>13</td>
<td>Kreeksluizen</td>
</tr>
<tr>
<td>14</td>
<td>Westerwoldsluizen</td>
</tr>
<tr>
<td>15</td>
<td>Sluiskomplex Huisweert</td>
</tr>
<tr>
<td>16</td>
<td>Volkeraksluizen</td>
</tr>
<tr>
<td>17</td>
<td>Kiemersluizen</td>
</tr>
<tr>
<td>18</td>
<td>Sluis Beeld</td>
</tr>
<tr>
<td>19</td>
<td>Sluis Sambrekk</td>
</tr>
<tr>
<td>20</td>
<td>Sluis Grave</td>
</tr>
<tr>
<td>21</td>
<td>Sluis Lith</td>
</tr>
<tr>
<td>22</td>
<td>Schutsluis te Born</td>
</tr>
<tr>
<td>23</td>
<td>Deelelingsluis te Maasbracht</td>
</tr>
<tr>
<td>24</td>
<td>Sluis Heel</td>
</tr>
<tr>
<td>25</td>
<td>Sluis te Linne</td>
</tr>
<tr>
<td>26</td>
<td>Juliana sluis te Gouda</td>
</tr>
<tr>
<td>27</td>
<td>Prins Irene sluis</td>
</tr>
<tr>
<td>28</td>
<td>Prins Bernhardsluis</td>
</tr>
<tr>
<td>29</td>
<td>Prinses Beatrixsluis</td>
</tr>
<tr>
<td>30</td>
<td>Houtbroodsluizen</td>
</tr>
<tr>
<td>31</td>
<td>Oranjesluizen te Schellingwoude</td>
</tr>
<tr>
<td>32</td>
<td>Sluizen te Utrecht</td>
</tr>
<tr>
<td>33</td>
<td>Krabbersgat sluisen</td>
</tr>
<tr>
<td>34</td>
<td>Lorentzsluizen</td>
</tr>
<tr>
<td>35</td>
<td>Steinsluizen</td>
</tr>
</tbody>
</table>

Bron: AVV, 1996
Telpunten

Figuur 3.1 IVS-telpunten op het hoofdvaarwegennet
Situation per 1 januari 1996

<table>
<thead>
<tr>
<th>Nr</th>
<th>IVS-'90-telpunt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Zeesluis te Delfzijl</td>
</tr>
<tr>
<td>2</td>
<td>Oostersluis te Groningen</td>
</tr>
<tr>
<td>3</td>
<td>Sluis te Caarkeuken</td>
</tr>
<tr>
<td>4</td>
<td>Prins Margrethsluis</td>
</tr>
<tr>
<td>5</td>
<td>Schutsluis te Eefde</td>
</tr>
<tr>
<td>6</td>
<td>Schutsluis te St. Andries</td>
</tr>
<tr>
<td>7</td>
<td>Schutsluis te Linel</td>
</tr>
<tr>
<td>8</td>
<td>Schutsluis te Amerongen</td>
</tr>
<tr>
<td>9</td>
<td>Schutsluis te Hagestein</td>
</tr>
<tr>
<td>10</td>
<td>Schutsluis te Weurt</td>
</tr>
<tr>
<td>11</td>
<td>Henriette sluis te Engelen</td>
</tr>
<tr>
<td>12</td>
<td>Sluiscomplex te Panheel</td>
</tr>
<tr>
<td>13</td>
<td>Kreekraksluisen</td>
</tr>
<tr>
<td>14</td>
<td>Westsluis Terneuzen</td>
</tr>
<tr>
<td>15</td>
<td>Sluiscomplex Hansweert</td>
</tr>
<tr>
<td>16</td>
<td>Volkeraksluisen</td>
</tr>
<tr>
<td>17</td>
<td>Krammersluisen</td>
</tr>
<tr>
<td>18</td>
<td>Sluis Belfeld</td>
</tr>
<tr>
<td>19</td>
<td>Sluis Sambek</td>
</tr>
<tr>
<td>20</td>
<td>Sluis Grave</td>
</tr>
<tr>
<td>21</td>
<td>Sluis Lith</td>
</tr>
<tr>
<td>22</td>
<td>Schutsluis te Born</td>
</tr>
<tr>
<td>23</td>
<td>Dinkelingsluis te Maastricht</td>
</tr>
<tr>
<td>24</td>
<td>Sluis Heel</td>
</tr>
<tr>
<td>25</td>
<td>Sluis te Linne</td>
</tr>
<tr>
<td>26</td>
<td>Juliana sluis te Gouda</td>
</tr>
<tr>
<td>27</td>
<td>Prinsen Imerensluis</td>
</tr>
<tr>
<td>28</td>
<td>Prins Bernhardsluis</td>
</tr>
<tr>
<td>29</td>
<td>Prinses Beatrixluis</td>
</tr>
<tr>
<td>30</td>
<td>Houtenluisen</td>
</tr>
<tr>
<td>31</td>
<td>Orangesluisen te Schellingwoude</td>
</tr>
<tr>
<td>32</td>
<td>Sluizen te Ummaden</td>
</tr>
<tr>
<td>33</td>
<td>Krabbersgatsluisen</td>
</tr>
<tr>
<td>34</td>
<td>Lorentszluisen</td>
</tr>
<tr>
<td>35</td>
<td>Stevinsluis</td>
</tr>
</tbody>
</table>

Bron: AVV, 1996
Bijlage  Regionale indeling verkeersgebieden in Nederland

Situatie in 1997

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Verkeersgebied</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Oost-Groningen</td>
</tr>
<tr>
<td>2</td>
<td>Overig Delfzijl e.o.</td>
</tr>
<tr>
<td>3</td>
<td>Delfzijl/Eemhaven</td>
</tr>
<tr>
<td>4</td>
<td>West-Groningen</td>
</tr>
<tr>
<td>5</td>
<td>Groningen (stad)</td>
</tr>
<tr>
<td>6</td>
<td>Overig noord-Friesland</td>
</tr>
<tr>
<td>7</td>
<td>Harlingen</td>
</tr>
<tr>
<td>8</td>
<td>Zuidoost-Friesland</td>
</tr>
<tr>
<td>9</td>
<td>Zuidoost-Friesland</td>
</tr>
<tr>
<td>10</td>
<td>Noord-Drenthe</td>
</tr>
<tr>
<td>11</td>
<td>Zuidoost-Drenthe</td>
</tr>
<tr>
<td>12</td>
<td>Zuidoost-Drenthe</td>
</tr>
<tr>
<td>13</td>
<td>Noord-Overijssel</td>
</tr>
<tr>
<td>14</td>
<td>Zuidoost-Overijssel</td>
</tr>
<tr>
<td>15</td>
<td>Twente</td>
</tr>
<tr>
<td>16</td>
<td>Veluwe</td>
</tr>
<tr>
<td>17</td>
<td>Achterhoek</td>
</tr>
<tr>
<td>18</td>
<td>Arnhem/Nijmegen</td>
</tr>
<tr>
<td>19</td>
<td>Zuidwest-Gelderland</td>
</tr>
<tr>
<td>20</td>
<td>Utrecht</td>
</tr>
<tr>
<td>21</td>
<td>Kop Noord-Holland</td>
</tr>
<tr>
<td>22</td>
<td>Alkmaar e.o.</td>
</tr>
<tr>
<td>23</td>
<td>Overig Noordzeekean, geb.</td>
</tr>
<tr>
<td>24</td>
<td>Ijmuiden/Velsen</td>
</tr>
<tr>
<td>25</td>
<td>Agglomeratie Haarlem</td>
</tr>
<tr>
<td>26</td>
<td>Zaanstreek</td>
</tr>
<tr>
<td>27</td>
<td>Overig Groot-Amsterdam</td>
</tr>
<tr>
<td>28</td>
<td>Amsterdam</td>
</tr>
<tr>
<td>29</td>
<td>Gooi/Vechtstreek</td>
</tr>
<tr>
<td>30</td>
<td>Leiden/Bollenstreek</td>
</tr>
<tr>
<td>31</td>
<td>Agglomer. 's-Gravenhage</td>
</tr>
<tr>
<td>32</td>
<td>Delft/Zuid-Holland</td>
</tr>
<tr>
<td>33</td>
<td>Oost-Zuid-Holland</td>
</tr>
<tr>
<td>34</td>
<td>Overig Groot-Rijnmond</td>
</tr>
<tr>
<td>35</td>
<td>Overig Rijnmond</td>
</tr>
<tr>
<td>36</td>
<td>Rotterdam</td>
</tr>
<tr>
<td>37</td>
<td>Vlaardingen</td>
</tr>
<tr>
<td>38</td>
<td>Overig zuidoost-Z.-Holland</td>
</tr>
<tr>
<td>39</td>
<td>Zoetermeer</td>
</tr>
<tr>
<td>40</td>
<td>Zwijndrecht</td>
</tr>
<tr>
<td>41</td>
<td>Overig ZW.-Vlaanderen</td>
</tr>
<tr>
<td>42</td>
<td>Terneuzen/Axel</td>
</tr>
<tr>
<td>43</td>
<td>Overig Zeeland</td>
</tr>
<tr>
<td>44</td>
<td>Vlissingen</td>
</tr>
<tr>
<td>45</td>
<td>Overig west-N.-Brabant</td>
</tr>
<tr>
<td>46</td>
<td>Middenlĳnders</td>
</tr>
<tr>
<td>47</td>
<td>Moerdijkgebied</td>
</tr>
<tr>
<td>48</td>
<td>Midden-N.-Brabant</td>
</tr>
<tr>
<td>49</td>
<td>Noordoost-N.-Brabant</td>
</tr>
<tr>
<td>50</td>
<td>Zuidoost-N.-Brabant</td>
</tr>
<tr>
<td>51</td>
<td>Noord-Limburg</td>
</tr>
<tr>
<td>52</td>
<td>Midden-Limburg</td>
</tr>
<tr>
<td>53</td>
<td>Zuid-Limburg</td>
</tr>
<tr>
<td>54</td>
<td>Flevoland</td>
</tr>
</tbody>
</table>
Bijlage Telpuntenkaart met vaarwegnummers

- IVS-telpunten
- Radar-scheepstreders
- Vaarwegnummer
- Hoofdvaarwegen
- Overige vaarwegen
6. A DYNAMIC TRIP TABLE ESTIMATION
6. A DYNAMIC TRIP TABLE ESTIMATION

6.1 Introduction

The purpose of this chapter is to introduce a new model for estimating a dynamic or time-dependent trip table and to test the model using traffic count data obtained from detectors on a sample of canals in a waterway network. So far the only reliable way to obtain an origin-destination trip table has been through surveys at origins and destinations. These surveys are expensive and they do not reflect changes in traffic flow that occur over time.

There have been a considerable number of studies conducted to determine if an origin-destination matrix could be derived from detectors, because these data can be collected and processed automatically. To develop more effective strategies for better management of canals it is necessary to improve the existing models for estimating O-D flows. In this chapter a reliable substitute to the traditional method of obtaining trip table is introduced.

The method proposed in this paper adds to existing literature first by introducing an alternative approach to estimating time dependent O-D volume which varies from the predecessor models in its functional relativity and sequential implementation. Second, the simultaneous inclusion of the possibility of multiple routes and allowing for trips with travel times exceeding one interval, makes this research different from earlier literature.

6.2 Recursive Generalized Least Square Estimation

The objective is to develop a method for estimating time-dependent origin-destination matrices. By time-dependent it is meant that if a time period is divided into smaller intervals (e.g. 8 hours each), the origin-destination matrices vary for each interval. The model development relies on formulating link flow as a fraction of origin-destination matrices during each time interval plus an error term, which reflects the deviation between the estimated link flow and the real link flow. An origin-destination matrix for an interval is expressed as a linear combination of the origin-destination matrix of the previous interval plus an error term.

The model is a linear system consisting of two equations, the link flow equation or measurement equation, and the origin-destination matrix equation or system equation. The objective is to find an estimated origin-destination matrix such that a quadratic measure is minimized. The estimate that minimizes the quadratic expression is the Least Square estimate. The procedure used to minimize the quadratic measure, or the objective function, by adjusting the origin-destination matrix based on the link flow estimate error variances and the origin-destination estimate error variances is the Generalized Least Square (GLS) Estimation method.
In this section some basic concepts of the recursive GLS algorithm are given. Start with the assumption that volume on any link $l$ during interval $j$, $z(l, j)$, is a fraction $h(r, t, l, j)$, of O-D volumes $x(r, t)$, leaving the origin during either the same or the previous interval and are contributing to the flow on link $l$ during interval $j$. In this formulation, $r$ represents origin-destination pair, and $t$ represents the departure time from the origin. Therefore, traffic flow on any link $l$ during interval $j$ can be expressed by the following equation:

$$z(l, j) = \sum_{r=1}^{n_{od}} \sum_{t=1}^{j} h(r, t, l, j) \cdot x(r, t) + v(l, j)$$  \hspace{1cm} (1)$$

where $n_{od}$ is the number of O-D pairs, $t$ represents intervals from 1 to interval $j$. $v(l, j)$ is assumed to be a zero-mean observation error term.

with covariance

$$\text{cov}(v(l, j)) = \sigma_{V} \delta_{jk}$$  \hspace{1cm} (2)$$

where

$V_{v}(j) =$ Variance of observation error terms during interval $j$.

$d_{jk} =$ Is equal to 1 if interval $j=k$, and equal to 0 if interval $j \neq k$.

In matrix form, equation (1) can be written as:

$$Z(j) = \sum_{t=1}^{j} H(t, j) X(t) + V(j)$$  \hspace{1cm} (3)$$

$Z(j)$ is an $(n_{l} \times 1)$ vector of link volumes during interval $j$, $H(t, j)$ is an $(n_{l} \times n_{od})$ assignment matrix during interval $j$ originating during interval $t$, $X(t)$ is an $(n_{od} \times 1)$ vector of O-D volumes originating during interval $t$, and $V(j)$ is an $(n_{l} \times 1)$ vector of error terms during interval $j$. Due to the short duration of total travel times from an origin to a destination in the proposed study area, only two analysis periods are considered in this study, the present interval ($j$), and the previous interval ($j-1$). Equation (1) is then modified by including the previous interval O-D volumes that are already estimated for interval ($j-1$).

$$z(l, j) = (\sum_{r=1}^{n_{od}} h(r, j-1, l, j) \cdot \hat{x}(r, j-1) + \sum_{r=1}^{n_{od}} h(r, j, l, j) \cdot x(r, j)) + v(l, j)$$  \hspace{1cm} (4)$$
where $\hat{x}(r, j-1)$ is the estimate of the previous interval O-D volume, $x(r, j)$ is the O-D volume to be estimated. Equation (4) can be rewritten as:

$$\tilde{z}(l, j) = \sum_{r=1}^{n_{od}} h(r, j-1, l, j) \cdot \hat{x}(r, j-1)$$ (6)

the matrix form of equation (5) is written as follow:

$$Z(j) = H(j, j) + X(j) + V(j)$$ (7)

$Z$ is an $(n_l \times 1)$ vector of link volumes, $H$ is an $(n_l \times n_{od})$ assignment matrix, $X$ is an $(n_{od} \times 1)$ O-D vector, and $V$ is an $(n_l \times 1)$ error vector. For $j=2$, $\tilde{Z}(j)$ can be written as:

for $j=1$, $\tilde{Z}(j)$ is equal to $Z(j)$.

The O-D volume for interval $j$ can be interpolated from the O-D estimates of interval $(j-1)$ as follows:

$$\hat{x}(j-1) = X(j) + W(j)$$ (9)

where $\hat{x}(j-1)$ is a vector of estimated O-D volumes for the previous interval; $X(j)$ is the present interval O-D vector to be estimated. The causal relationship defined in equation (9) is significant. It allows the analyst to compensate for the problem of rank deficiency, by converting the under-determined system into a full rank system. None of the columns of the $\hat{x}(j-1)$ matrix are linear combinations of the other columns. Therefore, in the optimization procedure there will not be any problems due to non-invertability of matrices that are either derived from or are functions of the O-D flow matrix. $W$ is a $(n_{od} \times 1)$ error or noise vector. The noise $w(j)$ is assumed to be a zero-mean estimation error term.

$$E\{W(j)\} = 0$$
$$\text{cov}\{W(j), W(k)\} = V_w(j, k)$$ (10)
\[ W(j) = \text{Variance of estimation error term during interval } j. \]
\[ d_{jk} = \text{Is equal to 1 if interval } j = k, \text{ and equal to 0 if interval } j \neq k. \]

we can adjoin the O-D estimates for interval \( \hat{X}(j-1) \) with \( \hat{Z}(j) \) to obtain

\[ \hat{Z}(j) = \hat{H}(j)\hat{X}(j) + \hat{V}(j) \quad (11) \]

where

\[
\hat{X}(j) = [\hat{Z}(j) \mid \hat{X}(j-1)]' \\
\hat{H}(j) = [H(j,j) \mid I]' \\
\hat{V} = [V(j) \mid 0] [I \mid W(j)]' \\
\]

The augmented matrices have the following dimensions: \( \hat{Z} \) is a \((n_l + n_{od}) \times 1\) augmented link volume vector, \( \hat{X} \) is a \((n_l + n_{od}) \times 1\) estimate of O-D matrix of the previous interval augmented with the link count estimates of present interval \( j \), \( \hat{H} \) is an \((n_l + n_{od}) \times (n_l + n_{od})\) assignment matrix. \( V(j) \) is a \((n_l \times n_l)\) link error matrix, \( I \) is an \((n_l \times n_{od})\) identity matrix with diagonal entries one, \( 0 \) is an \((n_{od} \times n_l)\) matrix of zeros, \( W(j) \) is an \((n_{od} \times n_{od})\) O-D error matrix. The augmented matrix, \( \hat{V} \) is an \((n_l+n_{od}) \times (n_l+n_{od})\) matrix.

The objective is to select an O-D estimate \( \hat{X}(j) \) such that the quadratic measure

\[ J(\hat{X}(j)) = \frac{1}{2}(\hat{Z}(j) - \hat{H}(j)\hat{X}(j))'\hat{V}^{-1}(\hat{Z}(j) - \hat{H}(j)\hat{X}(j)) \quad (12) \]

problem, the least square estimate is obtained by setting

\[ \frac{d(J(\hat{X}(j)))}{d(\hat{X}(j))} = 0 \quad (13) \]

By using equation (12) for \( \hat{X}(j) \) and carrying out the indicated partial differentiation, we obtain

\[ d(J(\hat{X}(j))) + d(X(j)) = \hat{H}(j)\hat{V}^{-1}(\hat{Z}(j) - \hat{H}(j)\hat{X}(j)) \quad (14) \]
The inclusion of error variance matrix results in a minimum variance unbiased estimate of the O-D matrix for interval $j$ expressed as:

$$\hat{X}(j) = (\hat{H}'(j)\hat{V}^{-1}\hat{H}(j))^{-1}\hat{H}'(j)\hat{V}^{-1}\hat{Z}(j)$$  \hspace{1cm} (15)

The test of the credibility of the results of the GLS model (estimated O-D volumes) is to find out how close the estimated link counts are to the observed link counts. The estimated link counts are not directly derived from the observed link counts, but rather are calculated using (OLS-GLS) estimated O-D volumes; therefore, the same data are not used both for evaluation and estimation. The statistic used to measure the degree of closeness between predicted and observed link counts is the root mean square error (rmse).

The rmse measures the degree of disagreement between two series of link flow values (estimated $\hat{z}(l,j)$ and observed $z(l,j)$) and it is given for each interval $j$ by:

$$rmse(j) = \sqrt{\frac{\sum_{l=1}^{n_j}(z(l,j) - \hat{z}(l,j))^2}{n_j}}$$  \hspace{1cm} (16)

The percent root mean square error is expressed as;

$$%rmse(j) = \left( \frac{rmse(j)}{Z(j)} \right) \times 100$$  \hspace{1cm} (17)

where

$$Z(j) = \left( \frac{1}{n_j} \right) \sum_{l=1}^{n_j} \hat{z}(l,j)$$

6.3 Parameter Estimation

The procedure for estimating the fraction of $r$th origin-destination volume that departed the origin during interval $t$, and is on link $l$ during interval $j$, $h(r, t, l, j)$. To estimate the fractions $h(r, t, l, j)$, link traffic counts must be expressed in terms of path flows. Letting $x_{path}(k, t)$ denote the traffic flow following path $k$ that left the origin during period $t$, $x_{path}(k, t)$ can be expressed as follows:
where \( p(k/t) \) is the probability of choosing path \( k \) going between an O-D pair \( r \) given that the trip left the origin during interval \( t \). The observed link traffic counts can be rewritten in terms of the path flow \( x_{path}(k, t) \) as follows:

\[
z(l, j) = \sum_{t = j-1}^{j} \sum_{k = 1}^{K} a(k, t, l, j) \cdot x_{path}(k, t) + v(l, j)
\]

(19)

where \( a(k, t, l, j) \) is the fraction of path flow \( x_{path}(k, t) \) contributing to link flow \( z(l, j) \). The assignment fraction \( h(r, t, l, j) \) can be expressed in terms of path-link incidence fractions \( a(k, t, l, j) \), and path choice probabilities \( p(k/t) \). Combining equation (18), and equation (19), the assignment fraction \( h(r, t, l, j) \) is expressed by the following relationship:

\[
h(r, t, l, j) = \sum_{k=1}^{K_r} a(k, t, l, j) \cdot p(k/t)
\]

(20)

\[
\tilde{h}(r, j-1, l, j) = \sum_{k=1}^{K_r} a(k, j-1, l, j) \cdot p(k/j-1)
\]

(22)

In the next section, steps taken to develop estimates of the parameter and \( a(k, t, l, j) \) is described. The estimation of parameter \( p(k/t) \) is already described in chapter 5 on assignment.

To estimate the path-incident fractions \( a(k, t, l, j) \) assume that the departure flow is uniformly spread over the interval length \( T \), and remains equally uniform while moving in the network. Based on the above assumptions, the fraction \( a(k, t, l, j) \) can be estimated as follows:
A freight traffic modeling approach for the waterway network in the Netherlands

\[ \alpha(k,t,l,j) = 1 - \frac{1}{T(t(k,t,l) - T(j - 1))} \quad \text{if} \quad T(t(k,t,l) - T(j - 1)) > 0 \]

\[ = 1 - \alpha(k,t,l,(j - 1)) \quad \text{if} \quad (j - 2)T < t(k,t,l) < (j - 1)T \]

\[ = 0 \quad \text{otherwise} \quad (24) \]

where \( t(k,t,l) \) is the arrival time on link \( l \) of the first vehicle on path \( k \) which departed the origin during interval \( t \), \( T \) is the length of an interval, and \( j \) is the present interval.

The procedure used to calculate the error variances uses several steps. In the first step, the residuals from the link flow estimates are used to calculate link flow error variances for each interval by using the normal distribution to group estimated link counts into low, medium, and high volume categories. The corresponding residuals are then grouped into low, medium, and high categories based on the grouping of the estimated link counts.

The logic for grouping the data is that observations with smaller variance receive a larger weight, and therefore have greater influence in the estimates obtained for the next interval.

The overall procedure is depicted in a flow chart below:

6.4 Model Validation

To test the validity of the procedure adopted for the time-dependent O-D estimation, the model was tested on a synthetic network. The network consists of 7 nodes and 7 links. Figure 1. The links are one directional from node 1 to node 7.

Data generated for this simulation study consists of the volumes on 6 links, plus travel time on all the links. Data was generated for 15 intervals of 10 minutes duration each. Data that needed to be calculated were link arrival times on each link on a path between each O-D pair, path choice probability for O-D pair (1-7) (the rest of the O-D pairs have only one route choice). The parameter \( r \) was estimated for the O-D pair (1-7) only. \( r \) was set equal to (1) for the other O-D pairs. This was due to the unidirectional links on the network.

6.4.1 Analysis Of The Results

An alternate OLS-GLS procedure was performed on the synthetic data. By an alternate OLS-GLS procedure it is meant that for each interval an OLS estimated O-D volumes and counts were found. These estimates were used to calculate the error variances. Next, a GLS estimate of origin-destination volumes was obtained using the inverse of the error variances as weights.
The alternate OLS-GLS procedure was performed for 2 days. Each day consists of 15 intervals of 10 minutes duration each. The estimated link counts were compared with the synthetic link counts. The measure of degree of closeness was the rmse, and %rmse. For this exercise, the true O-D volumes were known. Comparison between the estimated origin-destination volumes, and true origin-destination volumes indicated a close match, Table 1, Graphs 1-3.
There are 15 intervals of 10 minutes duration each. As is seen by the graph there are 6 links with simulated traffic and link 7 is just a dummy link. Link 7 is assumed to have zero traffic, and zero travel times. The simulated O-D volumes are not available, though at the time they were generated the same way as the traffic counts, and travel times. Therefore, it was possible to compare O-D estimates with the simulated O-D volumes.

There are 7 origin-destination pairs: (1,7), (1,3), (1,5), (2,3), (2,7), (4,5), (4,7).

<table>
<thead>
<tr>
<th>O-D Pair</th>
<th>Path One (links)</th>
<th>Path Two (links)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 7</td>
<td>1,2,3,7</td>
<td>4,5,6</td>
</tr>
<tr>
<td>1 - 3</td>
<td>1,2</td>
<td>none</td>
</tr>
<tr>
<td>1 - 5</td>
<td>4,5</td>
<td>none</td>
</tr>
<tr>
<td>2 - 3</td>
<td>2</td>
<td>none</td>
</tr>
<tr>
<td>2 - 7</td>
<td>2,3,7</td>
<td>none</td>
</tr>
<tr>
<td>4 - 5</td>
<td>5</td>
<td>none</td>
</tr>
<tr>
<td>4 - 7</td>
<td>4,6,7</td>
<td>none</td>
</tr>
</tbody>
</table>

**DAY 1 Simulated Counts**

<table>
<thead>
<tr>
<th>link 1</th>
<th>link 2</th>
<th>link 3</th>
<th>link 4</th>
<th>link 5</th>
<th>link 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>278</td>
<td>347</td>
<td>206</td>
<td>257</td>
<td>389</td>
<td>476</td>
</tr>
<tr>
<td>265</td>
<td>368</td>
<td>228</td>
<td>252</td>
<td>382</td>
<td>491</td>
</tr>
<tr>
<td>354</td>
<td>444</td>
<td>284</td>
<td>289</td>
<td>442</td>
<td>538</td>
</tr>
<tr>
<td>499</td>
<td>517</td>
<td>310</td>
<td>359</td>
<td>543</td>
<td>665</td>
</tr>
<tr>
<td>473</td>
<td>668</td>
<td>386</td>
<td>409</td>
<td>665</td>
<td>810</td>
</tr>
<tr>
<td>449</td>
<td>669</td>
<td>383</td>
<td>401</td>
<td>684</td>
<td>818</td>
</tr>
<tr>
<td>517</td>
<td>709</td>
<td>385</td>
<td>501</td>
<td>773</td>
<td>857</td>
</tr>
<tr>
<td>533</td>
<td>739</td>
<td>409</td>
<td>547</td>
<td>839</td>
<td>941</td>
</tr>
<tr>
<td>505</td>
<td>730</td>
<td>442</td>
<td>467</td>
<td>795</td>
<td>931</td>
</tr>
<tr>
<td>430</td>
<td>633</td>
<td>384</td>
<td>475</td>
<td>791</td>
<td>932</td>
</tr>
<tr>
<td>520</td>
<td>703</td>
<td>443</td>
<td>481</td>
<td>686</td>
<td>934</td>
</tr>
<tr>
<td>414</td>
<td>684</td>
<td>475</td>
<td>354</td>
<td>658</td>
<td>947</td>
</tr>
<tr>
<td>367</td>
<td>558</td>
<td>372</td>
<td>386</td>
<td>655</td>
<td>878</td>
</tr>
<tr>
<td>356</td>
<td>564</td>
<td>372</td>
<td>377</td>
<td>603</td>
<td>831</td>
</tr>
<tr>
<td>371</td>
<td>565</td>
<td>382</td>
<td>347</td>
<td>563</td>
<td>785</td>
</tr>
</tbody>
</table>

**Prior O-D Volumes for interval 1 Day 1**

<table>
<thead>
<tr>
<th>link 1</th>
<th>link 2</th>
<th>link 3</th>
<th>link 4</th>
<th>link 5</th>
<th>link 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>299</td>
<td>132</td>
<td>150</td>
<td>9</td>
<td>60</td>
<td>15</td>
</tr>
<tr>
<td>117</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Travel Times for 15 interval for Day 1**

<table>
<thead>
<tr>
<th>link 1</th>
<th>link 2</th>
<th>link 3</th>
<th>link 4</th>
<th>link 5</th>
<th>link 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.8735</td>
<td>0.8665</td>
<td>1.7120</td>
<td>2.0019</td>
<td>0.9832</td>
<td>1.3847</td>
</tr>
<tr>
<td>1.7901</td>
<td>0.8389</td>
<td>1.7859</td>
<td>1.9680</td>
<td>0.9875</td>
<td>1.4246</td>
</tr>
<tr>
<td>1.9689</td>
<td>0.9552</td>
<td>1.8976</td>
<td>2.1253</td>
<td>1.0163</td>
<td>1.6475</td>
</tr>
<tr>
<td>2.0040</td>
<td>1.0973</td>
<td>2.0640</td>
<td>2.1239</td>
<td>1.0510</td>
<td>1.9516</td>
</tr>
</tbody>
</table>
A modeling approach to freight traffic on the canals in the Netherlands

1.9547 1.1965 1.9438 2.0054 1.3111 1.7387
2.0044 1.3697 2.0401 2.0864 1.6526 1.6345
2.0022 1.8632 1.9823 2.1569 1.8065 1.7620
2.1957 2.1637 2.0392 2.2517 1.9849 1.8899
2.5206 1.8857 2.3659 2.7804 2.3259 1.7175
2.3991 2.2956 2.7804 2.3259 1.7175
2.0576 1.1830 1.8450 1.5032 1.7670
1.9486 1.0380 1.8450 1.5032 1.7670

DAY 2 Simulated Counts

<table>
<thead>
<tr>
<th>link 1</th>
<th>link 2</th>
<th>link 3</th>
<th>link 4</th>
<th>link 5</th>
<th>link 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>287</td>
<td>360</td>
<td>212</td>
<td>259</td>
<td>393</td>
<td>486</td>
</tr>
<tr>
<td>267</td>
<td>373</td>
<td>232</td>
<td>255</td>
<td>391</td>
<td>501</td>
</tr>
<tr>
<td>356</td>
<td>446</td>
<td>263</td>
<td>308</td>
<td>463</td>
<td>568</td>
</tr>
<tr>
<td>408</td>
<td>527</td>
<td>317</td>
<td>363</td>
<td>553</td>
<td>675</td>
</tr>
<tr>
<td>482</td>
<td>678</td>
<td>387</td>
<td>425</td>
<td>680</td>
<td>815</td>
</tr>
<tr>
<td>458</td>
<td>683</td>
<td>391</td>
<td>416</td>
<td>707</td>
<td>841</td>
</tr>
<tr>
<td>529</td>
<td>724</td>
<td>391</td>
<td>508</td>
<td>790</td>
<td>871</td>
</tr>
<tr>
<td>556</td>
<td>767</td>
<td>429</td>
<td>555</td>
<td>846</td>
<td>949</td>
</tr>
<tr>
<td>500</td>
<td>737</td>
<td>432</td>
<td>496</td>
<td>842</td>
<td>957</td>
</tr>
<tr>
<td>464</td>
<td>680</td>
<td>422</td>
<td>463</td>
<td>773</td>
<td>954</td>
</tr>
<tr>
<td>512</td>
<td>697</td>
<td>435</td>
<td>500</td>
<td>723</td>
<td>957</td>
</tr>
<tr>
<td>417</td>
<td>688</td>
<td>475</td>
<td>375</td>
<td>687</td>
<td>974</td>
</tr>
<tr>
<td>366</td>
<td>561</td>
<td>369</td>
<td>396</td>
<td>674</td>
<td>900</td>
</tr>
<tr>
<td>376</td>
<td>591</td>
<td>402</td>
<td>365</td>
<td>593</td>
<td>848</td>
</tr>
<tr>
<td>380</td>
<td>579</td>
<td>386</td>
<td>368</td>
<td>594</td>
<td>815</td>
</tr>
</tbody>
</table>

Travel Times for 15 interval for Day 2

<table>
<thead>
<tr>
<th>link 1</th>
<th>link 2</th>
<th>link 3</th>
<th>link 4</th>
<th>link 5</th>
<th>link 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.9100</td>
<td>0.9000</td>
<td>1.8000</td>
<td>2.1000</td>
<td>0.9900</td>
<td>1.4400</td>
</tr>
<tr>
<td>1.8700</td>
<td>0.8800</td>
<td>1.8500</td>
<td>2.0500</td>
<td>1.0300</td>
<td>1.4700</td>
</tr>
<tr>
<td>1.9900</td>
<td>0.9900</td>
<td>2.0000</td>
<td>2.1300</td>
<td>1.0200</td>
<td>1.7100</td>
</tr>
<tr>
<td>2.0500</td>
<td>1.1000</td>
<td>2.1000</td>
<td>2.1400</td>
<td>1.0800</td>
<td>2.0300</td>
</tr>
<tr>
<td>2.0400</td>
<td>1.2500</td>
<td>2.0800</td>
<td>2.1200</td>
<td>1.3600</td>
<td>1.8000</td>
</tr>
<tr>
<td>2.1000</td>
<td>1.4400</td>
<td>2.1000</td>
<td>2.1600</td>
<td>1.7100</td>
<td>1.7200</td>
</tr>
<tr>
<td>2.0800</td>
<td>1.9000</td>
<td>2.0700</td>
<td>2.2000</td>
<td>1.9000</td>
<td>1.8300</td>
</tr>
<tr>
<td>2.3500</td>
<td>2.2000</td>
<td>2.1000</td>
<td>2.3500</td>
<td>2.1000</td>
<td>2.0200</td>
</tr>
<tr>
<td>2.5500</td>
<td>2.8000</td>
<td>2.1300</td>
<td>2.4000</td>
<td>2.7000</td>
<td>2.2200</td>
</tr>
<tr>
<td>2.9100</td>
<td>3.0000</td>
<td>2.4000</td>
<td>2.9000</td>
<td>2.9000</td>
<td>2.3000</td>
</tr>
<tr>
<td>3.1300</td>
<td>2.6500</td>
<td>2.3800</td>
<td>3.2000</td>
<td>3.0000</td>
<td>1.8000</td>
</tr>
<tr>
<td>2.6000</td>
<td>1.9000</td>
<td>2.5000</td>
<td>2.8000</td>
<td>2.4000</td>
<td>1.7600</td>
</tr>
<tr>
<td>2.4500</td>
<td>1.6000</td>
<td>2.3300</td>
<td>2.2600</td>
<td>1.9000</td>
<td>1.8100</td>
</tr>
<tr>
<td>2.1300</td>
<td>1.2000</td>
<td>2.2200</td>
<td>1.9800</td>
<td>1.6000</td>
<td>1.8300</td>
</tr>
<tr>
<td>2.0000</td>
<td>1.1000</td>
<td>2.2500</td>
<td>2.0100</td>
<td>1.2000</td>
<td>2.0000</td>
</tr>
</tbody>
</table>

O-D Volumes Estimated for Day 1, 15 Intervals
A modeling approach to freight traffic on the canals in the Netherlands

<table>
<thead>
<tr>
<th>Q-D Volumes Estimated for Day 2, 15 Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>352  303  333  385  450  442  500  538  512  490  501  418  397  394  400</td>
</tr>
<tr>
<td>177  112  171  204  255  247  288  324  298  271  284  210  193  186  187</td>
</tr>
<tr>
<td>142  88   108  152  190  188  246  271  238  225  234  155  145  138  146</td>
</tr>
<tr>
<td>9     4    0    1    3    4    5    6    6    5    5    4    3    4    4</td>
</tr>
<tr>
<td>143  116  126  144  177  186  190  211  232  217  211  201  169  189  193</td>
</tr>
<tr>
<td>16    12   8    10   15   17   20   21   22   22   21   17   17   14   14</td>
</tr>
<tr>
<td>176  140  168  190  242  246  272  283  287  289  287  254  252  249  248</td>
</tr>
</tbody>
</table>
Table 1 - Comparison of Actual Versus Estimated Link Flows - GLS run - Sample Data - Day 1 and Day 2

<table>
<thead>
<tr>
<th>Interval</th>
<th>Day 1 RMSE</th>
<th>%RMSE</th>
<th>Day 2 RMSE</th>
<th>%RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>58</td>
<td>18</td>
<td>66</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>7</td>
<td>36</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>43</td>
<td>11</td>
<td>36</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>55</td>
<td>12</td>
<td>50</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>11</td>
<td>64</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>66</td>
<td>12</td>
<td>57</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>83</td>
<td>13</td>
<td>71</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>93</td>
<td>14</td>
<td>87</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>101</td>
<td>16</td>
<td>95</td>
<td>14</td>
</tr>
<tr>
<td>10</td>
<td>99</td>
<td>16</td>
<td>89</td>
<td>13</td>
</tr>
<tr>
<td>11</td>
<td>80</td>
<td>13</td>
<td>72</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>78</td>
<td>13</td>
<td>68</td>
<td>11</td>
</tr>
<tr>
<td>13</td>
<td>62</td>
<td>12</td>
<td>53</td>
<td>10</td>
</tr>
<tr>
<td>14</td>
<td>47</td>
<td>9</td>
<td>40</td>
<td>8</td>
</tr>
</tbody>
</table>
Flow Chart Depicting the Overall VIADUCT Model Procedure

1. Data Preparation
   - Calculation of link travel times
   - Calculation of path total travel times

2. Calculation of the Assignment Matrix
   - Probabilities of choosing path $k \in K$ during interval $(t)$
   - Interval $t=1$, day 1 only requires an Apriori O-D table
   - Link traffic counts interval $(t)$
   - Starting from $t>1$ prior interval $(t-1)$ O-D Estimates
   - Ordinary Least Square Estimation (OLS)
   - O-D Error Variance Calculation
   - Augmented Inverse Error Variance Calculation
   - Interval $t=1$, day 1 only requires an Apriori O-D table
   - Link traffic counts interval $(t)$
   - Starting from $t>1$ prior interval $(t-1)$ O-D Estimates
   - Generalized Least Square (GLS)
   - GLS O-D estimation interval $(t)$
   - Link estimation Procedure interval $(t)$

3. Check the validity of O-D estimates by calculating RMSE and %RMSE for the estimated link counts

$t=t+1$
7. FORECASTING MODELS
7. FORECASTING MODELS

7.1 Introduction

This chapter will expand more on the forecasting options given in chapter two. Two forecasting procedures were suggested in chapter one. The objective of both options is to forecast waterway freight traffic for the future. The difference between the two models is in the way that the endogenous variables are dealt with. As is common knowledge there are two types of variables that are used in modeling: the endogenous variables and the exogenous variables. The endogenous variables are those variables that are to be forecasted as part of the modeling exercise like the ship matrix (that is the number of ships that go between origin-destinations by ship type/size). The exogenous variables are those that are required to run the model and are originated externally to the models themselves. Typical examples are the network data, and the matrix of tons carried between origins and destinations (the description of this matrix is provided in chapter 2). The values of these variables should be provided for the base year and for the forecast years.

The importance of the exogenous variables in influencing the accuracy of the whole modeling exercise is very high. The specification errors are less significant than the errors in exogenous variables. In many cases the errors are the result of errors in official forecasts of exogenous data. The difference between the two models of forecasting suggested in chapter one is in the level of detail and the degree of disaggregation.

7.2 A Static Forecast Model

This is a static model with respect to the endogenous variables. The parameters that are used in the base year model are used in the forecast year model as well. These parameters are calibrated in the base year, and are used unchanged in the forecast year. The exogenous variables are externally provided as can be seen in the flow chart. In the base year the waterway freight trip table is an observed data. The objective is to modify the ship matrix to match the observed trip table data as closely as possible. This is ideal because here by trying to match the ship matrix with the trip table data an attempt is made to find the best estimate of the parameters involved. This is consistency on an aggregate level. On a disaggregate level, attempt is made to match the estimated counts of ships on the canals that have detectors with the counts obtained from the detectors on the canals. The estimation process is recursive and continues till the results are satisfactory. The details included in the base year model are an attempt to provide more reasonable estimates of the desired output namely the ship matrix.
The details of the box entitled dynamic logit assignment and trip table estimation is given in chapters five and six. In the forecast year the same parameters that are used in the base year ALMC are also used in the forecast year ALMC. In the forecast year there is no dynamic trip table estimation as it is not possible to forecast counts by time of day for the next thirty years with conventional methods that are used in transportation modeling. Therefore, a static assignment model would suffice in this case. The problem encountered in the forecasting model is that there is no way of verifying the results. Therefore, the objective here is to repeat the procedure until convergence is achieved, that means until the output can not be improved any further.

7.3 A Dynamic Forecast Model

The difference between this model and the previous one is that the same dynamic origin-destination matrix estimation and assignment model is used for the forecast year as well as the base year. As was mentioned earlier it is not easy to forecast counts for thirty years from now.
Here a method is suggested for forecasting counts by time of day for the forecast year. To forecast the counts by time of day the ideal approach is to use the method of dynamical systems. The principle branch of use suggested is the time-evolution law.

In the most general setting this is a rule that determines the state of the system at each moment of time \( t \) from its states at previous times. The system here is the counts by time of day. For such a system an autonomous dynamic system with time invariant evolution law can be used. Eventually a non autonomous model could be considered in future research.

In other words, the result of time evolution will depend only on the initial position of the system and on the length of the evolution but not on the moment when the state of the system was initially registered. Thus if the system was initially at a state \( x \in \mathcal{X} \), it will find itself after time \( t \) at a new state, which is uniquely determined by \( x \) and \( t \), and thus can be denoted by \( F(x,t) \).
If \( t \) is fixed then a transformation is obtained \( \varphi': x \rightarrow F(x,t) \) of the phase space into itself. These transformations for different \( t \) are related to each other. Namely, the evolution of the state \( x \) for time \( s+t \) can be accomplished by first applying the transformation \( \varphi' \) to \( x \) and then by applying \( \varphi' \) to the new state \( \varphi' (x) \). The most significant characteristics of using a dynamical system method is it's emphasis on asymptotic behavior, especially in the presence of non-trivial recurrence, that is properties related with the behavior as time goes to infinity.

Historically, smooth continuous-time dynamical systems appeared initially because of Newton’s discovery that the motions of mechanical objects can be described by second-order differential equations. But in more general view many other natural and social phenomena such as chemical reactions, population growth, or the number of ships moving at different points on a path between an origin-destination may be modeled with various degrees of accuracy by systems of ordinary differential equations. Since at any point in time the number of ships between any origin-destination pair is a function of market forces that react, and the behavior of carriers that use the ships to carry cargo as far as choosing a path and maintenance of the ships are concerned.

On the other hand the trip table data can be forecasted by any time series methods and does not require complicated forecasting procedure. The advantage of using the dynamic forecasting method is that it provides a detailed look at the state of the freight transport system in the waterway network in a far future. Many elements are considered to assure the reliability of such forecasts. These elements are: the counts data that are forecasted using a dynamical systems theory, the trip data that is forecasted through some time series method, and all other exogenous input variables that are forecasted by independent methods.

### Conclusion

Both flow charts provide a rough outlook of the procedure. The details are explained in various chapters of this report. At this point none of the forecasting methods have been implemented and are basically theoretical. Further research is required to validate the approaches suggested in this chapter. The use of the dynamical systems theory in analyzing the behavior of ships on canals by time of day should be researched. The overall methodologies though sound, still need to be implemented with real data.
8. CONCLUSION
A freight traffic modeling approach for the waterway network in the Netherlands
8. CONCLUSION

In this chapter the aim is to give general comments about the waterway freight transport and point out the requirements of future research in this area. Given road congestion, which in large part is caused by truck movement on the roads, facilitating the movement of freight on the waterway network system is likely to have a major impact on economic development. Considering the importance of waterway freight transport it is surprising that much less research has been undertaken on modeling this type of movement than the effort allocated to other types of freight movement.

There are several reasons for this:

Only few countries have useful network of canals and rivers that can be used for freight transport. Waterway freight transport involves many players. There are the industrial firms sending or receiving the goods, the shippers organizing the consignment, in some cases both are the same. Then there are the carriers who undertake the movement of goods. Given the number of players there are always conflicting objectives that are difficult to model in detail in practice.

The most traditional approach to modeling waterway freight transport is to use the conventional four-step aggregate demand models with minor adaptations specific to waterway freight transport. A typical approach would be to estimate freight generations and attractions by zone. Different methods are employed such as direct surveys of demand and supply of major products (sugar, petroleum products, iron ore, coal, cement, sand, fertilizers, grains, etc); or macroeconomic models; or growth factor methods which are typically used for forecasting future trip ends; or zonal multiple linear regression which is often used to get more aggregate measures of freight generations and attractions.

Then, the generated volumes are distributed to satisfy 'trip-end' generation and attraction constraints using typically a gravity model, but could also consider linear programming. The general form of a gravity model is as follows:

\[ T_{ij}^k = A_i^k B_j^k O_i^k D_j^k \exp(-B^k C_{ij}^k) \]

\( k \) represents good type
\( T_{ij}^k \) are tons of product \( k \) moved from \( i \) to \( j \)
\( A_i^k, B_j^k \) are balancing factors
\( O_i^k, D_j^k \) are supply and demand for product \( k \) at zone \( i \) or \( j \)
\( \beta^k \) are calibration parameters for product \( k \)
\( C_{ij} \) are generalized transport costs per ton of product \( k \) between zones \( i \) and \( j \).

Typically the cost function is presented as:

\[
C_{ij} = f_{ij} + b_1s_{ij} + b_2\sigma s_{ij} + b_3w_{ij} + b_4p_{ij}
\]

- \( f_{ij} \) is the out-of-pocket charge for using a service from \( i \) to \( j \)
- \( s_{ij} \) is the zone to zone travel time between \( i \) and \( j \)
- \( \sigma s_{ij} \) is the variability of travel time \( s_{ij} \)
- \( w_{ij} \) is the waiting time for delivery of goods
- \( p_{ij} \) is the probability of loss or damage to goods during transport
- \( b_n \) are proportional to the value of the goods

Another approach to distribution modeling is linear programming (LP). The constraints are usually supply and demand.

\[
\text{Minimize } Z = \sum_{k} T^k_y C^k_y
\]

subject to

\[
\sum_{i} T^k_y = D^k_j \\
\sum_{j} T^k_y = O^k_i \\
T^k_y \geq 0
\]

Overall it is perceived that the gravity model performs better when the transportation costs are not a major costs in the generalized cost function. The gravity model is considered to be flexible when dealing with several origins and destinations by being able to change the value of \( \beta \) and therefore vary the relative importance of cost compared with supply and demand constraints. The LP model is considered to perform better when transportation costs are a major portion of the generalized cost function such as the case of transporting low-cost high-bulk goods like cement, sand, gravel, etc. Also when there are few origins and destinations, meaning few fixed demand and supply. Ortuzar, and Willumsen 1977.

The ideas presented in this report seem to disagree with the prevailing belief that the gravity model is the best model to use in case of waterway freight transport. The problem is not simple but rather complex. There are many players involved in transporting goods as was mentioned earlier, therefore, there are many origins and destinations meaning many shippers who have demand for many destinations from many origins (this is due to the fact that warehouses are not concentrated and are dispersed). There are also by necessity many carriers. The transportation
costs are an important part of the generalized cost function but they are not the only important part, costs due to other factors such as maintenance cost, cost of travelling empty during return trips to origins or bases other than origin, depreciation costs, personnel costs are only few of the other costs involved on the part of the carrier. The costs incurred by the shipper influence the price he/she is willing to pay for the transport of goods, which directly affects the carrier. These costs should also be considered and incorporated in the models. In neither of these models the interaction between the shippers and the carriers is directly modeled.

The aim of this report is to incorporate the complexities involved in this problem into a mathematical model. For example the costs considered in the modified linear programming approach of this report are the generalized costs. The unit price of transporting products is determined based on the utility of the shipper or carrier, which is actually the generalized cost.
These unit costs are the coefficients in the objective function of the linear programming model. Two major ideas are incorporated in calculating these unit costs. One is the use of the Logit probability models to simulate both the carrier and the shippers choice process. Then the notion of the elasticity of demand is used to further calculate the unit costs. In modeling with linear programming typically it is the shipper’s activities that are modeled. It is assumed that it is a shipper’s decision as to which carrier should be used to deliver the goods to their destination. Here, this assumption is taken to be incorrect. In this report it is assumed that both shippers and carriers are active participants in the market and therefore, interact with each other. So, not only the shipper’s decisions are modeled but also the carriers. It is also assumed that only after the market reaches an equilibrium state that the goods will be moved to their destinations. By saying the market reaches equilibrium it is implied that both shippers and carriers interact in the market and at the end each shipper and carrier finds his or hers suitable match. At the end all participants in the market are matched. Those that are not matched will have to leave the market. The market will always reach equilibrium. Those who are not matched are not able to survive the forces of the market place.

Finally there is an assignment of origin-destination movements to ship types/sizes and routes. The assignment is now the carrier’s decision. It is the choice of the best route to take the goods from origin to destination. In the case of the waterway freight transport, there are two types of constraints: one is due to the physical constraints of the canals. These are water levels during different times of year, and canal’s depth. The second are constraints due to the characteristics of the ships. These constraints are width, length, depth, and size. An assignment model should be able to deal with these constraints. In this report the idea of an assignment model is modified. It is assumed that the number of possible paths is fixed. There are only a certain number of canals that connect origins and destinations. What can vary is the probability that a carrier might choose any of these paths at different times of day and during different times of the year. This probability is not only the function of travel time, but delay due to queuing and bridge/lock operation at the locks and bridges, and the number of stops at the locks and bridges due to queuing and operational factors. Other factors such as carrier’s deadline to deliver goods, carrier’s perception of which of the routes are the shortest, and his expectations as to the level/number of difficult maneuvers (number of curves, and abrupt changes in the canal’s width) play an important role in calculating the assignment probabilities. In short, the idea is that the shortest route is a relative concept. The route being short is relative to the time of day, carrier’s perception and expectation, and the level of stops and delays that one has to encounter.

8.1 FUTURE WORK

The modeling approach presented in this report has not been tested. Future research in this area would entail implementing the procedures suggested. Check for the validity and soundness of the
results and modify the modules as are necessary. The author did not have the opportunity to test different modules of the system that is proposed based on real data. This was due to the limited period of author's stay in the Technical University of Delft. Given the level of technological advancement and know how; the use of the conventional four-step modeling in a traditional manner is not anymore justified. Attempt should be made to modify the four-step planning models in order to make them compatible with the needs of today's societies. The present report aims at doing just that in the hope to open new doors and avenues of research and interest.
A freight traffic modeling approach for the waterway network in the Netherlands
9. APPENDIX A
10. APPENDIX B
Price And Output Under Perfect Competition With Calculus

The following shows with the use of calculus the first- and second order conditions for
the output that a perfectly competitive carrier must produce in order to maximize total
profit.

Total profit (π) are equal to total revenue (TR) minus total costs (TC). That is,

= TR - TC

where π, TR, and TC are all functions of output (number of vessels used) (Q). Note that
the type of vessel is not important here because the type of cargo already decides what
type vessel should used.

Taking the first derivative of π with respect to Q an setting it equal to zero gives

\[ \frac{d\pi}{dQ} = \frac{d(TR)}{dQ} - \frac{d(TC)}{dQ} = 0 \]

\[ \frac{d(TR)}{dQ} = \frac{d(TC)}{dQ} \]

\[ \frac{d(TR)}{dQ} = MR \quad \frac{d(TC)}{dQ} = MC \]

and

\[ MR = MC \]

Since under the perfect competition MR = P, the first-order condition for profit
maximization for a perfectly competitive carrier becomes

\[ P = MR = MC \]

The above is only the first-order condition for maximizartion (and minimization). The second-
order condition for profit maximization requires that the second derivative of π with respect to
negative.
That is

$$\frac{d^2 \pi}{d^2 Q} = \frac{d^2 (TR)}{d^2 Q} - \frac{d^2 (TC)}{d^2 Q} < 0$$

so that

$$\frac{d^2 (TR)}{d^2 Q} < \frac{d^2 (TC)}{d^2 Q}$$

Since under perfect competition the MR curve is horizontal, this means that the MC curve must be rising at the point where MR=MC for the carrier to maximize its total profits (or minimize its total losses).
11. APPENDIX D
Sample Space And Events

11.1.1 Random Experiments:

Any process of observation is referred to as an experiment. The results of an observation are called the outcomes of the experiment. An experiment is called a random experiment if the outcome cannot be predicted. Typical examples of random experiments are drawing a card from a deck, the roll of a die, or selecting signal for transmission from several messages.

11.1.2 Sample Space:

The set of all possible outcomes of the experiment is called the sample space (or universal set), and it is denoted by $S$. An element in $S$ is called a sample point. Each outcome of a random experiment corresponds to a sample point. Note that any particular experiment can often have many different sample spaces depending on the observation of interest. A sample space $S$ is said to be discrete if it consists of a finite number of sample points or countably infinite sample points. A set is called countable if its elements can be placed in a one-to-one correspondence with the positive integers. A sample space $S$ is said to be continuous if the sample points constitute a continuum.

11.1.3 Events

If $\xi$ is an element of $S$ (or belongs to $S$), then

$$\xi \in S$$

if $\xi$ is not an element of $S$, then $\xi \notin S$. A set $A$ is called a subset of $B$, denoted by $A \subseteq B$ if every element of $A$ is an element of $B$. Any subset of the sample space $S$ is called an event. A sample point of $S$ is often referred to as an elementary event. Note that the sample space $S$ is the subset of itself $S \subseteq S$. Since $S$ is the set of all possible outcomes, it is often called the certain event.

Example 1.1 As a result of the analysis of traffic on the roads in a section of Den Haag, Government has decided to build a bridge to connect two segments that were not previously connected. This action is anticipated to alleviate congestion on the roads. An event is when the bridge is in operation; i.e. it is lifted up to let a freight ship pass. Let's denote the sample space $S$. $S$ contains $n$ elementary events. Let $S = \{s_1, s_2, \ldots, s_n\}$. Let $\Omega$ be the family of all subsets of $S$. ($\Omega$ is sometimes referred to as the power set of $S$). Let $S_i$ be the set consisting of two statements, that is,
\( S_i = \{ \text{bridge in operation, the } s_i \text{ is in; bridge is not in operation the } s_i \text{ is not in} \} \)

Then \( \Omega \) can be represented as the Cartesian product

\[
\Omega = S_1 \times S_2 \times \ldots \times S_n \\
= \{(s_1, s_2, \ldots, s_n); s_i \in S_i \text{ for } i = 1, 2, \ldots, n\}
\]

Since each subset of \( S \) can be uniquely characterized by an element in the above Cartesian product, the number of elements in \( \Omega \) by

\[
N(\Omega) = n(S_1)n(S_2) \ldots n(S_n) = 2^n
\]

Where \( n(S_i) = \text{number of elements in } S_i = 2 \).

11.2 Algebra Of Sets

11.2.1 Set Operations

Union:

The union of two sets \( A \) and \( B \), denoted \( A \cup B \), is the set containing all elements in either \( A \) or \( B \) or both.

\[
A \cup B = \{ \xi : \xi \in A \text{ or } \xi \in B \}
\]

Intersection:

The intersection of sets \( A \) and \( B \), denoted \( A \cap B \), is the set containing all elements in both \( A \) and \( B \).

\[
A \cap B = \{ \xi : \xi \in A \text{ and } \xi \in B \}
\]

Mutually exclusive events:

Two sets \( A \) and \( B \) are called mutually exclusive if they contain no common element, that is,

\[
A \cap B = \emptyset
\]

\( \emptyset = \text{A set containing no elements (null set).} \)
11.3 The Notion Of Probability

When real numbers are assigned to the events defined in a sample space S they are known as the probability measures. Consider a sample space S of random experiments, and let A be a specific event defined in S. Then if the random experiment is repeated n times, and if event A occurs n(A) times, then the probability of event A, denoted by P(A), is defined as

\[ P(A) = \lim_{n \to \infty} \frac{n(A)}{n} \]

where

\[ 0 \leq \frac{n(A)}{n} \leq 1 \]

1. \( A \cap B = \emptyset \), then \( n(A \cup B) = n(A) + n(B) \)
2. \( n(A \cup B) / n = n(A)/n + n(B)/n \)

11.4 Equally Likely Events

Finite Sample Space:

Let S denote a sample space S with n finite elements

\[ S = \{ \xi_1, \xi_2, \ldots, \xi_n \} \]

The \( \xi_s \)s are elementary events. Let \( P(\xi_i) = P_i \), then

1. \( 0 \leq P_i \leq 1 \quad i = 1,2,\ldots,n \)
2. \( \sum P_i = P_1 + P_2 + \ldots + P_n = 1 \)

A = \( \cup \xi_i \), \( P(A) = \sum P(\xi_i) = \sum P_i \)

Equally likely Events:

If all elementary events \( \xi_i \), \( i = 1,2,\ldots,n \) are equally likely, which implies that

\[ P_1 = P_2 = \ldots = P_n \]

Then from condition 2 above, the probabilities are equal to:

\[ P_i = \frac{1}{n} \quad i = 1,2,\ldots,n \]
And \[ P(A) = \frac{n(A)}{n} \]

\( n(A) \) is the number of outcomes belonging to event \( A \) and \( n \) is the number of sample points in \( S \).

An Example:
A committee of 5 persons is to be selected randomly from a group of 5 men and 10 women.

Find the probability that a committee consists of 2 men and 3 women.
Find the probability that the committee consists of all women.

The sample space \( S \) consists of 15 people. The number of total outcomes is given
\[ n(S) = \binom{15}{5} \]

When selection is random it means that each of the outcomes is equally likely. Let \( A \) be the event that the committee consists of 5 men and 3 women. Then the number of outcomes belonging to event \( A \) is given by:
\[ n(A) = \binom{5}{2} \binom{10}{3} \]

Thus the probability of selecting 2 men and 3 women at random is
\[ P(A) = \frac{n(A)}{n(S)} = \frac{\binom{5}{2} \binom{10}{3}}{\binom{15}{5}} = \frac{400}{1001} \approx 0.4 \]

Let \( B \) be the event that the committee consists of all women. Then the number of outcomes belonging to \( B \) is
\[ n(B) = \binom{5}{0} \binom{10}{5} \]

Thus the probability that all women are selected is
\[ P(B) = \frac{n(B)}{n(S)} = \frac{\binom{5}{0} \binom{10}{5}}{\binom{15}{5}} = \frac{36}{429} = 0.084 \]
12. References


A freight traffic modeling approach for the waterway network in the Netherlands
A freight traffic modeling approach for the waterway network in the Netherlands

13. RECENT PUBLICATIONS OF THE TRANSPORTATION AND TRAFFIC ENGINEERING SECTION.


Dijker, Th., P.H.L. Bovy, and R.G.M.M. Vermijs (October 1997), *Car-following under non-congested and congested conditions*, Delft University of Technology, Transportation Planning and Traffic Engineering Section report vk2206.301.


These publications can be ordered from:
Delft University of Technology
Transportation Planning and Traffic Engineering Section
P.O. Box 5048
NL-2600 GA Delft, The Netherlands
Tel:   (+31) 15 278 1681
Fax:   (+31) 15 278 3179
e-mail: transportation@CTiG.tudelft.nl
web:   http://www.ct.tudelft.nl/verkeer/vk.html
De sectie Verkeerskunde houdt zich bezig met onderwijs en onderzoek op het gebied van planning, ontwerp en exploitatie van verkeerssystemen voor personen en goederen, alsmede het functioneel ontwerp van verkeersinfrastructuur.

De sectie Verkeerskunde participeert in de onderzoekschool TRAIL.