2472 S

Stellingen van Erik de Romph
behorende bij het proefschrift
A Dynamic Traffic Assignment Model
Theory & Applications

I. Het gebruik van een verkeersmodel voor het maken van een verkeersvoorspelling voor het komende uur is alleen nuttig als er verstoringen zijn, zoals ongevallen, wegwerkzaamheden of extreme weersomstandigheden. Tijdens normale omstandigheden geeft een historisch gemiddelde meestal een betere voorspelling.

II. Een goede visualisatie van de resultaten van een dynamisch toedelingsmodel is van essentieel belang. Gedurende de ontwikkeling van het model komen onvoorziene bijverschijnselen van een nieuwe theorie sneller aan het licht. Tijdens het uiteindelijke gebruik van het model wordt sneller inzicht verkregen in de effecten van verkeersbeheersingmaatregelen. Tevens verhoogt een goede grafische presentatie de aantrekkelijkheid en de kans op het daadwerkelijke gebruik van dit soort verkeersmodellen.

III. De introductie van telematica in de verkeerskunde vraagt om meer kennis van informatica in de opleiding tot verkeerskundige.

IV. Het vlot kunnen bedienen van een populair software pakket moet niet worden verward met informatica.

V. Een dynamisch toedelingsmodel moet in de eerste plaats beoordeeld worden op de dynamiek (evolutie van de verkeersstroom) en niet op bepaalde extra's, zoals afslagverboden.

VI. Wil men het dynamisch gebruikersevenwicht in analogie met het principe van Wardrop definiëren, dan moeten de routes gebaseerd zijn op de reistijden die men daadwerkelijk tijdens de route ondervindt, en niet op de reistijden die gelden op het moment van vertrek.

VII. Intelligent Vehicle/Highway Systems moeten niet gezien worden als de enige oplossing voor het congestie probleem. Het bevordert de veiligheid en verbetert de reisomstandigheden, maar dient gecombineerd te worden met lange termijn oplossingen, zoals het aanpassen van de ruimtelijke ordening en het stimuleren van alternatieve vervoerwijzen.

VIII. Het operating systeem Linux[1] is vele malen superieur aan het operating systeem MS-DOS, de grootste aantrekkingskracht van Linux zit echter in de mondiale samenwerking en het onzelfzuchtig enthousiasme waarin het tot stand gekomen is.

IX. Zolang sport commentatoren het begrip "professionele overtreding" hanteren, kan het begrip "sportiviteit" beter uit het woordenboek worden geschrapt.

X. De dromen die men heeft op jonge leeftijd en de berusting die men vaak heeft op oude leeftijd, pleiten voor een omgekeerd salarisverloop van een carrière.

XI. Het afschaffen van de militaire dienstplicht zal ongetwijfeld gevolgen hebben voor de instroom van AIO's aan de universiteiten in Nederland.

1. Linux is een UNIX variant voor PCs, en is volledig gratis beschikbaar via Internet.

A DYNAMIC TRAFFIC ASSIGNMENT MODEL

THEORY & APPLICATIONS

A DYNAMIC TRAFFIC ASSIGNMENT MODEL

THEORY & APPLICATIONS

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus,
Prof.ir. K.F. Wakker
in het openbaar te verdedigen
ten overstaan van een commissie
aangewezen door het College van Dekanen
op maandag 28 november 1994 om 16.00 uur

door

Erik DE ROMPH
informatica ingenieur
geboren te Velsen

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. ir. R. Hamerslag

Samenstelling promotiecommissie:

Prof. dr. P. van Beek
Prof. dr. ir. P.H.L. Bovy
Prof. drs. B.K. Brussaard
Prof. dr. ir. J.J.M. Evers
Prof. dr. ir. R.E.C.M. van der Heijden
Prof. dr. A.G. Hobeika

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Romph, Erik de

A dynamic traffic assignment model: theory and applications
/ Erik de Romph. - [S.1 : s.n.] Delft :
Universiteitsdrukkerij Delft). - Ill.
Thesis Technische Universiteit Delft. - With index, ref. -
With summary in Dutch.
ISBN: 90-9007710-3
NUGI: 849
Subject headings: dynamic traffic management / traffic
models

Printed by: Universiteitsdrukkerij Delft
Cover design: Yolande Weerdenburg / Erik de Romph

This dissertation was completed in July 1994;
The hottest July ever recorded to date in the Netherlands
(registration of meteorological data started in 1706).

Published by:

E. de Romph
Schrijnwerkerslaan 30
3454 DS De Meern
The Netherlands

Distributed by:

Delft University of Technology
Faculty of Civil Engineering
Department of Infrastructure
Transport Planning & Traffic Engineering Section
P.O. Box 5048
2600 GA Delft
The Netherlands
tel: (+) 31 15 784609
fax: (+) 31 15 783179

PREFACE

This research was carried out at the Traffic Section of the Department of Infrastructure at the Faculty of Civil Engineering of Delft University of Technology. It is part of a wider research project towards the possibilities and the realization of Road Transport Informatics and Intelligent Vehicle/Highway Systems. I would like to thank Rudi Hamerslag who saw the need for this research at an early stage and for his support and ideas throughout the project.

During the research several other departments and faculties were involved in this project. Most thanks go to Rik van Grol of the Computational Physics section at the Faculty of Applied Physics. His ideas, modelling experience and programming skills gave the project an enormous momentum. I would also like to thank Peter Maas from the same section, who did a lot of programming and research towards the discretization aspects of the model.
Within my own department I would like to thank Ben Immers for the support and the discussions we had, Nanne van der Zijpp for giving me more mathematical insight, and the work he did on the Amsterdam case, Marcel Westerman for his clear overview of Road Transport Informatics, Jim Stada, for the discussions we had about computers, and Lori Tavassy and Adri Nijsse for their criticism and support. Together with the other members of the department they provided me with an interesting environment to work in.

During the research I spent one semester at the *Center for Transportation Research* of *Virginia Polytechnic & State University (Virginia Tech)*. I would like to thank the director of the Center, Antoine Hobeika, for his hospitality and support, also Kaan Ozbay and Sam Subramaniam for the discussions we had and who made it fun to be there.
I am also thankful to the *Laboratory for Scientific Visualization* at Virginia Tech for making their hardware available and their enthusiasm for visualisation in traffic, and *Virginia Department of Transportation* (VDOT) and *Washington Metropolitan Council of Governments* (COG) who provided me with the necessary data. Financial support came from *NWO* (Netherlands Organization for Scientific Research) and *University funds Delft*.

Although I have never met him, I would like to thank Linus Torvalds, the author of the LINUX operating system. His efforts, to provide a free UNIX-like operating system for PC's, made it possible for me to do a lot of work at home, and increased my computer knowledge significantly. Furthermore, the time I spent on LINUX was fun.

Several people reviewed early parts of this dissertation. I am thankful to Jon Fricker and Eric van Berkum for doing this, and Miranda Aldham-Breary who checked my English.

Above all, I would like to thank my parents, for their support, interest and tolerance throughout the research.

Erik de Romph

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Problem Definition and Research Approach

The introduction of the automobile has changed our world. It has stimulated the economy and increased our independence and mobility. The enormous growth of traffic, however, has resulted in heavy congestion in and around urban areas, causing serious problems for the economy because of increased delays, and for the environment, due to increased emissions.
One solution to resolve congestion problems would be to expand the roadway system, but because of spatial and environmental limitations, especially in urban areas, such a solution is not always well advised. Another, more preferable, solution would be to reduce car use by changing land use patterns and encouraging people to travel by foot, bicycle, public transport and high occupancy vehicles. This approach is known as *demand management*. Changing land use patterns and human behaviour is, however, not easy and will only give results in the very long term.
A new, emerging way to alleviate congestion problems is to use the existing infrastructure more efficiently. Several approaches exist to improve the performance of freeways. These methods range from providing information to road users to completely automated highways. These developments are usually referred to as Dynamic Traffic Management (DTM)

> *"Dynamic Traffic Management employs technologies for real-time traffic monitoring, network wide management of traffic flows, and traffic-adaptive control to respond to changing traffic conditions while improving the efficiency, safety and travel conditions of a highway network".*

It is a mistake, however, to look at DTM only in terms of increased highway capacity. There are great benefits to be gained from user interaction with the system. People can now make travel decisions to maximize their individual benefits based on real-time information. In general DTM should improve travel conditions, increase safety, and reduce energy use and environmental impact.

Various instruments are available within traffic management systems and they can be divided into two basic types. *Road side based* instruments and *In-Vehicle* instruments. Road side based instruments are designed to manage and control traffic flow (ramp meters) and give road side information (variable message signs). In-Vehicle systems use electronic equipment in the car to provide information to the driver (navigation) or to control the behaviour of the car (speed and distance control). All these systems can work in a complementary way.

To introduce DTM in reality, it is necessary to distinguish three different levels of decision making: Firstly, decisions have to be made concerning *long term planning* (days to years) for DTM. Questions about the feasibility of DTM have to be answered, and the location and impact of its instruments have to be studied. Sometimes this has to be done in combination with traditional traffic engineering, such as the construction of new roads. Secondly, decisions have to be made concerning *medium term management* (a few hours) of DTM. When several DTM instruments are used, questions have to be answered about the siting and settings of ramp meters and changeable message signs. In the case of an accident, especially, advice is needed quickly about the impact of the accident and the measures to be taken. Rather than analyse and react to information, decisions have to be based on forecasts to take action prior to degradation of traffic flows. Thirdly, decisions have to be made concerning *short term control* (few minutes) of DTM. These decisions are usually made by computers without human interference. Examples are the control of a local ramp meter, or the alarm signal caused by a truck entering a tunnel that is too low for it.

Most of the proposed systems are very expensive to run and maintain, require new technologies, and have uncertain impact. To solve the planning, management and control problems involved with the introduction of DTM, and to test innovative strategies, all kinds of *traffic and transport models* are needed.

> *"The main advantages of using traffic models are, that they provide an environment where traffic control strategies can be tested and fine tuned without having to disturb traffic, they avoid risk of liability when problems in a strategy are detected only after implementation, and they save the capital required to acquire and install traffic control hardware so that strategies, which may or may not work, can be field tested [80]".*

Before a modelling approach is chosen, however, several aspects must be considered [61]:

- *The decision making context.* The results of a model need to support a number of different decisions. Decisions are required for DTM concerning quantification of benefits, which instruments to use, where to place them and how to manage and control them.
- *Level of accuracy required.* To support these decisions, different levels of accuracy are required. To gain insight into the possibilities and the impact of DTM for the long term, less accuracy is required than for a study into the effects of local ramp meters on traffic flow at a short road section.
- *The availability of suitable data.* A key factor when deciding on a modelling approach is the availability of suitable data. When insufficient data is available, decision makers are unlikely to have much confidence in the model.
- *The state of the art in modelling.* A model should be able to replicate the relevant aspects of the behaviour of a transport system. Irrelevant variables, or aspects too difficult to forecast, should be omitted from the model. Further, the results of the model should be tractable and reproducible.
- *Resources, requirements and skills.* The final aspect to consider, when choosing a modelling approach is availability of the required resources, such as money, time, computer hard-

ware and software, etc. In relation to computers, the required computational power and the amount of required and produced data must be considered. In relation to the operator, skills and experience to interpret the results are required to make the correct decisions.

Traditional traffic models hardly meet the requirements to justify their use to solve planning, design and management problems in relation to the introduction of DTM. Most models are not suitable for the problems caused by congestion in relation to networks, or are not capable of operating within a reasonable time. There is a need for much more advanced traffic models, especially within the emerging field of DTM. These models should be able to deal with congestion and the instruments DTM offers, including any combination of instruments. It should be possible to study the impact of these instruments on a network wide scale.

The different levels of decision making have direct consequences for the accuracy of the required data. Long term planning can be done with rather crude data, aggregated over longer periods of time. Much more detailed data is required, for medium term planning and management, with more actuality, accuracy and reliability. Short term control requires the most accurate and up-to-date data, with a high level of detail. Reliability of the system and the accuracy of the data is vital with no human control of the system.

With the increased computational power of modern computers and recent developments in electronics and telecommunications for the collection of the required data, research towards such models is now possible.

This dissertation focuses mainly on the theory and applications of a new traffic model to solve problems in relation to long term planning and medium term management of DTM. Therefore, the *main research issues* addressed in this dissertation are:

- The development of a traffic model to simulate traffic in a road network and to test DTM instruments for their ability to solve the congestion problem.
- The development of a traffic management system which can predict traffic flow in a freeway network including the effect of DTM instruments

It is preferable, for long term planning applications, if the new model fits neatly into the traditional four step traffic modelling approach. This makes it easier to integrate the new model into the existing traffic modelling methods. Fast availability of results is important, for medium term management of a network, because decisions have to be made promptly.

To answer the main research issues the following approach was chosen.

- The definition of DTM was studied. What differentiates DTM from traditional traffic management? What kinds of instruments are utilized by DTM and what are the benefits to be expected?
- DTM has initiated a lot of research effort towards the development of new models. These efforts were studied and categorized.
- A new traffic model was designed, implemented and evaluated. The development and evaluation of this model comprises the core of this research project.
- A framework to apply the model for long term planning applications and a framework to apply the model for medium term management applications is proposed and several modules of this system were developed further.

- The model was applied to the Washington Metropolitan Area, to test for long term planning purposes.
- Finally, to test the model for medium term management applications, the model with several additional modules were applied to the Amsterdam beltway in the Netherlands.

1.2 Outline of this Dissertation

This dissertation is organized in the order described by the research approach. The second chapter gives an introduction to Dynamic Traffic Management (DTM), the instruments that can be used and the benefits to be expected. A separate section is dedicated to projects initiated in Europe, the USA and Japan. The need for new advanced traffic models to solve planning, management and control problems in relation with DTM, is made clear by this chapter. Such models should be able to simulate traffic realistically and calculate effects of DTM instruments, not only on a network wide bases, but also as a combination of various instruments.

Chapter three reviews the relevant literature on traffic models. The models are divided into traffic assignment models and traffic simulation models. After an introduction, a brief overview of simulation models is given, and an overview of the traditional methods of traffic assignment. The developments towards models for DTM are mainly found in the category "dynamic traffic assignment models". It can be seen from this chapter, that several of these models lack certain basic requirements and that research towards a new or improved dynamic traffic assignment model is needed.

The fourth and the fifth chapters describe the main subject of this dissertation. Chapter four describes the underlying theory of a new dynamic traffic assignment model. All the different aspects of the model, such as pathfinding, assignment, and queuing, are described in detail. These aspects are evaluated, compared and calibrated in chapter five. The results are compared with another model, and the effects of discretization are investigated.

Chapter six explains the utilization of the model for long term planning applications, and medium term management applications. A framework for long term traffic planning and a framework for medium term traffic management is proposed. The structure of these frameworks is described, and several modules are explained in detail. Some modules are individual research projects of which only a brief description is given. The hardware and software developed is described briefly at the end of chapter six.

Chapters seven and eight describe two case studies. Chapter seven describes a case which was carried out during a four month visit to the Virginia Polytechnic Institute and State University, USA. This case describes an application of the model for long term planning purposes according to the traditional four step modelling approach. The primary objective of the study was to see whether the model can be used for planning purposes, whether there is an advantage over traditional models, and whether the model is a useful tool for investigating the effects of Dynamic Traffic Management (DTM). A secondary objective was to gain insight into the possibilities and problems associated with the application of the model to large networks.

Chapter eight describes the application of the model to the Amsterdam beltway. This case tests the framework for medium term management applications as proposed in chapter six. The primary objective of the case study was to see whether the model could give a reasonably accurate prediction of traffic situations, and how the model reacted to changing conditions, such as accidents and weather changes. A secondary objective of this case study, was to test if the model could provide predictions within a reasonable time.

The dissertation finishes with conclusions and recommendations for further research in chapter nine.

1. INTRODUCTION

CHAPTER 2

ROAD TRANSPORT INFORMATICS

2.1 Introduction

The objective of this chapter is to give some insight into the context of the problem addressed in this dissertation. First a brief overview of the IVHS program in the USA, the DRIVE program in Europe and research in Japan is given. These sections describe the worldwide interest in, and the magnitude of the investment in new techniques to solve the congestion problem. Sections 2.3 and 2.4 focuses on Dynamic Traffic Management and its instruments, and some results of Dynamic Traffic Management in the Netherlands are given. These sections will give some insight into the proposed solutions and the impacts. The benefits to be expected from DTM are discussed in section 2.5. Some remarks about these developments in general are given in section 2.6, and conclusions are presented in section 2.7.

2.2 International Programs

Worldwide there are three major programs. The DRIVE program in Europe, the IVHS program in the USA, and the VICS program in Japan. This section will describe how these three programs originated, how they are structured, and what kind of research is done. Information for this section was derived from Ian Catling's book on IVHS and ATT [20] and from several papers. Papers from Boesefeldt & Neuherz 1993 [9] and Klijnhout 1993 [42] were used for the European programs, and papers from Judycki & Euler 1993 [40], Euler 1990 [22], and Sussman 1992 [88] for the American programs. Information on the Japanese programs was derived from Upchurch 1993 [92].

2.2.1 Europe

In 1985 the EUREKA project, an open framework for cooperative projects between European firms and research institutes, was established by the European Commission. One of research programs within EUREKA was PROMETHEUS (PROgraM for European Traffic with Highest Efficiency and Unprecedented Safety). The program was initiated by the European car-manufacturers and was mainly focused on in-car electronics. It contained many projects, divided into four different functions.

These functions are:
- *Improved Driver Information*: Obstacle detection, Monitoring environment, Monitoring driver, Vision Enhancement.
- *Active Driver Support*: Safety margin determination, Critical course determination, Dynamic vehicle control, Supportive driver information.
- *Cooperative driving*: Intelligent manoeuvring and control, Intelligent cruise control, Intelligent intersection control, Medium range pre trip information, Emergency warnings.
- *Traffic/Fleet Management*: Static route guidance, Dynamic route guidance, Trip planning, Network control, Parking Management, Flow control, Demand management, Interface to public transport, Commercial fleet management.

A lot of effort was put into projects to fulfil these functions, and several projects have shown to be successful. To show the results of the research effort to the public (politicians in particular) *CEDs* (Common European Demonstrators) were introduced. CED's are vehicles equipped with two or more functions. The CED's were developed over a very short time span, and they certainly served their purpose; they received wide-spread press coverage and more important they produced a breakthrough in European policy.

In 1988 the DRIVE (Dedicated Road Infrastructure for Vehicle safety in Europe) program was founded by the EC (European Community). DRIVE is part of the Framework program which supports industrial development funded by the EC. The main objectives of DRIVE are to increase road safety, reduce the negative ecological impact of road traffic, and improve the efficiency of the infrastructure. This should lead towards an Integrated Road Transport Environment (IRTE). While PROMETHEUS was initiated by European Industry and emphasized the intelligent vehicle, DRIVE was initiated by the European authorities and focused on the intelligent infrastructure, the role of public authorities and on human behavioural aspects. Within the framework of DRIVE, many research projects were initiated and completed on mainly basic research, such as route guidance and information technology (e.g. SOCRATES). The results of DRIVE gave an idea of the technical possibilities of Road Transport Informatics and the consequences of the human-machine interaction. Some important steps towards standardization were also made.

The DRIVE II program was launched in January 1992, and is another three year program funded by the EC. In the past this program was often referred to as *Road Transport Informatics* (RTI), but more recently as *Advanced Transport Telematics* (ATT). While DRIVE I aimed at "exploring the options" the main target of DRIVE II is "preparing for implementation". The technologies developed earlier in the DRIVE I program are now being validated, by assessment and integration during field trials. DRIVE II started with 56 large scale projects, of which 30 were pilot projects and 26 were supporting Research and Development projects.

The following main themes were formulated for DRIVE II:
- Standardization of information and possible communication
- Pilots or real life tests to establish the acceptance of Road Transport Informatics by the public
- Evaluation of the effects of Road Transport Informatics systems in terms of safety, efficiency etc.
- Assessment of design safety.

Seven interrelated areas of major interest were defined to cover these main themes:

I. Demand Management
Demand management concerns area access control and pricing. The use of vehicles in towns and cities is discouraged by utilizing various parking restrictions, high parking prices and toll collection. Facilities such as park & ride and public transit are encouraged.

II. Traffic and Travel Information
This area comprises the collection, processing and distribution of information of direct use to people before and during the course of a journey. The information is not restricted to drivers of vehicles, but is also provided to public transport passengers and fleet operators. Various kinds of information, such as route and service information, pre trip and trip information, weather and road conditions and public transport information, should help to improve travel conditions

III. Integrated Urban Traffic Management
Traffic management in urban areas is mainly concerned with the management and control of urban traffic streams and urban traffic facilities, for examples coordinated traffic signal systems, route information, and parking management systems.

IV. Integrated Inter-Urban Traffic Management
This area covers systems for traffic management and control on inter urban highways and alternative parallel roads. It covers data collection, data processing, data diffusion and traffic control actions. Examples in this area are, incident detection, integrated network control, route control, traffic calming, and tidal flow.

V. Driver Assistance and Cooperative Driving
Systems to assist drivers with their primary driving tasks. Systems range from cruise control, through anti collision control, to intelligent manoeuvre control and cooperative intersection control.

VI. Freight and Fleet Management
This area should improve efficiency and safety for truck operators, with improved systems for scheduling and dynamic route planning, fleet monitoring and vehicle dispatching. Paper work should be minimized with automatic order processing.

VII. Public Transport Management
Travel conditions for public transport can be improved by providing publicly accessible databases with schedules, networks, fares, real-time information, and interchange information at terminals. Further improvement is possible by providing priority to public transport in the streets.
A transit service operator could improve its services by utilizing dynamic data and real time information about delays, to (re-)schedule and improve planning using travel demand forecasts.

Several pilot projects were started at a number of sites in Europe. The field trials of PROMETHEUS were integrated with DRIVE. Recently, an official framework between the DRIVE and PROMETHEUS programs has been established, called ERTICO (European Road Transport Telematics Implementation Coordination Organization).
At this time the first arrangements are being made for DRIVE III, which starts in 1995 and which will contain really large scale demonstrations.

2.2.2 USA

In 1986 an informal group of academics, federal and state transport officials, and representatives of the private sector began to discuss the future of the transport system in the United States. Various reasons could be given for the need of a new vision. Firstly the Interstate System, a 120 billion dollar project was about to be finished in 1991. Secondly congestion had grown dramatically over the past 20 to 30 years resulting in serious economical, safety and environmental problems. Thirdly Europe and Japan were advancing very quickly in the development of new technologies for transport systems, so national productivity and international competitiveness was at stake.
In 1990 the need for a formal organization became clear and IVHS America (Intelligent Vehicle/Highway Systems) was formed to act as a Federal Advisory Committee for the US Department of Transportation. In December 1991, the Intermodal Surface Transportation Efficiency Act (ISTEA) became law. It's purpose is to develop 'a national intermodal transport system that is economically sound, to provide the foundation for the nation to compete in the global economy, and to move people and goods in an energy-efficient manner'. IVHS is an integral part of ISTEA. In June 1992 IVHS America produced a 'Strategic Plan for Intelligent Vehicle Highway Systems in the US'.

The vision for IVHS was articulated as follows:
- a national system that operates consistently and efficiently across the U.S. to promote the safe, orderly and expeditious movement of people and goods.
- an efficient mass transit system that interacts smoothly with improved highway operations.
- a vigorous U.S. IVHS industry supplying both domestic and international needs.

Analogous to the seven areas of major interest defined in DRIVE, IVHS uses six predominant system areas:

I. Advanced Traffic Management Systems (ATMS)
ATMS is expected to integrate management of various roadway functions, predict traffic congestion and provide alternative routing instructions to vehicles and transit operators. Real time data will be collected, utilized and disseminated. Dynamic traffic control systems will respond in real time to changing conditions. Incident detection is seen as a critical function in this area.

II. Advanced Traveller Information Systems (ATIS)
ATIS involves providing data to travellers in their vehicle, in their home, or at their place at work. Information will include: location of incidents, weather problems, road conditions, optimal routing, lane restrictions, and in-vehicle signing. The information will be provided to both vehicle drivers and transit users.

III. Advanced Vehicle Control Systems (AVCS)
AVCS is viewed as an enhancement of the driver's control of the vehicle to make travel safer and more efficient. Systems range from cruise control and collision warning to completely automatically controlled vehicles in special lanes. Dramatic improvements in highway capacity are possible by assuming partial or total control of vehicles.

IV. Commercial Vehicle Operations (CVO)
Private operators of trucks, vans and taxis are adopting IVHS technologies to improve the productivity of their fleets and the efficiency of their operations. Faster dispatch-

ing, efficient routing, and more timely pick ups and deliveries are possible. Safety improvements are possible by tracking hazardous material transports, and monitoring driver fatigue.

v. Advanced Public Transport Systems (APTS)
APTS involves providing information to transit users concerning information on routes, schedules, delays and fares. The scheduling of public transport vehicles and the utilization of bus fleets can be improved by utilizing real time information. Another example is an automated fare collection system.

vi. Advanced Rural Transportation Systems (ARTS)
The special economic constraints of relatively low density roads and the question of how IVHS technologies can be applied in this environment is a challenge that has been undertaken by many rural states. The main focus is on highway safety. Other objectives are to provide in-vehicle route information, traffic and weather advisory information and the use of beacons to locate accidents or disabled vehicles in remote areas.

IVHS America has achieved a tremendous momentum and has spawned a number of projects whose scale will outstrip any currently being implemented in Europe. One of the key features of the program is the recognition of the need to establish a universal system architecture and a strategic plan for implementation, the communication links between vehicles, road side equipment and traffic centres are of critical importance in this universal system architecture.
The IVHS program, in comparison to the European DRIVE program, lacks an interest in Demand Management.

2.2.3 Japan

IVHS development in Japan can be traced back to the early 1970's. Japan has made substantial progress and is regarded the world-leader in several IVHS areas, due to its long term research into IVHS. One of the first programs in Japan was called the Comprehensive Automobile Traffic Control System (CACS), which started in the early 1970's. The CACS program was funded by the Ministry of International Trade and Industry.
A second program, the Road Automobile Communication System (RACS) was initiated in the mid 1980's. This program focused on in-vehicle navigation systems and was (partly) funded by the Ministry of Construction.
The Advanced Mobile Traffic Information and Communication System (AMTICS) program was initiated in the late 1980's. The program had many similarities with the RACS program and was funded by the National Police Agency. The AMTICS program included a large demonstration project in Osaka.
In 1991 a fourth program was started, the Vehicle Information and Communication System (VICS). This program is intended primarily to bring the RACS and AMTICS programs together. VICS is the current focus of Japanese efforts in IVHS.
Japan has already implemented several IVHS features. The country has 74 traffic control centres and 87 sub centres, virtually the entire urban roadway network is monitored via traffic control centres. Radio station traffic reporters have access to the traffic control centres to broadcast on site traffic information.
There are currently 12,000 signalled intersections in Tokyo, of which 7,000 send data to the traffic control centre. Additional information comes from 104 television camera's and from

ultrasonic detectors. Inductive loop detectors are rarely used [92].
At present ramp metering is not used for traffic management. This is because practically all the expressways are toll roads with the toll booths serving as traffic meters.
Travel time information is collected using video imaging. License plates are read at various places and travel time between two locations can thus be easily derived. Experiments have been done with vehicles equipped with special identification tags which are read by infra red devices. These devices also transmit travel times back to the vehicles.
Japan is also ahead in in-vehicle navigation and route guidance systems, some 250,000 to 300,000 of these systems have been sold in Japan.

2.3 Dynamic Traffic Management

The relevant areas for this dissertation are within DRIVE, the Integrated (Inter-) Urban Traffic Management area and within IVHS, the Advanced Traffic Management Systems.
These systems can be defined as follows:

> *"Dynamic Traffic Management employs technologies for real-time traffic monitoring, network wide management of traffic flows, and traffic-adaptive control to respond to changing traffic conditions while improving the efficiency, safety and travel conditions of a highway network".*

Three stages of development for ATMS are recognized in IVHS. The first stage, building blocks 1992-2000, requires basic research on driver behaviour and travel information needs, and the development of new traffic models (dynamic traffic assignment and simulation models). The development of new detection hardware, and the exploration of artificial intelligence techniques for incident detection and control strategy selection is also required. The second stage, advanced strategy development, 1995-2005, will use insights derived from the previous stage to develop advanced traffic management strategies, including short term forecasting techniques, rapid response incident detection techniques, demand management strategies, and integrated adaptive traffic control strategies. The third stage, ultimate traffic control, 2000-2005, envisages an interactive traffic control system in which traveller information systems (ATIS) are integrated into traffic management centres.

When a closer look is taken at these areas it is possible to distinguish three different kinds of systems. Firstly there is a need for systems to support *planning decisions*, secondly for systems that provide information to support *management decisions*, and thirdly various *control systems* are needed.
Adopting IVHS terminology these three systems can be called respectively: Advanced Traffic Planning Systems, Advanced Traffic Management Systems and Advanced Traffic Control Systems. A similar division is proposed by Westerman & Hamerslag [99].

Advanced Traffic Planning Systems (ATPS)
: ATPS deals with providing information about the possibilities of IVHS. Advanced models to support decisions about the construction of new roads and the *placement* of variable message signs and ramp meters are required to study the impact DTM have on the current traffic patterns.

Advanced Traffic Management Systems (ATMS)[1]
: Traffic Management improves the efficiency of a network. An ATMS gives advice, based on real time data, for taking measures to improve the throughput and/or increase safety in

the network. These systems provide traffic operators with information on which they can make decisions. Management systems require human interaction, and based on the available information, using sophisticated models, a traffic operator takes the decisions, not the system. One of the goals of ATMS is to prevent congestion, so rather than analyse and react to information, forecasts are required to take action prior to degradation of the traffic flows.

Advanced Traffic Control Systems (ATCS)

Traffic Control Systems improve the efficiency at certain (critical) parts of the network. These systems can operate without any human interaction. With these systems, it is the system that takes the decisions. The data requirements are much higher for these systems, and usually the impact of such is system is only local.

The main reason for making these divisions are the different levels of decision making. ATPS is concerned with *long term planning* (days to years), ATMS with *medium term management* (few hours) and ATCS with *short term control* (few minutes). The divisions have direct consequences for the accuracy of the required data. ATPS can do with rather crude data, aggregated over longer periods of time. Much more detailed data is required for ATMS, with more accuracy and reliability. ATCS requires the most accurate data, with a high level of detail. While ATMS still involve a person for decision making, no human interference is present with ATCS, so reliability of the system and the data plays a major role.

In the third stage of ATMS, described as the "ultimate traffic control stage"(2000-2005), ATMS and ATIS finally end up in one integrated system. By providing more and better information to (individual) travellers, ATIS will have to take the actions of ATMS into consideration. ATMS could use ATIS to improve the overall performance of the system, and the information available in ATIS from *probes* (vehicles, equipped with navigation devices and two-way communication) will be of great use for ATMS. A central system could take global optimization into consideration when providing information to individual travellers.

2.4 Instruments of Dynamic Traffic Management

Within traffic management systems various instruments are available. They can be divided into two basic types. *Road-side based systems* and *In-Vehicle systems*. Road side based systems are designed to manage and control traffic flow (ramp meters) and give road side information (variable message signs). In-Vehicle systems use electronic equipment in the car to provide information to the driver (navigation) or to control the behaviour of the car (speed and distance control). All these systems can work in a complementary way. The division is adopted from Lambrechtsen & Westerman 1993 [44].

2.4.1 Road-side Based Systems

Incident detection systems, ramp metering and variable message signs are examples of road-side based systems. Data for these systems is usually collected by detectors. A large variety of detectors is available, such as mechanical, optical, magnetic, acoustical, radar and video camera's.

1. Caution: ATMS is redefined here, and narrowed down to the actual management of traffic flow. Automatic traffic control systems are excluded.

Incident Detection. An important cause of congestion are incidents, thus automatic incident detection and congestion monitoring are important functions within dynamic traffic management. The function of incident management is to detect incidents and warn upcoming traffic. More comprehensive systems include route advice and warning the police and other services. A large part of incident management is concerned with organizing the support vehicles and removing the cause of delay.

In the Netherlands the Motorway Control and Signalling System (MCSS) has been in operation since 1983. A sign is placed above each lane which can display a number (speed restriction), an arrow (merge to next lane), a red cross (lane closed) and an end of messages sign. A sign can be supported by flashing lights. The system is fed with information from double induction loop detectors. Speeds and flows are collected at a road-side computer. Loops and signs are spaced 500m to 1000m from each other. In the case of congestion the system automatically generates speed restrictions to warn up-stream traffic, for accidents or road-works a lane can be closed using the arrow and the cross signs. In this case the signs are controlled from the nearest traffic control centre.

The MCSS system is very successful in the Netherlands, evaluation studies, have shown that the number of accidents has decreased by 23.6 percent and the number of secondary accidents by 46.2 percent (Van Willigen 1992 [102]).

The system is also being used for experiments with fog detection and homogenising traffic flows. In the case of fog detection the system is fed by special fog detectors. Experiments with recommended speeds have been done to homogenize traffic flow. See Heidemij 1993 [34].

Ramp-metering. Metering traffic at on ramps was applied for the first time in the USA in 1968. The system is meant to relieve the traffic flow from too many merging vehicles at once. This is achieved by placing a traffic signal at the on ramp which allows one vehicle per cycle, with a high cycle rate. Different approaches exist to control the traffic signal. Most systems use induction loops in the main stream to control the traffic light frequency. Some use detection loops down stream of the on ramp to control the performance of the section (USA), some use upstream loops to detect 'holes' in the stream to allow cars to merge (Europe). The latter methods sometimes allows more than one car to merge. Different studies have been done to evaluate the methods. See Middelham 1991 [59], Salem et al. 1988 [79] and Owens 1988 [62].

Four ramp metering installations are in operation in the Netherlands, in contrast to Los Angeles where almost a thousand ramp meters are installed.

Ramp-metering installations are usually *local* systems. It is, however, possible to utilize several ramp meters for *coordinated ramp-metering*. The in-flow of traffic for an entire freeway could be controlled with ramp meters at every on-ramp. More elaborate systems are proposed that use special *buffers* at on ramps and the *metering of the freeway* at large freeway intersections. At this time no experience is known about such applications.

Variable Message Signs are signs that give information about current or future traffic conditions. In principle the nature of the information given can be divided into two: descriptive information and prescriptive information (van Berkum & van der Mede 1993 [7]). Descriptive information provides drivers with information about the conditions in the network (e.g. Potomac bridge closed, or congestion on I-66 North). Prescriptive information tells drivers what to do (Arlington: take I-395 West). Of course descriptive and prescriptive information may be mixed. In Japan large graphic signs are used to display levels of congestion on the downstream roadway of the network pictorially.

In the Netherlands variable message signs have been in operation since 1991. The signs are descriptive and give information on traffic jam length. Several evaluations have been done on the effects of the signs, and in general the results are positive. See Witteveen & Bos 1992 [103] and BGC 1992 [15]. Van Berkum & van der Mede did an extensive study on the impact of traffic information which resulted in a dissertation [7].

2.4.2 In-Vehicle Systems

The main interest for in-vehicle systems is focused on electronic navigation systems and information systems, they can, however, also be applied to traffic management. These navigation systems can be divided into three different levels. These levels are based on increased communication capability. The most simple navigation system is based on static information, a more advanced system is based on one-way communication, while the most advanced system is based on two-way communication. Other in-vehicle systems are systems that provide only traffic information like the Radio Data System (RDS).

Electronic Navigation. Nowadays a number of different navigation systems are on the market, and most of these systems use an LCD display and/or a synthetic voice to give route information. The system navigates the driver to a given destination. In the first systems the navigation system was used as a map reader, and told the driver where to turn left and right. The directions were based on the calculated shortest route without taking possible congestion into account. Later systems include a communication device to inform the system about possible delays. This information can be used for route choice decisions and en-route diversions. Systems with two-way communication represent a significant enhancement in terms of capabilities. Individual vehicles can become traffic *probes* and supplement the information provided by traffic sensors and other detection systems by sending data to the traffic control centre. These traffic probes become a valuable source of information for various systems.
There are several competing systems in Europe. Philips developed a system called CARIN (CAR Information and Navigation system), and Bosch/Blaupunkt developed a system called Travelpilot, both these systems work with wheel-sensors and an electronic compass to maintain the position on the map. Map matching and dead reckoning is used for accuracy (Thoone, 1987 [89]). A CD-ROM is used for static map information. The third system is ALI-SCOUT/ LISB which works with beacons to determine the location of the car on a map. There is no map in the car and the beacons provide the necessary map information in the vicinity of the car. Shortest route calculations are done at a central traffic centre.
In Japan, Toyota has developed their GPS Voice Navigation System. In 1992 they produced 1500 cars per month equipped with this system (additional cost: $6,000.-). The system works with a touch-screen mounted in the dashboard that displays maps and sketches of upcoming intersections. The driver gets instructions from a synthetic voice. The position of the car on the map is maintained using two Global Positioning Systems, map matching and dead reckoning. The maps are stored on CD-ROM.

Radio Data System (RDS). A normal FM radio signal has the possibility to broadcast extra digital information without disturbing the normal radio broadcast. This extra information becomes available with a special RDS-decoder in the car radio. The extra digital information can be used for various functions, such as automatic frequency switching, disaster reports, switching from cassette to radio for traffic messages, etc. The Traffic Message Channel

(TMC), is a special usage of RDS and the most relevant for dynamic traffic management. Digitally coded traffic information is transmitted and stored in the memory of the car radio. The decoded information is available to the driver at any time. The navigation systems can also use this service to update travel times. Car radio's equipped with RDS have been available for several years now. A pilot project using TMC is taking place in Germany and The Netherlands.

2.5 Benefits of Dynamic Traffic Management

Several DTM systems have been discussed in the former sections. All these systems provide certain benefits. To measure the impact of these systems some standard of measurement for evaluating the current situation must be established. One approach is to establish measures of effectiveness. Roughly six categories of benefits can be determined and for each benefit some measures of effectiveness can be listed. It is possible to express the benefits in real numbers using these measures of effectiveness [52].

Benefit 1: Improved capacity and operational efficiency
- Throughput, travel speeds, travel times, delay times, incident detection time, number of stops, vehicle occupancy, predictability of travel time

Benefit 2: Improved safety
- Number of accidents, number of fatalities, number of secondary accidents, accident costs, incident response times, driver fatigue

Benefit 3: Reduced use of energy and environmental impact
- Fuel efficiency, fuel consumption, vehicle emission, noise pollution

Benefit 4: Increased productivity for motor carriers and service providers (taxis, couriers)
- Operating costs, volume of goods moved over existing facilities, just-in-time delivery

Benefit 5: Increased comfort and convenience of travel
- Motorist stress, motorist confusion, driver fatigue, predictability of travel time

Benefit 6: Improved cooperation between transport system operators.
- Incident & congestion information, sharing of information between agencies, consultation on control strategy implementation

Several studies have been done to express these benefits in real numbers. Different studies come up with different numbers. Reduction in congestion costs ranges from 0 to 40 percent in different studies. It is very difficult to claim certain percentages of benefits and field trials are needed to gain some insight into what the real benefits are. In the USA, for the next 20 years a total of 230 billion dollars will be spent on IVHS with 80% from the private sector.

2.6 Feasibility of IVHS/ATT

Is IVHS/ATT a buzzword that will quickly disappear, or will it lead to a substantial improvement in our transport systems? The real cause of the problem of congestion lies in land use policies that encourage car dependency. It might be stated that IVHS/ATT systems to alleviate congestion is a 'band-aid' approach. Travel demand is not a static commodity, it responds dynamically to the capacity of the supply. Increased capacity tends to relieve congestion in the short term but soon stimulates additional travel demand. Numerous examples of this phenomenon are known. As Carl Thor [90] puts it nicely:

> *"When the limits of what IVHS can do for traffic capacity are reached, we will still be in the same boat but further up the creek".*

Although IVHS/ATT might be beneficial for alleviating some congestion and may increase safety and improve travel conditions, it should not be viewed as a cure-all. At least as much effort should be put into alternative solutions, which will improve the long term situation. IVHS/ATT should be accompanied by measures to reduce automobile use by changing land use patterns and encouraging people to travel by foot, bicycle, public transit and high occupancy vehicles.

2.7 Conclusions

Although there is some legitimate scepticism on the feasibility of IVHS/ATT, it has become a fact. In Europe, USA and Japan the governments are investing huge sums of money in these technologies, and the first pilot projects are becoming a reality. Although the benefits of increased capacity might be debatable, improvements in safety and travel conditions are feasible.

Within DTM, three different levels of decision making can be discerned: Advanced Traffic Planning Systems (ATPS), Advanced Traffic Management Systems (ATMS) and Advanced Traffic Control Systems (ATCS). ATPS looks at long term planning (days to years), ATMS deals with medium term planning (few hours) and ATCS with short term decisions (few minutes). These divisions have direct consequences for the accuracy of the required data. ATPS can do with rather crude data, aggregated over longer periods of time. Much more detailed data is required for ATMS, with more accuracy and reliability. ATCS requires the most accurate data, with a low level of aggregation. While ATMS still involves a person for decision making, with ATCS no human interference is present, so reliability of the system and the data plays a major role.

Within the vision of ATPS, ATMS and ATCS, the need for new advanced traffic models is clear. To solve planning, management and control problems there is a need for more advanced traffic models. These models should be able to simulate traffic realistically and be suitable for calculating the effects of dynamic traffic management measures. Fast availability of the results is also very important for ATMS applications.

CHAPTER 3

LITERATURE REVIEW

3.1 Introduction

The objective of this chapter is to review the current state of traffic flow prediction models, with emphasis on dynamic assignment models. The chapter will make clear what type of model has the best potential to operate within DTM.
Several classifications and reviews of dynamic assignment models can be found in the recent literature. Some are made with the intended application field in mind, e.g. the ability to model the impact of traffic information (van Berkum & van der Mede [7]) or the type of road, e.g. signalled intersection versus freeways (A. May [53]), some are based on to what extend a model is able to represent a number of basic types of traffic operations, e.g. supply, demand, route choice, departure time, classes, etc. (D. Watling [98]) and others are based on the techniques used (e.g. van Grol [29], S. Peeta [69] and I.A. Kaysi [41]).

In this review, the traffic flow prediction models are divided in two categories: simulation models and assignment models. Neither of these models have a generally accepted definition. In this review, simulation models are time propagation based models which do *not* consider route choice, and assignment models are not necessarily time propagation based models that *do* consider route-choice.

A brief description of simulation models is given in section 3.2. The remainder of this chapter focuses on traffic assignment models, with an emphasis on the dynamic variant. The variables used throughout this dissertation are listed in appendix E.

3.2 Simulation Models

Simulation is a numerical technique for conducting experiments on a computer, which may include stochastic characteristics, and involve mathematical models that describe the behaviour of a transport system over extended periods of time. Simulation models are typically time propagation models, which calculate the state of the model at time $t+1$ based on the state at time t. In this review, simulation models do not consider route choice.
There are two basic concepts in traffic flow simulation: *Microscopic simulation* models, based on individual driver behaviour and the interaction between vehicles, and *Macroscopic simulation* models based on hydrodynamic theory.

3.2.1 Microscopic Simulation Models

A microscopic simulation model describes the interaction between individual drivers based on the drivers behaviour. These models are also described as car-following models. Some models are based on the mechanical properties of the cars, others on the perception of the driver, the more advanced models are based on both.
The advantage of these models is, that complete information about the state of each vehicle is available. A typical application of a microscopic simulation is to investigate local behaviour of traffic. Usually short road sections and short time periods are considered. A description of various microscopic simulation models is out of the scope of this review.

3.2.2 Macroscopic Simulation Models

One of the first macroscopic simulation models is the simple continuum model developed by Lighthill and Witham [46] in 1955. Since then much progress has been made in the development of macroscopic continuum models.
A distinction can be made between *simple* continuum models and *high-order* continuum models. Continuum models are based on the conservation equation. The high-order variant also takes into account the dynamic effects of inertia and acceleration of traffic mass, by adding a momentum-equation.
A nice overview is given by A. Lyrintzis et al. 1994 [48]. They distinguish the simple model Lighthill & Witham 1955 [46], the original high-order model, Payne 1971 [67],1977 [68], the improved high-order model, Papageorgiou 1983 [64],1989 [65], the semi-viscous model, Michalopoulos et al. 1991 [57], 1993 [58], and the viscous model, Michalopoulos et al. 1991 [56].
Continuum models are usually used to study local traffic behaviour. Normally small networks with only one route between OD-pairs and short time periods are used. Typically a link length of 30 meters and a time period of 10 seconds are used.

3.3 Assignment Models

A transport system basically consists of two elements: *transport supply* and *travel demand*:

- Transport supply is the set of facilities and means available to the users of the traffic network. The supply is usually identified with the road network, represented by a directed graph where arcs correspond to network links, and nodes to junctions. With each link a travel cost is defined. If the travel cost is a function of the link flow, a travel cost function is introduced.
- Travel demand is expressed by the number of users using the network for a given reason (work or leisure), using a certain transport mode (car, train), at a given time of day. The demand is usually identified by an Origin-Destination matrix (OD-matrix), specifying the trips from an origin to a destination.

The interaction between transport supply and travel demand produces a flow pattern on the network links. A representation of transport supply and a model of travel demand are necessary to compute the network flow pattern and a model has to be build to simulate the interaction between the two elements. This model is known as a *traffic assignment model* [84].

Traffic assignment models have been used in traffic engineering for several decades. It usually is the last step in the classic four-step approach. The results of traffic assignments are used for various reasons. The basic application is to get insight into the state of the network for planning and design purposes. Planning scenarios can be evaluated by changing road-attributes and adding (or deleting) road-segments. In the last few years the traffic assignment model has also been used to provide the basis for environmental studies.

It is possible to classify traffic assignment models on different levels:

I. based on theoretical assumptions concerning *demand* and *supply*,
   on the transport *supply* side assumptions are made concerning:
   - a fixed, limited capacity of a road segment (queuing), or a flow dependent capacity.
   - the conditions along the length of the link. (homogeneous or divided into a flowing part and a queuing part).
   - a traffic flow dependent travel time.
   - all kinds of variability (static or dynamic, deterministic or stochastic) in service characteristics (capacity, free flow speed, traffic lights, etc.).

   On the travel *demand* side (assuming a known OD-matrix) assumptions are made concerning:
   - dependency between the level of demand and the quality of service (elastic demand, dynamic departure times).
   - homogeneous behaviour of travellers (e.g. concerning route choice).
   - all kinds of variability (static or dynamic, deterministic or stochastic) in demand characteristics (travel cost, perception, value of time, etc.).

II. based on assumptions made about the interaction between supply and demand; usually expressed as *the users behaviour*, or *the model*. The two most studied models are: the user equilibrium model and the system optimal model.

III. based on the solution techniques to solve the model. Several different techniques are known for solving a model, e.g. analytical methods, iterative methods, methods using path-enumeration, etc.

One of the most fundamental theoretical assumptions is the treatment of the time aspect. A division is possible into: static assignment models and dynamic assignment models. A *static* traffic assignment model treats time as a constant, travel demand and costs are time invariant. In contrast, a *dynamic* traffic assignment model treats time as a variable, travel demand and cost now vary over time.
Section 3.3.1 will describe static assignment models, and section 3.3.2 dynamic assignment models.

3.3.1 Static Assignment Models

Static assignment models assume that travel demand and link cost functions remain constant during the time span of interest. As a consequence, link flows and link trip times are time independent.
Taking the possible theoretical assumptions concerning demand and supply into account, several different models are possible, for example the all-or-nothing assignment is a model in which all travellers between an OD-pair take the same shortest path. It assumes a deterministic demand, deterministic and flow independent supply, has one class and a deterministic route choice. The user-equilibrium method is a model in which travellers between an OD-pair can take different routes. It assumes a deterministic demand, deterministic and flow dependent

supply, has one class and a deterministic route choice.

The *user-equilibrium* model is probably the most studied method. At equilibrium the distribution of travellers from one OD-pair is defined by the first criterion of Wardrop [97]:

"The journey times on all routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route."

The mathematical program, developed by Beckman in 1956 [4], allows the derivation of the existence of the Wardrop user equilibrium criterion, and the uniqueness properties of the solution.

Beckman's mathematical program:

$$Min\, z(x) = \sum_{ij} \int_0^{x_{ij}} t_{ij}(\omega)\, d(\omega) \qquad \text{(eq 1)}$$

The objective function, $z(x)$ is the sum of the integrals of the link performance functions $t_{ij}(x)$ for all links ij. The flow on link ij is denoted with x_{ij}.

The function is a mathematical construct that can be used to solve equilibrium problems [84]. The objective function is subject to several constraints. The first constraint is:

$$\sum_k f_k^{rs} = q_{rs} \qquad \text{(eq 2)}$$

This equation represents the flow conservation constraints. The flow f_k^{rs} on all paths k, connecting each OD-pair rs, should be equal to the OD trip rate, q_{rs}.
The second constraint is:

$$f_k^{rs} \geq 0 \qquad \text{(eq 3)}$$

This equation ensures that the flows are not negative.

The program is subject to the following **definitional** constraint:

$$x_{ij} = \sum_{rs} \sum_k f_k^{rs} \cdot \delta_{ij,k}^{rs} \qquad \text{(eq 4)}$$

This definitional constraint defines the incidence relationship. It expresses link flows in terms of path flows. The parameter $\delta_{ij,k}^{rs}$ is an indicator variable, equal to one if link ij is on path k between OD-pair r-s, and zero otherwise. This incidence relationship (e.g., the values δ) is presumed given. When $t_{ij}(x)$ is strictly convex and increasing, it can be shown that the objective function $z(x)$ is strictly convex, and that the constraints are convex. So there exists a unique solution to the above program.
To prove that the above program is indeed equal to the user equilibrium, it is sufficient to derive the first order conditions of the Lagrangian of the equivalent minimization program. (see 4.8.1)
A second well-known model, that can be solved with a mathematical program, is the *system-*

optimum model. The mathematical program for the system-optimum is similar to the program for user-equilibrium. The difference lies in the definition of the objective function, which usually represents the total system travel costs.

$$Min\, z(x) = \sum_{ij} x_{ij} \cdot t_{ij}(x_{ij}) \qquad \text{(eq 5)}$$

Sheffi [84] gives a comprehensive treatment of the static user-equilibrium assignment and system optimum, addressing the conceptual, mathematical, algorithmic and computational aspects.

Static assignment models are a well established and widely accepted class of models. Despite their widespread use they show a number of shortcomings. These shortcomings were recognized a long time ago, but the models were considered to be a sufficient approximation for the studies they were used for. Most of these shortcomings are related to the fact that the network conditions, such as flows and travel times, are assumed constant in time. The shortcomings became more apparent when congestion started to play a more important role in traffic studies (Ben-Akiva 1985 [5]). In a static assignment model trips are assigned to specific routes for the entire time period. This means that when a peak period is modelled, the same cars are assumed present at the same time at two consecutive bottlenecks (Hamerslag 1987 [31]). In reality they would have been delayed at the first bottleneck and thus could not contribute to the second. Apart from incorrect modelling of congestion, the usual response to congestion, i.e. to change departure time, cannot be modelled by static assignment models.
With the increasing problems of congestion the need for better (dynamic) models became clear.

3.3.2 Dynamic Assignment Models

Dynamic assignment models have developed at a rapid pace. Several approaches have been proposed and many others are being developed. At this point (1994) no single methodology has been generally accepted. The dynamics in the models are introduced by using discrete time periods, or continuous time functions. The existence of time variance by itself, however, does not imply that the model is dynamic, unless events in one period can influence events in the next.
Dynamics can be introduced into the assignment model at several levels:

- **Link-level**: The most elementary level of dynamics is the "temporal evolution" of the traffic flow pattern on a network. A vehicle should be assigned to a link in the network, at the time the vehicle actually traverses that link. This means that events occurring at a certain time should influence events occurring at a later time.
- **Route-level**: The second level of dynamics concerns route choice. Optimal routes vary with time and are either based on conditions valid during the trip, *experienced route choice,* or on conditions valid at the departure time of the trip, *instantaneous route choice*. Several variations (heuristic/stochastic) are possible at this level, e.g. by recomputing the instantaneous shortest paths at each decision point (junction), or incorporating day-to-day dynamics where route choice decisions are based on the experience of the previous day(s) and/or exogenous information.
- **Demand-level**: A third level of dynamics, concerns the dependency between the level of congestion and demand (e.g. dynamic departure time choice or dynamic mode choice).

When the level of congestion of the previous day(s) is taken into account, another form of day-to-day dynamics is attained.

In this review, a traffic assignment model is considered to be *truly* dynamic if at least the first two levels of dynamics are present in the model, i.e. an event in one period influences an event in a later period, and route choice varies over time.

Taking the different levels of dynamics into account and the various assumptions about demand and supply, a wide variety of dynamic traffic assignment models are possible.
A classification of dynamic assignment models based on these assumptions seems logical. Most models, however, can be extended or changed, to facilitate different assumptions. As can be observed from the literature, most models have been extended to deal with dynamic departure times and with exogenous information.
Still, there are basic differences between the models at the first two levels of dynamics, the evolution of traffic and the route choice. For this reason, it seems better to focus on the techniques used to introduce the dynamics in the model.

The field of dynamic assignment models can be divided into two large groups. The first group can be classified as the *heuristic* approach, and the second as the *mathematical* approach. This classification does not mean that heuristic approaches contain little mathematics or that mathematical approaches are not pragmatic.

3.3.2.1 Heuristic Approach

Several models have been developed following a pragmatic or heuristic approach. The main interest of these approaches is the ability to model traffic flow in a realistic way, without bothering too much about convergence of algorithms or uniqueness of solutions. Several models in this group have been implemented and are available as software products.
Some models can be seen as *extended static models*. These models still use many of the techniques known for static assignment models, but are truly dynamic. A second group are *packet-based* models, these models handle the traffic in groups called packets. A third group of models uses simulation techniques to move traffic, and are described as *assignment-simulation* models. All models that do not fall into one of these groups are placed in the group: *remaining models*.

Extended Static Models

The model closest to static assignment is SATURN (Simulation and Assignment of Traffic to Urban Road Networks) (Hall et al. 1980 [30], Van Vliet 1982 [95]), which is roughly a series of almost independent static assignments. The total time-span is divided into periods of a certain length. The network and the OD-matrix are divided accordingly. The traffic is assigned to each period independently, in a similar fashion to static assignments models. The only interaction between periods occurs when the flow on a link exceeds capacity. When this happens, excess traffic is assigned to the same link in the following period. SATURN was developed to find out how traffic flows will change in response to a traffic management scheme, especially the influence of changing timings of traffic signals in a network. The model consists basically of two modules, an assignment module, and a simulation module. The two modules are applied in an iterative scheme, in which the assignment module provides the flows for the simulation module, and the simulation module the delays (speed-flow curves) for the assignment module. Each module, however, handles traffic within one time-period, largely independently of other time-periods. Therefore, the model is not considered to be truly dynamic and is sometimes called quasi dynamic.

In 1987 Hamerslag proposed a model which had more interaction between the periods, called: "dynamic assignment in the three dimensional timespace" [31]. The model determines the time-varying traffic conditions, flow and travel time, for every link during a certain simulation period. The simulation period is divided into periods of equal duration, and the properties of the network and travel demand are presumed given. In the model the links are defined by the nodes i and j and period p. The travel time on a link ij during a period p is denoted as t_{ij}^{p}, the traffic flow during period p as x_{ij}^{p}, and the trips departing from node r to node s in period d as q_{rs}^{d}.

The model determines an equilibrium by minimizing:

$$Min \sum_{p}\sum_{ij}\int_{0}^{x_{ij}^{p}} t_{ij}^{p}(x)\,dx \qquad \text{(eq 6)}$$

The link flows are equal to the sum of all OD-pairs traversing the link:

$$x_{ij}^{p} = \sum_{d}\sum_{rs}\sum_{k} q_{rsk}^{d}\delta_{rsk}^{dp} \qquad \text{for } d \leq p \qquad \text{(eq 7)}$$

where δ_{rsk}^{dp} is an indicator variable, equal to one if in period p, link ij is on route k between OD-pair r-s, departing in period d, and zero otherwise. The sum of the trips on all the routes used between r and s must equal q_{rs}^{d} :

$$q_{rs}^{d} = \sum_{k} q_{rsk}^{d} \qquad \text{(eq 8)}$$

where q_{rsk}^{d} denotes the number of trips of the OD-pair r-s departing in period d, taking route k.

The model is solved using an iterative process. In each iteration the shortest paths in the network are calculated for all OD-pairs and for every departure period. This basic iteration scheme is essentially the same as for *static* assignment models. An extra loop over the departure period is required, and shortest pathfinding and assignment have to be performed *in time*. The paths are defined using the travel time on a link in the period in which the traffic actually traverses the link, i.e. the trajectory the traffic follows in time is calculated. The network is loaded, based on these trajectories.
The link delay t_{ij}^{p} is defined in a similar way as in static assignment models. The delay function t is a function of x_{ij}^{p}, Q_{ij}^{p} , C_{ij}^{p} and $t^{\circ}{}_{ij}^{p}$. Where Q_{ij}^{p} denotes the number of waiting vehicles from previous periods, C_{ij}^{p} the period capacity, and $t^{\circ}{}_{ij}^{p}$ the initial (unloaded) travel time on link ij during period p. The difference with the delay function in static assignment models is that the overloaded links in previous periods (Q_{ij}^{p}) influence the delay in later periods.

Packet Based

One of the first dynamic assignment models used for several studies is the model developed by the Transport and Road Research Laboratory in the U.K. (TRRL) and is known as CONTRAM (Continuous TRaffic Assignment Model) (Leonard et al. 1978 [45]). The model handles the traffic in groups called *packets*. Each packet (typically one to ten vehicles) is treated as a single entity and is assumed to experience the same traffic conditions. Each packet is loaded on the network by a kind of incremental assignment. The progression of packets through the network is calculated according to statistical and queuing theory relationships between flow increments, capacity at junctions and journey time. The conditions a packet will encounter on its route depend on the order in which it is assigned to the network, because after loading each packet, conditions are updated. In an iterative process packets are removed and re-assigned to the network in consecutive iterations, with the hope of converging to a dynamic equilibrium. CONTRAM has withstood time because of several extensions and improvements. The latest improvements involve changes to handle ATIS and ATMS.

Cascetta and Cantarella 1991 [18][19] have proposed a day-to-day and within-day dynamic traffic assignment model. The day-to-day dynamics capture the variations occurring between successive days, (or successive morning peak hours). The model follows a non-equilibrium approach, in which day-to-day fluctuations are modelled as a stochastic process. Route and departure time decisions are adjusted from one day to the next.
In this model time is modelled by discrete time periods and all travellers who leave from the same origin at the same time and follow the same route are grouped together. A group may be split into smaller groups if necessary. It is assumed that all members of a group experience the same travel conditions, and the length of a group is ignored. The network is defined using nodes and links. Two types of links are defined: running links and queuing links. The conditions on a link are considered to be homogeneous during a time interval.
Routes are based on a nested logit model. The utility function includes terms representing perceived travel times by route and departure time as well as penalties for early and late arrivals, relative to a target arrival time at the destination. During each time interval groups are moved along running links or spend time in queuing links, or a combination of the two. The location of each group is stored in memory.
As the travel time along a link depends on the flow on that link, and vice-versa, a fixed point problem has to be solved.

Another approach has been suggested by Ben-Akiva [6] and extended by Vythoulkas [96]. This equilibrium based approach uses discrete time periods and moves groups of vehicles along the links of the network, based on OD-flow rates, which are defined by path and departure time. The conditions on a link are homogeneous and the time-varying link inflows, outflows, and densities are based on link flow conservation equations. As the paths are based on time-dependent link travel times a fixed point problem has to be solved.

Assignment-Simulation Models

This group of models adopts car following simulation techniques. These approaches relate speed and density on a section to macroscopic relations, but move vehicles individually or in groups. The approach is situated somewhere between microscopic and macroscopic models and is sometimes called mesoscopic. The mesoscopic approach has the ability to keep track of individual vehicles, which offers an advantage when modelling traffic lights and the effect of in-vehicle information systems. SATURN, which has been placed under "extended static models", uses a similar approach.

Two models using this approach are DYNASMART (Mahmassani et al.[51]) and INTEGRATION (VanAerde et al. [93]). These models are designed in the context of in-vehicle electronic route guidance systems. The DYNASMART model simulates individual vehicles, moving them at a speed determined by the total flow on the link. DYNASMART uses a modified Greenshields' speed-density function to determine the speed on a link. INTEGRATION estimates the speed of a vehicle based on the headway in front of each vehicle. A speed-flow function is used to estimate the speed. Various speed-flow functions are possible in INTEGRATION. The function is determined using four different user-defined parameters (free-flow speed, speed at capacity, capacity and jam density). Routes are based on current conditions, and are therefore instantaneous. Routes have to be recalculated frequently, e.g. every 15 seconds, to react to prevailing conditions. Intersection delays (traffic lights) are modelled explicitly, and several different classes of vehicles are possible. INTEGRATION identifies four different classes of vehicles:

- Vehicle type 1 (background traffic): can have its routes selected by the user, or can be generated internally using the static Frank-Wolfe algorithm at user specified intervals.
- Vehicle type 2 (guided vehicles): uses instantaneous information to generate vehicle routes at user specified intervals.
- Vehicle type 3 (anticipatory routing): uses a combination of historical and real-time information in generating the routes. These vehicles attempt to minimize their respective travel times based on a weighted link travel time relationship. Vehicles that are expected to arrive at a link later in time place more emphasis on the forecasted future link travel time than on the current best estimate of link travel time.
- Vehicle 4 (TravTek vehicle): uses TravTek route guidance system logic to model vehicles.

Remaining Models

Janson 1990 [38] presents a dynamic assignment algorithm which consists of a once through, incremental assignment. The method is time propagation based. It starts by calculating the shortest paths for the traffic departing in the first period. The shortest paths are based on the travel times in the periods in which they actually traverse a link. Since the future link impedances are unknown at the time of calculation, the paths are based on *projected* link impedances. The projected link impedances are obtained by multiplying current link volumes by factors accounting for changing levels of demand in future periods. According to the resulting shortest paths, the OD matrix is assigned to these paths. The algorithm repeats this for all departure periods, assigning only a fraction of the OD-matrix. After assigning all OD-relations for all periods, the process is repeated, assigning a next fraction to the network. Different routes may now be found. This incremental assignment is repeated until the entire OD-matrix is assigned.

3.3.2.2 Mathematical Approach

Several models have been developed using a mathematical approach. The main interest of these models is to give a sound mathematical formulation of the assignment problem. Proof of convergence and the existence of a unique solution are the main interest. The literature contains very limited implementation or computational work, and testing has been done on very

small networks.
Models which belong to this group either seek a system optimum or a user equilibrium. The definition of a dynamic user equilibrium can be obtained by directly extending Wardrop's user equilibrium criterion. At equilibrium for travellers from one OD-pair departing at a certain time the criterion becomes:

> *"The journey times experienced on all routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route."*

In this definition route choice should be based on *experienced* travel times and *not* on instantaneous travel times. The definition can be extended by adding the choice of departure time to it.

The field is divided into models using a *discrete time mathematical programming approach* and models based on a *continuous time optimal control theory.* The main difference between these models is the treatment of time. Discrete time periods versus continuous time functions.

Mathematical Programming

One of the first attempts in the field of discrete time dynamic assignment was the formulation of the dynamic system-optimal problem by Merchant and Nemhauser [54][55]. They originally presented the model for a many-to-one network. Later Carey [17] reformulated the problem for multiple destinations as a convex, non-linear problem with non-linear constraints for each link. He also proposed extensions to handle multiple destinations.

The system optimal approach minimizes total system costs by minimizing the summation over all links a, and all time periods p, of the cost function $t_{ij}(x)$ incurred when link ij contains x_{ij}^p vehicles in time period p.

$$\text{Min } z(x) = \sum_p \sum_{ij} t_{ij}(x_{ij}^p) \qquad \text{(eq 9)}$$

The objective function is subject to a relation insuring that the number of vehicles exiting a node n to node c, denoted by m_{nc}^p, is equal to the number of vehicles entering node n from node b, μ_{bn}^p, within one period p, corrected by the amount of traffic generated, I_n^p, and attracted at node n, O_n^p, in period p.

$$\sum_b \mu_{bn}^p = \sum_c m_{nc}^p + I_n^p - O_n^p \qquad \text{(eq 10)}$$

The second important relation for this program states that the amount of traffic present at link ij in period p, x_{ij}^p, is the result of the amount present at link ij in period p-1, x_{ij}^{p-1}, corrected with the amount of traffic entering the link in period p, μ_{ij}^p, and the amount of traffic exiting the link, m_{ij}^p.

$$x_{ij}^{p-1} = x_{ij}^p - m_{ij}^p + \mu_{ij}^p \qquad \text{(eq 11)}$$

Further, the number of cars allowed to exit a link is subject to a function of the amount currently present at the link, and all the parameters in the program have to be positive.
Most of the work in finding the system optimum using a mathematical programming approach is based on the above formulated program. Several approaches are known to extend the model

with more sophisticated queuing algorithms (Ghali & Smith 1992 [26]) or other features.

Janson 1991 [39] presents the first attempt of a mathematical program to solve the dynamic user equilibrium, defining an objective function, with three constraints and seven (!) definitional constraints. Janson proves equivalence to the user equilibrium with the first order conditions of the Lagrangian.

The objective function is defined as:

$$Min\ z(x) = \sum_{ij}\sum_{p} \int_{0}^{x_{ij}^{p}} t_{ij}(\omega)d\omega \qquad \text{(eq 12)}$$

The objective function is similar to the objective function for the static case (eq. 1). An extra summation is needed over all periods p, and the index p has been added to x, to denote the proper period.

The objective function is subject to several constraints:

$$x_{ij}^{p} = \sum_{r}\sum_{d \le p} v_{rij}^{dp}\,\alpha_{ri}^{dp} \qquad \text{(eq 13)}$$

In this equation, v_{rij}^{dp} denotes the flow on link ij in period p for the trips departing from r in departure period d, and α_{ri}^{dp} is a zero-one variable indicating whether trips departing from origin zone r in departure period d reach node i in period p.
The constraint states that the total flow on link ij in period p is equal to the sum of flows departing from any origin r in any departure period $d \le p$, reaching node j via node i in period p.

The second constraint:

$$q_{rn}^{d} = \sum_{p \ge d}\left[\sum_{in} v_{rin}^{dp}\,\alpha_{ri}^{dp} - \sum_{nj} v_{rnj}^{dp}\,\alpha_{rn}^{dp}\right] \qquad \text{(eq 14)}$$

In this equation q_{rn}^{d} denotes the trip rate between origin r and node n departing in period d. The constraint states that the trips departing from r in departure period d going to n should be equal to all the trips that flow into n minus the trips that flow out of n, all coming from r, for every period $p \ge d$ in which trips enter and leave n.
The third constraint defines that all link flows should be greater than zero:

$$v_{rij}^{dp} \ge 0 \qquad \text{(eq 15)}$$

Several definitional constraints are added to this program to ensure temporally continuous paths, i.e. that trips are assigned to a link in the correct period. Furthermore extra constraints are necessary for the equivalence proof.

Optimal Control

A number of researchers have approached the dynamic traffic assignment problem by formulating it as a continuous time optimal control problem. One of the first efforts was initiated by Luque and Friesz 1980 [47], who reformulated the Merchant-Nemhauser model of a dynamic *system optimum* as a *continuous-time optimal control* problem, assuming known continuous dynamic OD-demand flows. Later Friesz 1989 [24] and Ran et al. 1990 [72] developed formulations for the *user-equilibrium* problem.

The objective function and the constraints are similar to the discrete case. Summation over the time periods has to be replaced by an integral over the total time span T. This results in minimizing the summation over all links a, of the integral of the link cost function:

$$\text{Min } z(x) = \sum_a \int_0^T \mathrm{f}[x_a(t)]\,dt \qquad \text{(eq 16)}$$

The parameter t now denotes a time *moment*. The cost function is denoted by $\mathrm{f}[x_a(t)]$, and represents the instantaneous cost on link a when it contains x vehicles at time t. The function is considered to be a continuous, increasing and convex function.

The first constraint ensures that, at a certain time t, the number of vehicles exiting node n to node c, $m_{nc}(t)$, is equal to the number of vehicles entering node n from node b, $\mu_{bn}(t)$, corrected by the amount of traffic generated, I_n, and attracted, O_n, at the node at time t.

$$\sum_b \mu_{bn}(t) = \sum_c m_{nc}(t) + I_n(t) - O_n(t) \qquad \text{(eq 17)}$$

All the nodes b have a link to node n, and all the nodes c have a link that starts at node n.

The second constraint is known as the state equation, relating the increase of traffic on an link a at time t to the difference between the instantaneous in- and out-flow.

$$\frac{dx_a(t)}{dt} = u_a(t) - v_a(t) \qquad \text{(eq 18)}$$

Where $x_a(t)$ is the number of vehicles on link a at time t, $u_a(t)$ the inflow and $v_a(t)$ the outflow of link a at time t.

3.4 Conclusions

Chapter two concluded with:

> *"Within the vision of ATPS, ATMS and ATCS, the need for new advanced traffic models is clear. To solve planning, management and control problems there is a need for more advanced traffic models. These models should be able to simulate traffic realistically and be suitable for calculating the effects of dynamic traffic management measures".*

Several traffic models are reviewed in chapter three and divided into two categories: simulation models and assignment models. Simulation models are in general not well suited for network problems. They do not consider route choice and are designed for small networks. They are well suited for studying the local impact of dynamic traffic management measures, such as the local impact of ramp metering, which makes these models suitable for ATCS applications. For ATPS and ATMS applications, models are required that can manage entire networks and consider route choice, to study rerouting measures and the network wide impact of ramp metering (coordinated ramp metering)

More appropriate models are assignment models. Assignment models are divided into static assignment models and dynamic assignment models. Static assignment models, however, lack the ability to model traffic realistically. The absence of dynamics in these models makes them unsuitable for modelling the impact of traffic jams. Several examples are known in which static assignment models can give wrong results [31].

Dynamic assignment models show several promising features. These models simulate the evolvement of traffic along the links of the network and can therefore model traffic more realistically. Several models have incorporated mechanisms to model dynamic traffic management measures. Dynamic assignment models were divided into heuristic models and mathematical models. The mathematical approach contributes several theoretically sound models. Proof of convergence and the existence of a unique solution are provided. One consequence of the strict mathematics is, however, that the amount of realism in these models is rather limited, and until now the literature contains very few implementation and little computational work, testing has only been done with very small networks.

Several of the heuristic dynamic assignment models have the potential to meet the requirements stated in chapter two. Some of the described models, however, lack the ability to calculate routes that are based on travel times while traversing the links, i.e. experienced route-choice. Experienced route choice is very important for short term predictions of the traffic situation, because with instantaneous routes, the traffic is not diverted optimally. One of the few models that incorporates dynamic experienced route choice is the model developed by Hamerslag [31]. This model was chosen as a starting point for the model described in the next chapters.

CHAPTER 4

THREE DIMENSIONAL ASSIGNMENT

4.1 Introduction

This chapter describes the dynamic assignment model: 3DAS (3-Dimensional ASsignment). The main objective of the research is to model the *evolution of traffic flows* in a network as realistically as possible. Further, the model should be able to simulate several instruments used for dynamic traffic management, such as ramp metering and variable message signs. The secondary objective of the research is that the model should be able to give a prediction of traffic flows within a reasonable period of time.

The model determines the time varying traffic conditions, flow and travel time, for every link during a certain simulation period. Like most other dynamic assignment models, the simulation period is divided into intervals of equal duration, referred to as periods. The link parameters may be defined separately for each period. The network and the travel demand are presumed given.
Two commonly used assumptions are made concerning route choice:
- all travellers are completely informed and are therefore taking future congestion into account.
- all travellers choose their minimum cost path.

These assumptions imply a user optimized approach. In a realistic situation, however, these assumptions are debatable. Although the traffic flow might lean towards user optimum due to habit of route choice, in the case of an accident, traffic often queues at the accident and does not divert to different routes automatically. If a certain route remains congested for a long time, or if rerouting information is provided, only then do travellers start to take different routes.
The theory of the model, as it is described in this chapter, is primarily constructed as a user optimal model; travellers choose their minimum cost path, and take *future* congestion into account. A second approach would be to choose the *instantaneous* shortest path. This path is based on the conditions in the network at the moment of departure or at the moment of a choice, i.e. at a junction. Instantaneous route choice does *not* take future congestion into account. A third strategy, to model more realistic route choice, is to use *historic* information for route choice.

The user optimal approach is the most complex method and is therefore described as the main approach. Instantaneous path finding is described as a simplification, and path finding based on historical information as an extra feature.

As discrete time units are used and for simplicity, the following two assumptions concerning the links are made:

- traffic conditions are homogeneous along a link and constant for the duration of each time period.
- the variables of a link, flow (q), density (ρ) and speed (v), in a certain period p are subject to a fixed relation described by:

$$q^p = \rho^p \cdot v^p \qquad \text{(eq 19)}$$

An outline of the model and its iteration process is described in section 4.2. Section 4.3 describes the input requirements. Several parts of the model are explained in detail in sections 4.4 through 4.6. Additions to the model, to simulate ramp metering and rerouting and dynamic departure times are described in section 4.7. Convergence properties are discussed in section 4.8.

4.2 Outline of the Model

The (user optimal) model determines the conditions in the network for every link during every period. An iterative process is used to solve the model. In every individual iteration, the route for every OD-relation in each departure period is calculated. According to these routes, the OD-matrix is assigned to the network. New network conditions are determined for every period, before each iteration, based on the conditions resulting from the previous iteration. The iteration process stops when the stop criterion is reached, i.e., when the loading of the network remains essentially unchanged from one iteration to the next.
The iteration process is basically the same as the one used for static equilibrium assignment methods, and is shown as a flow chart in figure 1.

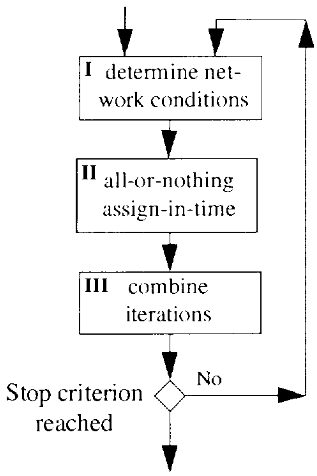


Figure 1: Iteration process flow chart

Each block in the flow chart is referred to by a roman number.

I. The cost to traverse a link is calculated for each link and each period. Cost is based on the conditions in the network as a result of the previous iterations. In the first iteration, when the network is empty, costs are based on free flow conditions. The relationship between the load on a link and the cost is determined by cost functions and is discussed in section 4.4.

II. An *All-or-Nothing-Assignment-in-Time* is performed to assign all OD-pairs for all departure periods to the network, based on the determined network conditions. The shortest path in time is calculated for each OD-relation and each departure period. In accordance with these shortest paths, the OD-flow is assigned to each link in the path. This *All-or-Nothing-Assignment-in-Time* is further explained by the flow diagram given in figure 2.

III. The end result of the current iteration is obtained by combining the results of the last iteration with the results of the previous iteration. The combination is attained by taking a certain fraction (1-λ) of the link loads of the previous iteration and adding the remaining fraction (λ) of the currently calculated link loads. This is described further in section 4.6.

These three steps are performed in an iterative process. During each iteration, the travel conditions for every link and each time period are updated, an *All-or-Nothing-Assignment-in-Time* is performed and the result is combined with the previous result. The stop criterion is applied at the end of each iteration. The process stops if the results of the last iteration are sufficiently close to the link flows of the previous iteration.

In an *All-or-Nothing-Assignment-in-Time* (**II**), the shortest paths are calculated for all OD-relations and all departure periods. Trips, obtained from the OD-matrix, are assigned to these paths. This procedure is explained in the flow diagram given in figure 2:

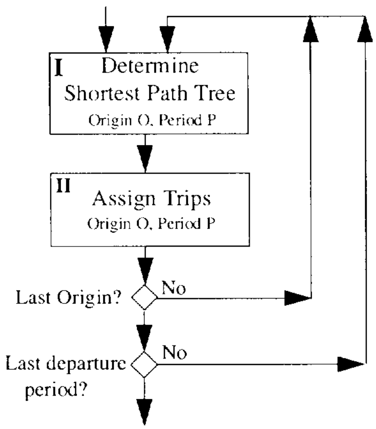


Figure 2: All-or-Nothing-in-Time flow chart

I. The shortest path tree for the origin to all destinations is calculated, starting with the first origin and the first departure period. These shortest paths are based on the travel costs of the links for the periods in which trips actually traverse these links. Pathfinding is explained in section 4.5.1.

II. Trips for the current origin and departure period, specified by the OD-matrix, are assigned to the network based on the shortest path tree. Trips are assigned to a link in the period in which they actually traverses the link. Assignment is explained in section 4.5.2.

This iteration process is repeated until all origins and all departure periods are treated. The sequence in which the origins and the periods are dealt with is not important. Each combination of origin and departure period is treated as totally independent, because there is no updating of travel conditions between the calculations. Overloading of a link within one all-or-nothing assignment is allowed. A special queuing procedure will take care of the overflow in the next iteration. This queuing procedure is described in section 4.7.

4.3 Input Requirements

The basic input requirements of the model are a network and an OD-matrix. Special input requirements to model traffic management instruments are described in section 4.7.

4.3.1 Network

The network is given in the form of a directed graph $G(N,A)$, with A the set of links to represent the road segments and N the set of nodes to connect the links. The network is *connected*, meaning that it is possible to get from each node to any other node by following a *path*. A path is a sequence of links leading from one node to another. The links are given parameters to represent network properties in general and road characteristics in particular. Some link parameters are constant in time, e.g., the link length; others may change, e.g., the number of lanes, due to an accident or road construction. Several link parameters are therefore defined separately for each period, however, each parameter is constant over the length of the link and for the duration of the period.
In principle there is no minimum or maximum link length, neither are there any requirements concerning the variance in link lengths, long links however are less sensitive to a fluctuating traffic demand and therefore to congestion, due to the assumed homogeneous traffic conditions along the links. The influence of link length is discussed in section 5.2.

4.3.2 OD-Matrix

Travel demand is specified by an OD-matrix. A time dependent OD-matrix is required for the model because the model takes time variance into account. An OD-matrix that is similar to the OD-matrix used for static assignment models is needed for each period. This results in a three dimensional OD-matrix, in which the trip q_{rs}^{d}, denotes the trips from r to s departing in period d.

Normally, such a three dimensional matrix is not available, and a traditional matrix with trips q_{rs} is used. This matrix is defined for a larger time span and is divided in periods with a *departure time function.* Such a function can be defined for each origin and denoted as a fraction ϕ_r^d. The fractions ϕ_r^d specify, for each departure period d, the percentage of trips that depart from origin r.

$$q_{rs}^d = q_{rs} \cdot \phi_r^d$$
$$\sum_d \phi_r^d = 1 \qquad \forall d \tag{eq 20}$$

There are several methods for defining the departure time function. The simplest method is to use the same function for all OD-entries. A more complicated method takes the travel time to the destination into account and calculates an individual function for every trip. A practical method is to define certain groups of OD-entries that will travel according to the same departure function. When n groups are distinguished, only n departure functions are required, this process is illustrated in figure 3.

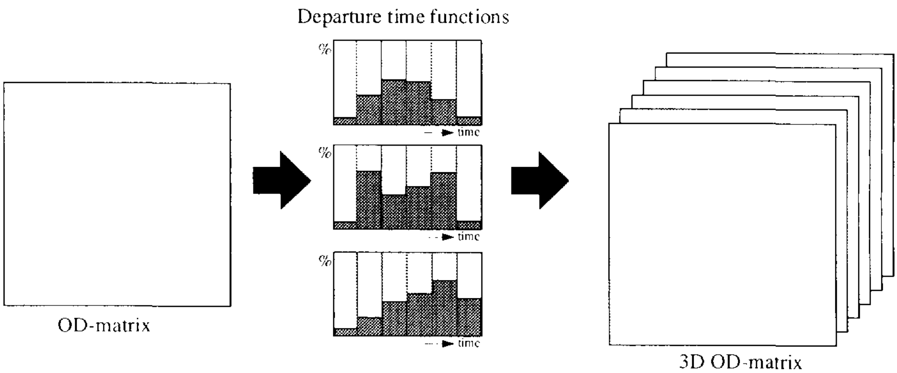


Figure 3: Determining a 3-dimensional matrix using departure time functions, in this example three departure time functions are used and six periods.

The accuracy required for the OD-matrix depends very much on the intended application. The methods described are sufficiently accurate for ATPS (see chapter 8), more accurate methods are required for ATMS and ATCS. The need for new, more advanced models for the estimation of dynamic OD-matrices has been recognized in the literature and they are now being developed at a rapid pace.

4.4 Travel Time Functions

Following from section 4.1, the travel cost for each link in each period is based on the traffic conditions on a link in the previous iteration (figure 1). Usually travel time is determined using a function describing the relationship between flow and travel time (e.g. the BPR function).

In a dynamic assignment model, however, a traveltime-flow function causes inconsistencies, because of an incorrect relationship between flow and travel time.

If flow, density and speed on a road section are observed in reality, a different relationship is found. In non-congested situations, traffic will flow freely and with increasing density and flow, travel time increases. In congested situations speed will drop rapidly, flow will **decrease** and travel time will increase. The observed relationship is illustrated as a dotted line in figure 4.

In reality flow declines with congestion which makes a traveltime-flow function **unsuitable** for describing a congested situation. A true traveltime-flow function would show duality, because with one flow, two travel times are possible (see figure 4).

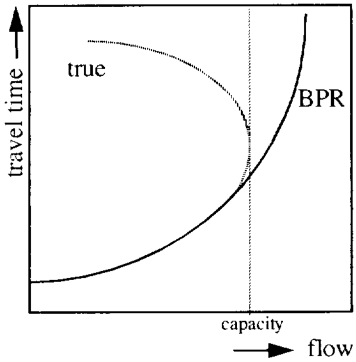


Figure 4: Traveltime-flow diagram

The relationship between **density** and travel time does not exhibit this duality, making it more suitable for dynamic assignment. Instead of using flow as the explanatory variable, density is used.

In a traveltime-density function, travel time increases monotonously. The true form of the traveltime-density function was subjected to further research (see section 5.3). The function used during model development was a speed-density function adopted from Smulders [86]. The speed v is given by two functions of the density ρ:

$$v(\rho) = \begin{cases} v^{max} \cdot \left[1 - \dfrac{\rho}{\rho^{max}}\right] & 0 < \rho < \rho^{crit} \\ \phi \cdot \left[\dfrac{1}{\rho} - \dfrac{1}{\rho^{max}}\right] & \rho^{crit} < \rho < \rho^{max} \end{cases} \qquad \text{(eq 21)}$$

In equation 21, v^{max} is the free flow speed, ρ^{crit} is the critical density and ρ^{max} is the maximum density. Maximum density represents a no-motion traffic jam. Maximum flow (capacity) is reached at critical density. The parameter ϕ is chosen to make the function continuous at ρ^{crit} ($\phi = v^{max}\rho^{max}$).

This speed-density relationship is plotted in figure 5 along with the matching flow-density and speed-flow relation. When the critical density is reached, speed drops more rapidly and the corresponding flow also decreases.

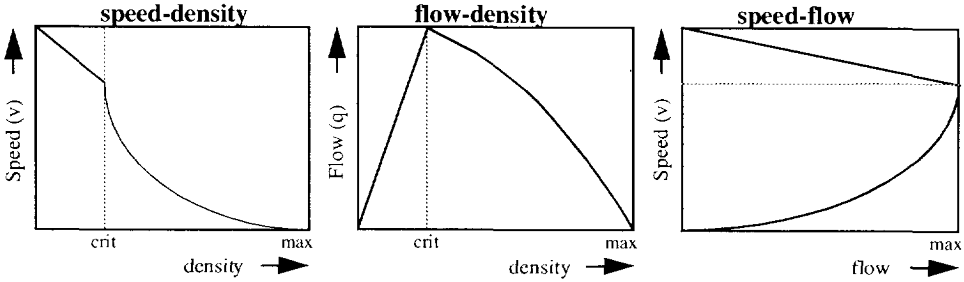


Figure 5: Speed-density function and flow-density function

In accordance with the density on a link, the matching speed (or equivalent travel time) can be found using this *speed-density* function. According to Smulders [86], critical density should be 24% of the maximum density.

4.5 All-or-Nothing-Assignment-in-Time

This section describes the *All-or-Nothing-Assignment-in-Time*, which incorporates the path-finding method and the assignment method. Four different path finding methods and two different assignment methods are described.

4.5.1 Pathfinding Methods

Paths are determined by the arrival time[1] at each node of the path. As in static assignment models, the arrival time at a node is calculated by summing the consecutive travel times of the links in the path. In a dynamic assignment model, taking future congestion into account, a difference arises because travel time on a link can vary over periods.

Pathfinding is illustrated by the following example (see figure 6):

1. The cost to traverse a link is in this dissertation denoted by travel time. Any other variable or combination of variables is possible, and a more general term would be 'generalized cost'. For the sake of clarity, however, travel time is used.

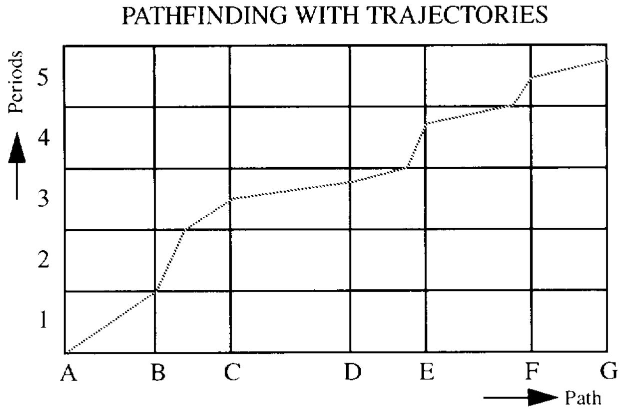


Figure 6: Pathfinding with trajectories

EXAMPLE: In figure 6, the path in time is drawn as a trajectory (distance-time) graph. The x-axis shows the path. The path consists of the nodes A,B,C,..G. It starts at node A (Origin) and ends in node G (Destination). The y-axis shows time, which is divided into 5 periods. A dotted line represents the trajectory of the trip.
Assume that trips that enter link A-B encounter a travel time (in this example) exactly equal to the duration of one period. So trips leaving node A in the first period, arrive in node B at the beginning of the second period. After link A-B they traverse link B-C (the next link in the path). At link B-C they will encounter a travel time that is valid for period two. The angle of the trajectory is adjusted because the travel time is different. In this example the travel time for link B-C in period two is greater, so the trajectory becomes steeper. Halfway along link B-C, the travellers enter the third period which also has a different travel time, so the trajectory is again adjusted to the travel speed that is valid for link B-C in the third period. Consequently link B-C is traversed in ±1½ time periods and the arrival time at node C is halfway through the third period. The travel time to D is ±¼ period, so the travellers arrive in node D at ±¾ through the third period, and so on.

Arrival times at the nodes are determined by following the trajectories. These node arrival times are used to define the shortest path tree in the same way as other (static) shortest path finding algorithms. The links that assemble a path from node i to node j with the earliest arrival time in node j, belong to the shortest path from node i to node j.
The example illustrates a method to calculate the arrival times at the nodes. Considering that traffic conditions are homogeneous along a link and constant for the duration of a period, this method is the most accurate. During the research, three other methods were developed to calculate arrival times at the nodes. The first method is similar to the method used by Hamerslag in [31]. The second and the third methods are improvements on this method. The method described in the example above will be referred to as the fourth method.

In general the time to traverse link a from node i to j is determined by calculating the arrival time at node j. If the arrival time at node i is denoted as d_i and the arrival time in node j is denoted as d_j, then, in the most simplistic way:

$$d_j = d_i + \text{travel time from } i \text{ to } j \qquad \text{(eq 22)}$$

The time to traverse link ij may, however, take several periods, and each period can have a different travel time. The travel time for link ij in period p is denoted as t_{ij}^p. Further, the actual *time period of departure* from node i is denoted as p_i.

The travel time from node i to node j is calculated differently and with increasing complexity and accuracy by each of the four methods. Each method is explained below using an example graph. The travel times in these graphs are the same for each method.

I. The first method is the simplest method. The travel time to traverse link ij is the travel time valid in the period p_i when the trip enters the link, p_i denotes the period in which the trip departs from node i.
This travel time is equal to $t_{ij}^{p_i}$. and remains unchanged until the end of the link is reached. If the next period is entered, there is no adjustment of the trajectory. The arrival time in node j is given in equation 23.

$$d_j = d_i + t_{ij}^{p_i} \qquad \text{(eq 23)}$$

This method is illustrated in figure 7.

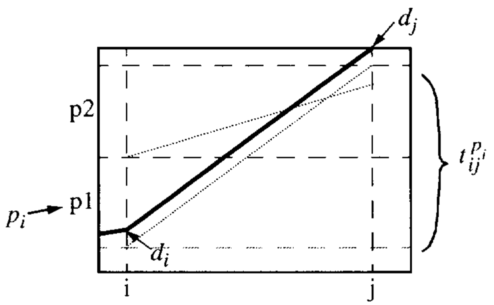


Figure 7: Example of method I for calculating arrival times at the nodes. The travel times valid in period p1 and p2 for link ij are given as dotted lines. The dashed lines represent period and link borders. The solid line represents the trip trajectory

The method is computationally very cheap, but problems occur with this method when the travel time in the next period differs very much from the current period. It is possible for traffic that departs one period later to arrive earlier. This can be observed in figure 7. The traffic that departs in period *p2* will follow the dotted trajectory which arrives at node j earlier than the traffic departing in period *p1*, which follows the solid line trajectory.

II. The second method tries to deal with radical changes in successive periods by checking whether *skipping* to the next period will result in a shorter travel time. If the uniform period length is denoted as T and the number of periods that are skipped is denoted as λ, e the arrival time at node j is given in equation 24

$$d_j = d_i + \lambda T + t_{ij}^{p_i + \lambda} \qquad (\text{for Min } \lambda \geq 0) \qquad \text{(eq 24)}$$

The value of λ is chosen in such a way that d_j is minimized.

The method is illustrated in figure 8.

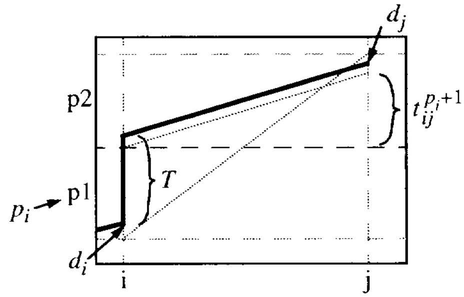


Figure 8: Example of method II for calculating arrival times in the nodes, the travel times valid in period p1 and p2 for link ij are given as dotted lines, the dashed lines represent period and link borders, the solid line represents the chosen trajectory, λ is equal to one.

A better adaptation to sudden changes is achieved using this method, but it is still possible to depart later and arrive earlier.

III. The third method is almost equal to the second method. Instead of skipping a full period, the method skips to the *beginning* of the next period. The arrival time at node j is now given by equation 25.

$$d_j = \begin{cases} d_i + t_{ij}^{p_i} \\ (p_i + \lambda)\,T + t_{ij}^{p_i+\lambda} \end{cases} \quad \text{(for Min } \lambda \geq 0) \qquad \text{(eq 25)}$$

The method is illustrated in figure 9.

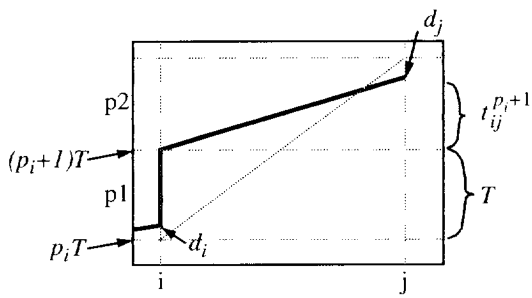


Figure 9: Example of method III for calculating arrival times in the nodes, the travel times valid in period p1 and p2 for link ij are given as dotted lines, the dashed lines represent period and link borders, the solid line represents the chosen trajectory, λ is equal to one.

With this method it is no longer possible to depart later and arrive earlier.

IV. The fourth method adapts the travel time every time a new period is entered. This method was described in the example at the beginning of this section. The mathematical formulation of this method is rather complex. The total travel time to traverse link ij becomes the sum of several different travel times because a different travel time is possible for each period.

Assume that during each period p a certain distance is travelled on link ij. This distance, expressed as a fraction of the link length covered in period p, is denoted α^p and can be calculated using the travel time and the period length:

$$\alpha^p = \frac{T}{t_{ij}^p} \qquad \text{(eq 26)}$$

The moment of departure from node i can be at any time during a period, so the fraction of distance travelled in the first period p_i can be less than α^p. This fraction is equal to σ_i and denoted in equation 27.

$$\sigma_i = \left[p_i - \frac{d_i}{T}\right]\alpha^{p_i} \qquad \text{(eq 27)}$$

The same occurs for the period p_j in which node j is reached. The fraction of distance travelled on link a in that period can be less than α^p, because the end of the link has been reached. The fraction travelled is equal to σ_j and denoted by equation 28.

$$\sigma_j = \left[\frac{d_j}{T} - (p_j - 1)\right]\alpha^{p_j} \qquad \text{(eq 28)}$$

Using equations 26 and 28, the arrival time at node j is:

$$d_j = (p_j - 1)T + \sigma_j t_{ij}^{p_j} \qquad \text{(eq 29)}$$

According to equation 29, to calculate d_j the arrival period p_j has to be known. To calculate this period, the distance travelled on link ij must be calculated. Thus, the sum of the fractions travelled on link ij has to be equal to one. This constraint is given in equation 30.

$$\sigma_i + \left[\sum_{p = p_i + 1}^{p_j - 1} \alpha^p\right] + \sigma_j = 1 \qquad \text{(eq 30)}$$

The method is illustrated in figure 10.

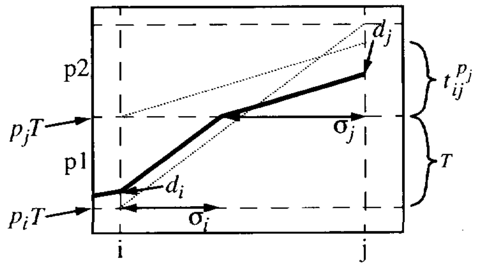


Figure 10: Example of method IV for calculating arrival times at the nodes, the travel times valid in period p1 and p2 for link ij are given as dotted lines, the dashed lines are period and link borders, the solid line represents the chosen trajectory.

Although this method is computationally more expensive, it represents the most accurate calculation of the arrival time at node j, and it is no longer possible to depart later and arrive earlier.

The first pathfinding method is the cheapest for implementation purposes, the fourth pathfinding method is theoretically the most accurate, but involves more computational effort. The different methods are tested in section 5.5.

When *instantaneous route choice* is considered, the method simplifies. Instantaneous route choice is based on travel times which are valid for the period of departure. This means that it is not necessary to follow the trajectory through time. An arrival time at the destination is derived by summing the travel times which are valid in the period of departure. Trajectories are still required however, to assign trips to the paths.

4.5.2 Assignment Method

The OD-matrix is assigned for each origin and each departure period according to the calculated trajectories. To calculate the contribution of a traveller to the density on link a in period p, the *duration of presence* on link a in period p must be calculated.

If we focus on the travellers specified by one OD-relation, then two situations can occur:

- **I.** *Several links are traversed during one period.* In this case the travellers are only present for a part of the period on the link, and therefore should only be assigned for this part.
- **II.** *One link is traversed during several periods.* Travellers are present on the link during multiple periods and should be assigned entirely for each individual period.

Two different methods were used during the development of the model: The '*trajectory method*' and the '*surface method*'.

The trajectory method assumes that all traffic departs at the beginning of the period. This method is computationally less demanding than the surface method.

The method is illustrated by the example given in figure 11:

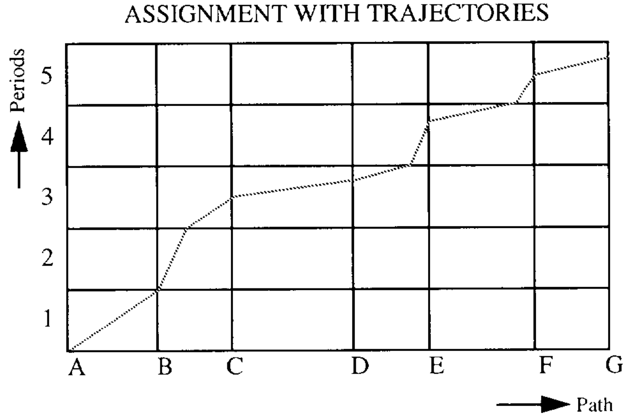


Figure 11: Assignment with trajectories

EXAMPLE: The example given in figure 11 uses the same trajectory as in figure 6 and assumes that pathfinding method IV was used. The x-axis shows the path along the nodes A,B,C,..G, and the y-axis shows time periods one through five. The same travel times as in figure 6 apply and trips departing from node A in the first period with destination G will follow the trajectory calculated by the pathfinding method. The trajectory is displayed as a dotted line. During the first period, the trips traverse the entire link A-B. So in the first period, all the cars are present on link A-B and are all assigned to link A-B in the first period. In the second period, the cars are present on the *first part* of link B-C during the entire duration of the period, so they will all be assigned to link B-C in the second period. In the third period the cars are *still* partly present on link B-C, but also partly on C-D and D-E. So in the third period, they are assigned according to the *duration* of the time they are present on each of these links. In this example this results in ±50% to B-C, ±30% to C-D, and ±20% to D-E. When we assume that the demand for OD-relation A-G departing in the first period is equal to 100 cars, then the final result is 100 cars on link A-B in the first period, 100 cars on link B-C in the second period, 50 cars on link B-C in the third period, 30 on link C-D in the third period and 20 on link D-E in the third period, and so on.

The method described in the example given above has to be carried out for all origins and all departure periods (as described in the flow diagram of figure 2 on page 35). So the density ρ_a^p on a link a in period p is a summation of the contribution of all OD-relations and all departure periods d, that make a contribution to link a (a is the link between node i and node j). This summation is given as:

$$\rho_a^p = \sum_r \sum_s \sum_d \rho_{ars}^{pd} \qquad \text{(eq 31)}$$

Where ρ_{ars}^{pd} represents the contribution of the OD-pair q_{rs}^d from origin r to destination s departing in period d to the density on link a in period p. The definition of ρ_{ars}^{pd} is given as:

$$\rho_{ars}^{pd} = \gamma_{ars}^{pd} \cdot q_{rs}^d \cdot \delta_{ars}^d \qquad \text{(eq 32)}$$

where δ_{ars}^d is equal to one if link a is on the shortest path for OD-pair q_{rs}^d and zero otherwise. The fraction in which OD-pair q_{rs}^d contributes to the density on link a in period p is denoted by γ_{ars}^{pd}, and specified in equation 33.

$$\gamma_{ars}^{pd} = \frac{d_j^p - d_i^p}{T} \qquad \text{(eq 33)}$$

T is the uniform period length and the parameters d_i^p and d_j^p are given in equations 34 and 35, which denote the departure time in period p and the arrival time in period p while travelling between nodes i and j.

$$d_i^p = \max\,[d_i,\ (p-1)T\,] \qquad \text{(eq 34)}$$

$$d_j^p = \min\,[d_j,\ p \cdot T\,] \qquad \text{(eq 35)}$$

The interpretation of d_j^p - d_i^p is the duration of time for which traffic is present on link a in period p. If the traffic departs from node i during period p then $d_i^p = d_i$; otherwise the traffic is already on its way to j and d_i^p is the start time of the current period p, which is equal to $(p\text{-}1)T$. For d_j^p the value is either the arrival time in node j, which is d_j or, when node j is not yet reached in period p, it is the moment when the period is left, which is equal to $p \cdot T$.

EXAMPLE: If equations 34 and 35 are illustrated using the situation of figure 11, then for the first period, p=*1*, the value for d_A^1 is equal to 0 and d_B^1 is equal to 1. So $\gamma_{(AB)AG}^{11} = 1$. (This is γ for link *(AB)* in period *1* for OD-relation *A-G* and departure period *1*). For the second period, d_B^2 is equal to 1, and d_C^2 is equal to 2, because d_C is equal to 2½. So $\gamma_{(BC)AG}^{21} = 1$. For the third period d_B^3 is equal to 2 and d_C^3 is equal to 2½, so $\gamma_{(BC)AG}^{31}$ = ½, d_D^3 is equal to 2¾, so $\gamma_{(CD)AG}^{31}$ = ¼, and d_E^3 is equal to 3, because d_E is 3¾. So $\gamma_{(DE)AG}^{31}$ = ¼. The first part of figure 11 is duplicated in figure 12 and the appropriate labels added.

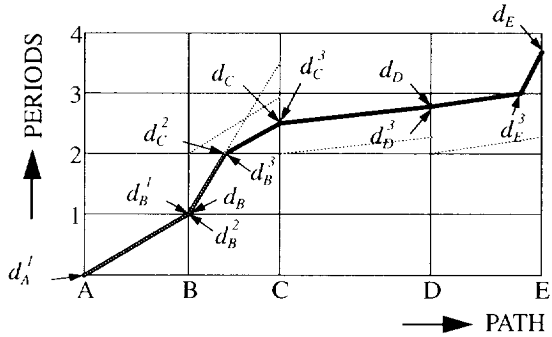


Figure 12: Example to calculate the duration of presence

The second assignment method, the *surface method*, assumes that all travellers, departing in one period leave the origin uniformly spread over the duration of the period. Travellers are assigned according to the *surfaces* between the trajectory the first car follows and the trajectory the last car follows. A similar approach has been described by Kroes [43].

The assignment method is illustrated by the example given in figure 13:

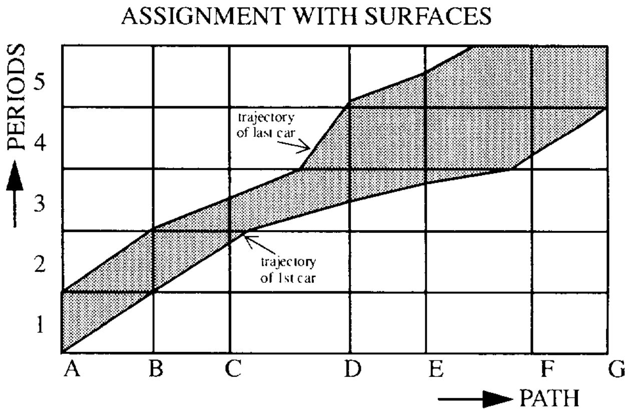


Figure 13: Assignment with surfaces

EXAMPLE: Figure 13 displays a path-time diagram. The x-axis shows the path, consisting of the nodes A,B,..G, and the y-axis shows the time periods, numbered from one to five. The trajectories are plotted in a fashion similar to that explained in figure 11. A trajectory is plotted for the first car leaving node A in period one, and for the last car leaving node A in period one. The trajectory for the last car in a period is the same as the trajectory for the first car in the *next* period. As the travel time on link A-B is equal to the duration of one period, the *first* car leaving node A in the first period arrives at node B at the end of period one. The travel time for link A-B in the second period is, in this example, also equal to the duration of one period, so the *last* car leaving in the first period arrives at the end of the second period at node B. All the *remaining* cars that depart in this period are spread between these two trajectories. So the traffic is assigned according to the (grey) surface between the two trajectories.
In this example it means that, in the first period, 50% of the traffic is assigned to link A-B. (The first period is special, because only half the traffic has departed). In the second period, the total surface of 100% covers A-B for ±50%, B-C for ±45% and C-D for ±5%. Traffic, departing from origin A and heading for destination G in time period one, is assigned to these links according to these percentages. In the third period the traffic is assigned to B-C, C-D, D-E and E-F. As density is assigned, the total amount of cars present in each period should be 100%, except for the period of departure and the period of arrival.
The link C-D shows that there can be a significant difference in travel time between the last car and the first car. Thus the distance between cars can increase or decrease, and the demand given by a certain OD-relation can spread over several links and periods.

The final density on link a in period p is a summation of the contribution of all OD-pairs from all departure periods that traverse link a in period p, as described by equations 31 and 32. The parameter γ is now calculated according to the example described in figure 13.

Using densities instead of flows offers, besides a more realistic representation of the travel time function, the possibility to check the conservation of traffic present in the system. The amount of traffic present on the network should be equal for each period; except in the period of departure from the origin and in the period the traffic reaches its destination, because, in the departure period a certain percentage has not entered the network and in the arrival period a certain percentage has already left the network.

The trajectory and surface assignment methods are evaluated in section 5.5.

4.6 Iteration Process

One iteration can be called an *All-or-Nothing-Assignment* because all the traffic has been assigned to one physical route, however, because the assignment is dynamic one iteration has no practical meaning due to temporal discontinuities (as Janson [38] calls them). A temporal discontinuity denotes the difference between the travel time used to assign the traffic and the travel time resulting from the assigned load. In a dynamic assignment a discontinuity can mean that traffic has been assigned in the wrong periods. Even if there is only one physical route, several iterations are needed to eliminate the discontinuity. A temporal discontinuity is eliminated when densities, and with densities, travel times, do not change between subsequent iterations, only then has a stable solution been reached and the iteration process can be stopped.

The iteration process takes care of the distribution of traffic over different routes and also the distribution of traffic over different periods. Even when no alternative routes are considered the iteration process is required to assign the travellers to the correct periods.

> **EXAMPLE:** Consider a linear network with ten links, in which the fifth link represents a bottleneck. The OD-matrix defines 100 cars, departing in the first period, travelling from left to right. In the first iteration free flow travel times are used to assign the cars to the links. As there is no delay at the bottleneck in the first iteration, all the 100 cars traverse the last link in period x. After several iterations, however, the delay at the bottleneck has increased, and the 100 cars traverse the last link in period x+1. This simple example illustrates that several iterations are required to get the 100 cars into the correct period(s).

Experiments were carried out using two methods to gain some understanding of the iteration process:

I. The first method calculates new routes before every iteration.
II. The second method uses sub-iterations. During the sub-iterations the physical routes remain constant. After the traffic is assigned to the correct periods on the current chosen route, new routes are calculated.

Both methods converged to the same solution, with the first method converging faster.

The subsequent iterations are weighted using:

$$\rho_a^P = (1-\lambda)\cdot\rho_a^{p,\,old} + \lambda\cdot\rho_a^{p,\,new} \qquad \text{(eq 36)}$$

Where ρ_a^p represents the density on link a in period p, $\rho_a^{p,\,old}$ is the result of all previous iterations and $\rho_a^{p,\,new}$ the result of the last iteration. In work to date, λ has been fixed (e.g. 0.2 or $\lambda = 1/(i+0.5)$, where i is the iteration number). Experiments have been done with an heuristically derived λ, based on the difference between the link travel times and path travel times.

The solution reached in this way tends to a user equilibrium. No mathematical proof of uniqueness or convergence is given. Several experiments with different networks and OD-matrices, however, have shown a convergence in all cases after 20 to 40 iterations, and equal path travel times on alternative routes, when used. The largest number of iterations were needed for networks with heavy congestion, because the initial iteration assigns too much traffic too far downstream. It should be realised that the number of iterations also determines the maximum number of alternative routes, because one route is calculated per iteration for each

OD-relation.
Convergence properties of the model are discussed in section 4.8 and examples are given to show that realization of a user equilibrium is plausible.

4.7 Additional Features of the Model

Some extra features are needed in the model for a realistic description of traffic and to be able to model traffic management instruments. These features are categorized in three sections. Section 4.7.1 explains extra features to improve the *performance* of the model, such as traffic jam building upstream, more realistic pathfinding and ramp merge behaviour. Section 4.7.2 explains additions to model several management instruments, such as ramp metering, rerouting, speed control and tidal flow. Dynamic departure times and multi class, multi mode and park & ride features are described for planning purposes in section 4.7.3.
These extra features have been added to the model on behalf of certain applications of the model. The importance of these features is very much secondary to the core of the model, the evolvement of traffic in time, described in sections 4.2 to 4.6.

4.7.1 Performance Features

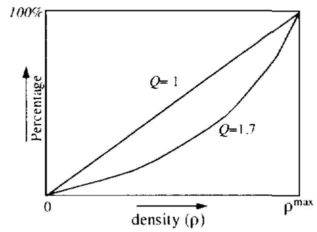
This section describes three features to improve the performance of the model.

Traffic jam building upstream
Two different methods for traffic jam building upstream were developed:

- In the model and in reality, density on a link has a maximum. Maximum density represents a no-motion traffic jam. When, during the calculation, maximum density has been reached on a link, no more trips can be assigned to it. Within one iteration, however, more than the maximum amount of traffic is allowed on a link. This will result in a higher density than maximum density. Travel time on the link will now increase considerably and the routes chosen in the next iteration will most likely not include the overloaded link. When the same route appears to be the shortest again, the traffic is not assigned to that link but to the preceding link in the path in the same period. After each iteration the previous results are combined with the current and the density of the overloaded link will decrease again. Eventually no overflow will occur. Upstream of the overloaded link, as a result of the reassigned traffic, density will increase and because of increasing density, travel time on the upstream link will increase. Depending on demand, this effect will proceed upstream and results in a queuing effect, which simulates the traffic jam building upstream.
- The second method does not wait for the density to reach maximum density, and starts to queue an increasing percentage of the traffic with increasing densities. The method is similar to the first method, but instead of queuing 100% of the traffic to the preceding link when maximum density is reached, this method determines a certain amount of traffic to be assigned to the preceding link as a function of the density.
  The function[1] determines the percentage to be assigned to the preceding link in the next iteration according to equation 37.

1. Upon the completion of this research more and better data became available to validate this function. The function has been improved and is discussed in section 9.3.

$$F(\rho) = \left(\frac{\rho}{\rho^{max}}\right)^{Q}$$ (eq 37)



The parameter Q influences the *speed* with which the queue builds upstream. With increasing density on a link, more traffic is assigned to the preceding link. When traffic is assigned to a link for several iterations, the density will still reach maximum density. At this point the function F will be 100% so maximum density will not be exceeded, because in the next iteration *all* the traffic is assigned to the preceding link.

As a result of the assignment of traffic to a preceding link, a discontinuity in flow will appear. The discontinuity involves only the links that are involved in the queuing process, up and downstream from that section no discontinuity will appear. Moreover, because of the influence of queuing on the travel times, blocking back in the final iterations is minimal, resulting in only minor discontinuities.

More Realistic Route choice

The shortest paths are calculated in every iteration, based on the assumption that drivers are completely informed about future travel times. In the event of an accident, represented as a decreased capacity, the traffic will reroute automatically. In reality this will not happen. Instead, the traffic will queue on the route with the accident. Only when information of the accident is available to the drivers or when the duration of the congestion lasts for a long time will traffic start to take alternative routes. For the model, this means that route choice should not be based on travel times resulting from the previous iteration, but rather on the travel times from before the accident. This suggests the use of *old* travel times for route choice. To implement this, two travel times are used for each link. An *historic* travel time and an *actual* travel time. Historic travel times are used for route choice, and actual travel times for assignment. Since in reality, most travellers will eventually be informed (by radio) of the accident, historic travel times are updated by actual travel times, if there is a significant difference over a long time period. To maintain diversion of traffic over different routes, travel times are only corrected slightly towards the actual travel time. On links where rerouting information is available, historic travel times are more quickly updated by actual travel times.

On-ramp link dependency (merging)

At sections of a freeway with on-ramps, less efficient throughput takes place. The more cars that enter the freeway at the on-ramp, the more delay occurs. Delay on the freeway is influenced by the ratio of the traffic on the on-ramp and the traffic on the main lanes. If the merging process is without delay the capacity of the freeway is not influenced, however, when the flow on the on-ramp becomes larger, the capacity of the freeway decreases, because of the merging process.

More realistic behaviour is achieved by making the capacity of the section of the freeway with the on-ramp dependent on this ratio. It is sufficient to adapt the critical density because maximum density determines the capacity of a link.

The relation used in the model to achieve the dependency is given in equation 38.

$$\rho_m^{crit} = \rho^{crit} - M\rho_r\rho_m\rho^{crit} \qquad \text{(eq 38)}$$

Where M is a parameter to adjust the impact of the dependency, ρ_m^{crit} is the new critical density valid for the main section. ρ^{crit} is the normal critical density of the main section, ρ_r is the current density at the ramp, and ρ_m is the current density at the main section.

M is a value between 0 and 1 and determines the influence the ramp has. The dependency is visualized in figure 14 for two different values for M, as a function of the density on the main section of the freeway and the density on the ramp

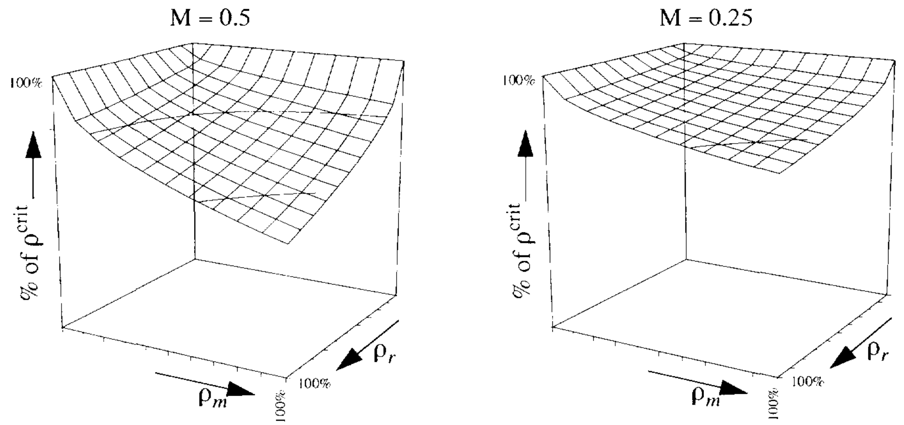


Figure 14: On-ramp link dependency

If the demand on the freeway is close to 100% and the demand on the on-ramp is close to 0%, the influence on critical density is minimal, however, when there is 100% demand on the on-ramp and 100% demand on the freeway, the critical density of the section is lowest. The true form of the dependency relation is probably not symmetric and needs further research.

4.7.2 Management Features

This section describes four features to add management capabilities to the model.

Rerouting

When it is necessary to force a certain percentage of the traffic to comply with rerouting measures, the model allows the entering of fixed routes between two nodes for certain groups of OD-pairs. The percentage of the trips to be assigned to the fixed route can be specified.
If the impact of rerouting using a certain percentage of equipped vehicles (probes) is required, the OD-matrix can be split into two parts. One part to represent the equipped vehicles and the other part to represent the unequipped vehicles. Different routing strategies are possible for both parts.

Ramp metering
Ramp metering is implemented in the model as a limited capacity of an on-ramp. Since ramp metering causes a more efficient merging process, the on-ramp link dependency explained above can be made less strict, by decreasing the value of M. More research towards behaviour at ramp metering installations is needed.

Speed control measures
The speed on certain links in the network can be controlled per period. The relationship between speed and density can also be changed. A different speed-density function can be used for homogenizing measures, and for severe weather conditions it may also be necessary to change the relationship between speed and density.
Some research on the influence of the speed-density relationship is described in section 6.3.3.

Tidal flow (reversible lanes)
The effects of tidal flow measures can be investigated by changing the maximum densities of links. A two lane link is easily transformed into a three lane link by increasing the maximum density of the link.

4.7.3 Planning Features

This section describes two *potential* features for planning applications.

Dynamic departure times
One of the first reactions of travellers who are exposed to recurrent congestion is to change their departure time. So, for planning purposes or to model the impact of dynamic traffic management, the assumption that the departure times remain fixed is not very realistic. Timing of commuting trips, particularly home to work, is strongly influenced by arrival time preferences and because the arrival time depends on the travel time, there is a strong correlation between departure time, travel time and route choice.
There is a growing literature on dynamic departure times, including Abkowitz 1981 [2], De Palma et al. 1983 [63], Hendrickson and Plank 1984 [35], Ben Akiva et al.1986 [6], Mahmassani and Herman 1992 [50], Boyce et al.1992 [12] and Janson 1993 [39].
In general departure time is a function of departure time preferences and expected travel times. This means that route choice and departure time choice have to be done simultaneously. Normally to incorporate this into a model, an extra penalty is put on early or late departure. This extra penalty can be seen as an extra link with a certain cost. The cost increases with later and earlier departure times. The function has a direct influence on the final results.
The simultaneous dynamic departure time choice is easily built into the pathfinding method by building the extra cost of early or late departures into the pathfinding algorithm as an extra link to a different departure period. Pathfinding tries to find minimal paths and automatically determines optimal departure times. See e.g. Kroes & Hamerslag [43].
More sophisticated algorithms have been developed by several authors. Janson 1993 [39], for example, describes a model in which the departure time choice depends on expected arrival time. The cost of departing later or earlier is a function of the arrival time. In this model, an extra penalty is put on early or late arrival, instead of early or late departure. This assumption is probably more realistic, but the model becomes somewhat more complex.

Multi Class, Multi Mode, Park & Ride

The effects of dynamic traffic management are not limited to changing routes, or changing departure time. A dynamic traffic management system can improve utilization of the overall capacity of the traffic system. This includes a redistribution of traffic to different modes and the awareness of services such as park and ride. The joint effect of dynamic traffic management on route, departure time and mode choice needs to be integrated into one model and handled simultaneously. Some models actually have all these capabilities, e.g. Boyce et al. 1993 [12].

Although the number of assumptions are fairly large in such a model, the required data very hard to get, and the interaction between the different choices largely unknown, they do have their value for decision making, because some insight is gained into the interaction between these aspects. The detailed results of these models, however, should be used with caution.

Multi class features are easily added by specifying extra links for different vehicle classes/modes, with their class specific cost functions. Some classes require separate links and some classes can use the same links. Park & ride requires extra costs to change mode/class, this can be implemented by adding a special link for the transit part. Discerning different classes of users also requires extra OD-information for the various classes. Each class has its own characteristics and should be modelled differently. These features are especially useful for studying the effect of high occupancy vehicles. The network can have links that are accessible for all classes and links that are only accessible for certain classes.

4.8 Convergence Properties

As stated in the introduction, the main objective of the 3DAS model is to simulate traffic flows as realistically as possible, and to make the model suitable for use in reality. At the same time it is desirable that the underlying model is mathematically sound. During the development of the model it became clear that the main objective could not be reached, when the model was defined as a mathematical program. Several heuristics had to be added to the model to achieve a realistic simulation of traffic flows.

A disadvantage of this approach is that the convergence of the model to a unique solution cannot be proven; however, to make convergence and a unique solution plausible, the model can be simplified and written down as a mathematical program.

The static user equilibrium assignment case is explained first. A simplified form of the 3DAS model is described as a mathematical program in the second part of this section. This program is derived from the mathematical program developed by Janson [38], his description is similar to Hamerslag's original model. To get even closer to the 3DAS model, several of Janson's constraints are relaxed, despite this, the 3DAS model in its present state could not be described as a mathematical program.

It is shown with empirical results that the model indeed converges to the same solution with various different starting conditions, and that the solution represents a dynamic user equilibrium. A small linear network and a network with several alternative routes are used for these tests. The networks are kept fairly small to allow for easier reproduction.

4.8.1 Static User Equilibrium Assignment as a Mathematical Program

Unique solution and convergence are well defined for static assignment problems.

Beckman showed that finding user equilibrium in a network is equivalent to finding the solution of the following mathematical program, known as Beckman's transformation [84]:

The following symbols are used:
A = set of links
q_{rs} = trip rate between origin r and destination s
x_{ij} = flow on link ij
$t_{ij}(x)$ = travel time function for link ij
f_k^{rs} = flow on path k connecting OD-pair r-s
$\delta_{ij,k}^{rs}$ = indicator variable, equal to 1 if link ij is on path k between OD-pair r-s, 0 otherwise.

Beckman's mathematical program is given in equation 39:

$$Min\, z(x) = \sum_{ij \in A} \int_0^{x_{ij}} t_{ij}(\omega)\, d(\omega) \qquad \text{(eq 39)}$$

The objective function $z(x)$ is the sum of the integrals of the link performance functions and can be seen as a mathematical construct that is utilized to solve equilibrium problems [84]. The objective function is subject to several constraints. The first constraint is:

$$\sum_k f_k^{rs} = q_{rs} \qquad \text{(eq 40)}$$

This equation represents the flow conservation constraints. Flow on all paths connecting each OD-pair should be equal to the OD trip rate.
The second constraint is:

$$f_k^{rs} \geq 0 \qquad \text{(eq 41)}$$

This equation ensures that flows are not negative.

The program is subject to the following **definitional** constraint:

$$x_{ij} = \sum_{rs} \sum_k f_k^{rs} \cdot \delta_{ij,k}^{rs} \qquad \text{(eq 42)}$$

This definitional constraint defines the incidence relationship and expresses link flows in terms of path flows. The incidence relationship (e.g., the values δ) is presumed given.
When $t_{ij}(x)$ is strictly convex and increasing, it can be shown that the objective function $z(x)$ is strictly convex, and that the constraints are convex. Thus a unique solution exists to the above program.
To prove that the above program is indeed equal to the user equilibrium, it is sufficient to

derive the first order conditions of the Lagrangian of the equivalent minimization program:

$$L(f, \boldsymbol{u}) = z[x(f)] + \sum_r \sum_s u_{rs}\left(q_{rs} - \sum_k f_k^{rs}\right) \qquad \text{(eq 43)}$$

The dual variable u_{rs} is associated with the flow conservation constraint for OD-pair r-s.
At the stationary point of the Lagrangian, the following conditions (eq 44) have to hold with respect to the path flow variables.

$$f_k^{rs}\frac{\partial L(f, \boldsymbol{u})}{\partial f_k^{rs}} = 0 \qquad \frac{\partial L(f, \boldsymbol{u})}{\partial f_k^{rs}} \geq 0 \qquad \forall\, k, r, s \qquad \text{(eq 44)}$$

The following conditions (eq 45) have to hold with respect to the dual variables:

$$\frac{\partial L(f, \boldsymbol{u})}{\partial u_{rs}} = 0 \qquad \forall\, r, s \qquad \text{(eq 45)}$$

Whenever these partial derivatives are derived, and using the following substitution

$$\sum_{ij} t_{ij}\delta_{ij,k}^{rs} = c_k^{rs} \qquad \text{(eq 46)}$$

where c_k^{rs} represents the travel time on path k from r to s, the following relations will result:

$$\begin{aligned} f_k^{rs}(c_k^{rs} - u_{rs}) &= 0 \\ c_k^{rs} - u_{rs} &\geq 0 \\ \sum_k f_k^{rs} &= q_{rs} \\ f_k^{rs} &\geq 0 \end{aligned} \qquad \text{(eq 47)}$$

The first relation in equation 47 denotes that for every path k either f_k^{rs} is equal to zero or $c_k^{rs} - u_{rs}$ is equal to zero. This means that if f_k^{rs} is not equal to zero, c_k^{rs} is equal to u_{rs} for any k. Which states that on all paths k between r and s, the travel time is equal to the Lagrange multiplier u_{rs}. When f_k^{rs} is equal to zero, then the second relation states that $c_k^{rs} - u_{rs}$ should be greater or equal to zero. This means that the travel time c_k^{rs} on the unused path cannot be smaller than u_{rs}. So all paths from r to s have an equal travel time and all unused paths have a longer travel time. Thus the first and the second relationships state the user equilibrium conditions as defined by Wardrop [97]. The third and the fourth expression just repeat the flow conservation constraints and the non-negativity constraints.

4.8.2 Dynamic Assignment as a Mathematical Program

To expand the above program for dynamic assignment models, Janson described a **dynamic** assignment model as a mathematical program; the dynamic user equilibrium assignment problem [38]. The 3DAS model described earlier has strong similarities with Janson's model, but there are some significant differences that make it difficult to give a mathematical description of the 3DAS model. First the Janson model is explained, then the differences between the two models are described.

4.8.2.1 Janson's Dynamic Traffic Assignment

Some of the notation used by Janson has been changed to remain consistent with the static notation.

The following notation is used:

A = set of all links.
D = set of all periods.
N = set of all nodes
Z = set of all zones.

T = period length
p = period number

$t_{ij}^{p}(x)$ = travel time function for link ij in period p.

x_{ij}^{p} = flow on link ij in period p.

v_{rij}^{dp} = flow on link ij in period p for the trips departing from r in departure period d.

α_{ri}^{dp} = zero-one variable indicating whether trips departing from origin zone r in departure period d reach node i in period p.

q_{rn}^{d} = trip rate between origin r and node n departing in period d.

b_{ri}^{d} = travel time along any path used from origin zone r to node i by trip departing in time period d.

The objective function is defined as:

$$Min\ z(x) \sum_{ij\in A}\sum_{p\in D}\int_{0}^{x_{ij}^{p}} t_{ij}^{p}(\omega)d\omega \qquad \text{(eq 48)}$$

An extra summation is needed over all departure periods $p \in D$, and an extra index p has been added to x and t to denote the proper period.

The objective function is subject to several constraints. The first constraint is:

$$x_{ij}^{p} = \sum_{r\in Z}\sum_{d\le p} v_{rij}^{dp}\,\alpha_{ri}^{dp} \qquad \text{(eq 49)}$$

This constraint states that the total flow on link ij in period p is equal to the sum of flows departing from any origin r in any departure period $d \le p$, reaching node i in period p. The second constraint

$$q^{d}_{rn} = \sum_{p \geq d} \left[\sum_{in \in A} v^{dp}_{rin} \alpha^{dp}_{ri} - \sum_{nj \in A} v^{dp}_{rnj} \alpha^{dp}_{rn} \right] \qquad \text{(eq 50)}$$

states that the traffic departing from r in departure period d going to n should be equal to all the traffic that flows into n minus the traffic that flows out of n, all coming from r, for every period $p \leq d$ in which traffic enters and leaves n.

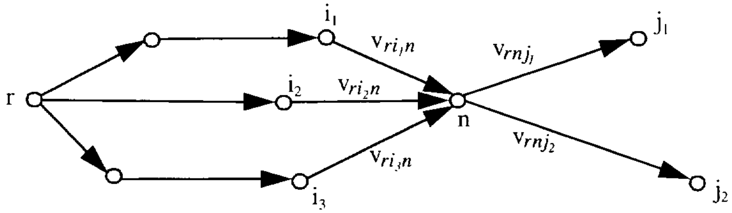


Figure 15: Graphic example of equation 50, the value of q^{d}_{rn} is equal to the sum over every period $p \geq d$ that traffic enters node n that has departed at period d from r, is equal to $(v_{ri_1n} + v_{ri_2n} + v_{ri_3n}) - (v_{rnj_1} + v_{rnj_2})$.

The third constraint (eq 51) defines that all link flows should be greater than zero:

$$v^{dp}_{rij} \geq 0 \qquad \text{(eq 51)}$$

The following **definitional** constraints are defined for the program. Constraint four and five are:

$$\alpha^{dp}_{ri} \in \{0, 1\}$$

$$\sum_{p \in D} \alpha^{dp}_{ri} \leq 1 \qquad \text{(eq 52)}$$

The value of α should be either one or zero, α indicates whether the traffic departing from r in departure period d, reaches node i in period p. Since traffic can reach each node only once, there is only one α equal to one for all p.

If the parameter b^{d}_{ri} denotes the travel time for the traffic departing in period d going from r to i, using any path, then the sixth constraint (eq 53) states:

$$(b^{d}_{rj} - b^{d}_{ri}) \times \alpha^{dp}_{ri} = (t^{p}_{ij}(x^{p}_{ij}) \times \alpha^{dp}_{ri}) + \tau^{dp}_{rij} \qquad \text{(eq 53)}$$

This constraint ensures temporally continuous paths. It states that the travel time for the traffic departing in period d going from r to j minus the travel time for the traffic going from r to i, i.e. the travel time on link ij, should be equal to the travel time $t^{p}_{ij}(x)$ on link ij resulting from the flow x on ij. The parameter τ is required for the equivalence proof and has no further meaning.

The seventh and eighth constraints (eq 54) in which T denotes the period length; are

$$
\begin{aligned}
&[b_{ri}^{d} - pT]\,\alpha_{ri}^{dp} \le 0 \\
&[b_{ri}^{d} - (p-1)T]\,\alpha_{ri}^{dp} \ge 0
\end{aligned}
\qquad \text{(eq 54)}
$$

These two constraints ensure that the correct periods p are chosen, i.e., that α is equal to one in the correct period.

When the following example applies, α is equal to one in the rectangles represented in gray.

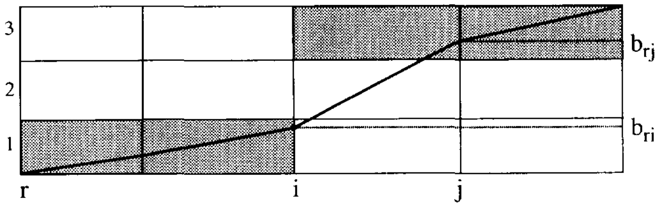


Figure 16: Graphic example of equation 54, the value for α_{ri} is equal to one in the first period, the value for α_{rj} is equal to one in the third period, so $\alpha_{ri}^{11} = 1$ and $\alpha_{ri}^{13} = 1$.

The last two definitional constraints, 9 & 10, are necessary for the equivalence proof, and have no further meaning.

$$
\begin{aligned}
&\tau_{rij}^{dp} \cdot v_{rij}^{dp} = 0 \\
&\tau_{rij}^{dp} \ge 0
\end{aligned}
\qquad \text{(eq 55)}
$$

Janson proves equivalence to user equilibrium using the first order conditions of the Lagrangian (eq 56).

$$
\begin{aligned}
L(x, v, \lambda, \mu, \tau) = {} & \sum_{ij \in A} \sum_{p \in D} \int_{0}^{x_{ij}^{p}} t_{ij}^{p}(\omega)\, d\omega \\
& + \sum_{ij \in A} \sum_{p \in D} \lambda_{ij}^{p} \left(x_{ij}^{p} - \sum_{r \in Z} \sum_{d \in D} v_{rij}^{dp} \alpha_{ri}^{dp} \right) \\
& + \sum_{r \in Z} \sum_{n \in N} \sum_{d \in D} \mu_{ij}^{d} \left(q_{rn}^{d} - \sum_{p \ge d} \left(\sum_{in \in A} v_{rin}^{dp} \alpha_{ri}^{dp} - \sum_{nj \in A} v_{rnj}^{dp} \alpha_{rn}^{dp} \right) \right) \\
& + \sum_{r \in Z} \sum_{n \in N} \sum_{d \in D} \sum_{p \in D} \tau_{ij}^{d} (-v_{rij}^{dp})
\end{aligned}
\qquad \text{(eq 56)}
$$

Deriving the first order conditions of this Lagrangian results in the following three relations:

$$\frac{dL}{dx_{ij}^{p}} = t_{ij}^{p}(x_{ij}^{p}) = \lambda_{ij}^{p}$$

$$\frac{dL}{dv_{rij}^{dp}} = (\mu_{rj}^{d} - \mu_{ri}^{d})\,\alpha_{ri}^{dp} = \lambda_{ij}^{p}\alpha_{ri}^{dp} + \tau_{rij}^{dp} \qquad \text{(eq 57)}$$

$$\tau_{rij}^{dp}\,v_{rij}^{dt} = 0$$

The first relation states that the optimal solution has a unique travel time function for each link ij in each interval p. The next two relations state that, for any given pair of nodes ij, all used paths from a given origin r and a given departure period d have the same travel time. Any unused path between these nodes cannot have a lower travel time. This last result is achieved by the introduction of τ in the program. It ensures that, when a path is not used, the travel time on this path is always larger.

4.8.2.2 3DAS Dynamic Traffic Assignment

The difference between the program described by Janson and the 3DAS model is primarily based on the difference in the link use parameters (α). In the 3DAS model α does not have to be equal to one or zero, but can be a fraction.

If the same example given in figure 16 is taken, trips will be assigned to the gray rectangles according to the values given in the squares of figure 17.

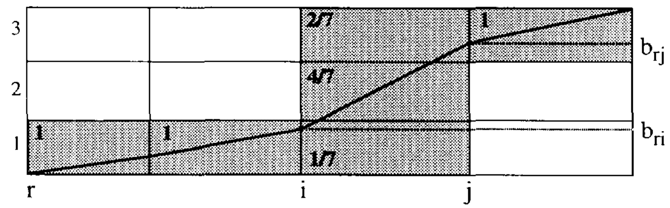


Figure 17: Graphical example of using a fractional α

So on link ij, traffic is partitioned over three different periods. This method gives a more realistic result, but causes a lot of difficulties in the mathematical program.

If we assume that the travel time on a link is always equal to a period or to multiple periods, the mathematical program can be adjusted.

The fourth constraint (equation 52) changes to:

$$\alpha_{rij}^{dp} \in [0, 1] \qquad \text{(eq 58)}$$

which states that α can take any value from zero to one.

The 7th and the 8th constraint (in equation 54) will change to:

$$[b_{rj}^{d} - p\Delta p]\, \alpha_{ri}^{dp} \geq 0$$
$$[b_{rj}^{d} - (b_{rj}^{d} - b_{ri}^{d}) - p\Delta p]\, \alpha_{ri}^{dp} \leq 0 \qquad \text{(eq 59)}$$

The following constraint (8a, eq 60) must be added:

$$\alpha_{rij}^{dp} = \frac{t_{ij}^{p}(x_{ij}^{p})}{\Delta p} \qquad \text{(eq 60)}$$

In this case, the same solution to the first order derivatives of the Lagrangian applies, thus the program also satisfies a dynamic user equilibrium.

The disadvantage of this program is the constraint that the travel time on a link is always equal to a period length or multiple period lengths. In 3DAS this is not the case. If this assumption is relaxed it becomes impossible to give a formulation of α (constraint 8a, eq 60), because traffic can arrive at any time during a period.

4.8.3 Empirical Demonstration of Convergence

To make convergence plausible, several tests are done. The duality gap is a measure for convergence and is defined by:

$$DG = 100\% \frac{\sum_{p}\sum_{ij} x_{ij}^{p}\, t_{ij}^{p}(x_{ij}^{p}) - \sum_{r}\sum_{s}\sum_{d} q_{rs}^{d}\zeta_{rs}^{d}}{\sum_{r}\sum_{s}\sum_{d} q_{rs}^{d}\zeta_{rs}^{d}} \qquad \text{(eq 61)}$$

Where x_{ij}^{p} represents the traffic flow on link ij for period p and $t_{ij}^{p}(x_{ij}^{p})$ is the travel time function for link a for period p, q_{rs}^{d} is the OD value from r to s departing in period d, and ζ_{rs}^{d} represents the travel time for the path from r to s departing in period d. The first term in the numerator represents the travel time used by all travellers as the sum over the travel times used in all links separately. The second term also represents the travel time used by all travellers, but now as the sum of the travel times between all OD-pairs.
At user equilibrium, the duality gap is equal to zero. Whether this equilibrium is unique has not been proven. The solution converged to the same results, in all the convergence tests performed during the research.

Two different networks are used to demonstrate convergence. The first network is a simple network with no alternative routes. The network has ten links of one kilometre and is shown in figure 18. There is one origin O and one destination D.

Figure 18: Simple linear network of ten links, three lanes merge to two lanes at the sixth node.

The network is also used in section 5.4.1. The exact definition of the data is given in appendix C. Several different initial travel times for each link were tested, and in all cases the solution converged to exactly the same values.

The plot given in figure 19 demonstrates the number of iterations needed for the algorithm to converge for the simple "bottleneck" network of figure 18. The duality gap is plotted against the number of iterations for three randomly generated different initial travel times.

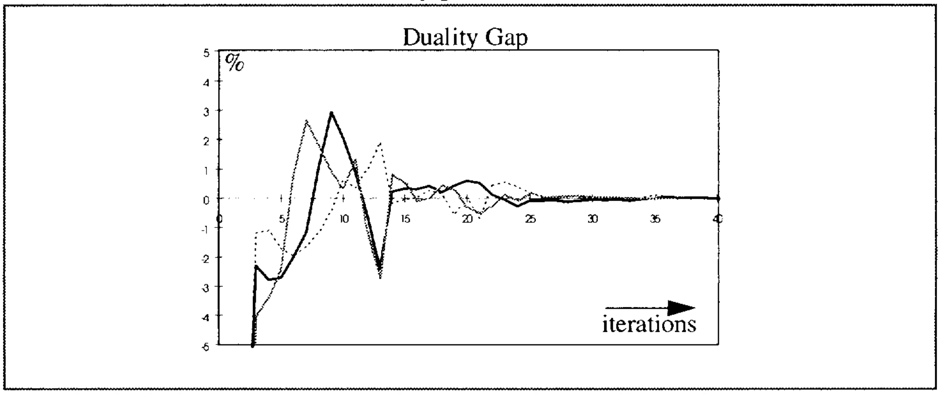


Figure 19: The duality gap as a measure of convergence for a small network

The duality gap converged to zero for all three start situations and the link volumes were equal for all three end solutions. The solution was found after an average of thirty iterations.

The second network is a network with various alternative routes. The network is given in figure 20. Traffic only travels from left to right, and there are three origins O1, O2 and O3 and two destinations D1 and D2. The departure time functions for all OD-relations are equal. The exact definition of the data is described in appendix C.

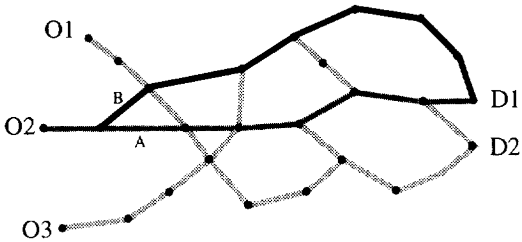


Figure 20: A simple network with various alternative routes.

Figure 21 shows the duality gap for three different initial travel times for this network. In all cases the solution converged to exactly the same solution. The value of the duality gap did not converge exactly to zero, but to 0.2%. This deviation is caused by rounding errors.

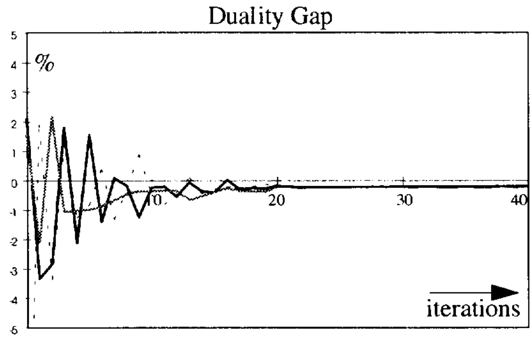


Figure 21: The duality gap as a measure of convergence for a network with alternative routes

An average of thirty iterations was enough to reach convergence.

In the network given in figure 20 two or more routes are possible for each OD-relation. The traffic was divided over two routes for OD-relation O2-D1. In figure 20 these routes are marked as a 'striped' route and a 'solid black' route. The first link on both routes is the same. At user equilibrium, the path travel times on these routes should be equal for each departure period (when both routes are used). Figure 22 shows the path-travel time needed to travel from O2 to D1 at different departure times. The x-axis shows the departure period and the y-axis shows travel time in minutes. The figure shows that initially the 'striped' route is shorter for trips that depart during the first 10 periods. Travel time on both routes increases due to increasing demand. Between the 11th and the 24th period the travel time on the striped route equals the travel time on the black route. At this point both routes are equal and a certain amount of trips take the black route. The amount of trips now present on the solid route keeps both travel times equal. Travel time on both routes still increases. After the 18th period travel time decreases and for trips that depart later than the 24th period, the striped route again becomes shorter, and the solid route is no longer chosen.

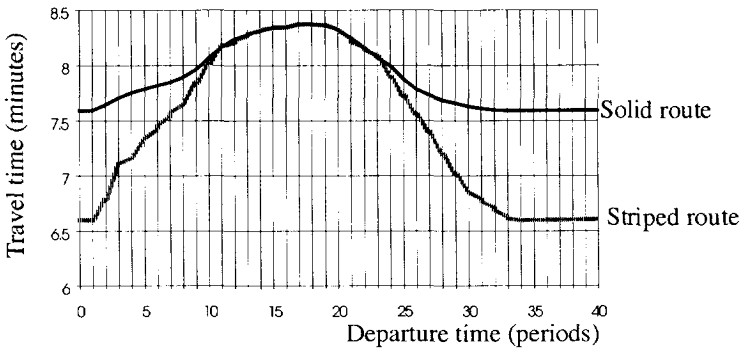


Figure 22: Travel time on two different routes for OD-pair O2-D1.

According to figure 22 trips should use the solid route between the 10th and the 24th period, that this is indeed the case is demonstrated by figure 23. The figure shows the pattern in time on link A and link B. Link A shows the trips that use the striped route and link B shows the trips that use the solid route. As the figure shows for link B, a number of trips use the solid route between the 10th and the 24th period, during other periods there are no trips. The number of trips is enough to keep the travel times equal.

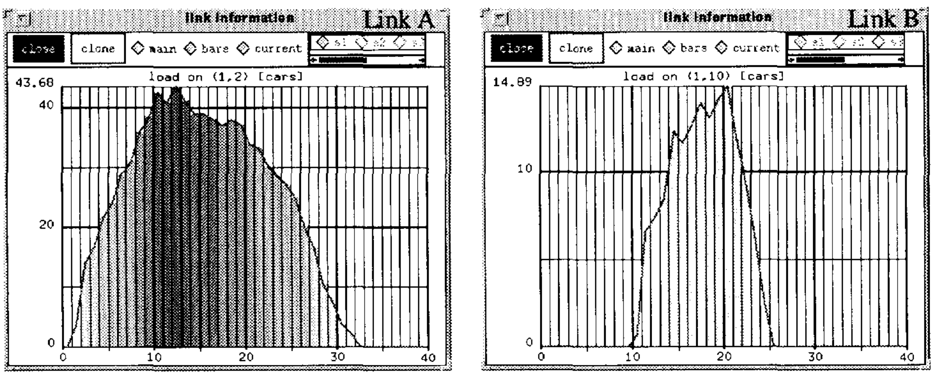


Figure 23: Pattern in time (load) on link A and B

The trips are distributed in such a way that the resulting travel time remains exactly equal on both routes, whenever both routes are used.

4.9 Conclusions

This chapter describes the theory of the 3DAS model. The model was developed as a heuristic model, because a mathematically sound model implied too many restrictions on reality. The model is solved using an iteration based method. Using several examples convergence to a dynamic user equilibrium is made plausible.

Like most other dynamic assignment models, the simulation period is divided into intervals of

equal duration, referred to as periods. The link parameters may be defined separately for each period. Network and travel demand are presumed given.
Two commonly used assumptions were made concerning route choice:

- all travellers are completely informed and take future congestion into account.
- all travellers choose a minimum cost path in time.

This user optimized approach was chosen for the initial theory of the model. Travellers choose their minimum cost path, and take *future* congestion into account. A second strategy is to choose the *instantaneous* shortest path. This path is based on the travel times at the moment of departure or at the moment of a choice, i.e. at a junction. Instantaneous route choice does *not* take future congestion into account. A third strategy, to model a more realistic route choice, is to use *historic* information for route choice.
The user optimal approach is the most complex method and was chosen as the main approach. Instantaneous path finding is described as a simplification, and path finding based on historical information as an extra feature.

As discrete time units are used and for simplicity, the following two assumptions concerning the links were made

- traffic conditions are homogeneous along a link and constant for the duration of each time period.
- the variables of a link, flow (q), density (ρ) and speed (v), in a certain period p are subject to a fixed relation described by:

$$q^p = \rho^p \cdot v^p \qquad \text{(eq 62)}$$

An iterative process is used to solve the model. In every individual iteration, the shortest route for every OD-relation in each departure period is calculated. According to these routes, the OD-matrix is assigned to the network. Before each iteration, new link travel times are calculated for every period, based on the conditions resulting from the previous iteration. The iteration process stops when the stop criterion is reached, i.e., when the loading of the network remains essentially unchanged from one iteration to the next. The model uses density as an explanatory parameter instead of flow, because a dual relationship exists between flow and travel time, while the relation between density and travel time is not. This makes it possible to model a decreasing flow when a critical density has been reached, thus representing congestion.
Four different methods for pathfinding and two different methods for assignment were developed. A method for traffic jam building and a method for merging traffic have been added to the model for a realistic representation of traffic. Several extra features have been appended to the model, such as ramp metering, and rerouting, for DTM applications.
The different methods, and the basics of the model, are compared and tested in chapter 5.

CHAPTER 5

SENSITIVITY AND CALIBRATION

5.1 Introduction

Several different aspects of the 3DAS model are investigated further in this chapter. Firstly, the consequences of discretization on time and space are discussed by investigating the influence of period length and link length. Secondly, the travel time function is discussed, and the preference for a speed-density function rather than a time-flow function is clarified. Several possible speed-density relationships are discussed and how different weather conditions, different locations, and accidents influence the relationship between speed and density. The third section gives a comparison of the 3DAS model and the microscopic simulation model FOSIM [81][82]. A bottleneck and an on-ramp situation are compared. The pathfinding method, the assignment method, the queuing algorithm and the merging algorithm are compared and evaluated using these situations. The calibration results of this comparison are described in the fourth section. The chapter ends with conclusions.

5.2 Consequences of Discretization

The following assumptions were made concerning the links in section 4.1:

- traffic conditions are homogeneous along the length of a link and constant for the duration of a period.
- The variables of a link, flow (q), density (ρ) and speed (v), are subject to a fixed relationship: $q^p = \rho^p \cdot v^p$

These assumptions will influence the accuracy of the calculation. As a consequence of homogeneous conditions along a link, no distinction is made between trips that have just arrived on the link and the trips that are just about to leave the link. All information about the position of trips on the link is lost, and the trips on one link result in a uniform traffic density over the entire length of the link. A similar argument holds for the assumption that traffic conditions are constant for the duration of the period.

Several experiments have been done[1] to quantify these consequences and *difference with actual travel time* was chosen as a measure of error. Travel time t_a^p on a link a in period p is

1. The experiments described in 5.2.1 and 5.2.3 were carried out by P.C.F. Maas and are also described in his masters thesis [49]

given by equation 63,

$$t_a^p = \frac{l_a}{v(\rho_a^p)} \qquad \text{(eq 63)}$$

where l_a is the length of link a, and $v(\rho_a^p)$ is the speed on link a as a function of the density in period p. The unit of t_a^p is 'time'. Travel time on a link can also be expressed in periods, denoted as τ_a^p. The unit becomes *number of periods* (not necessarily an integer value). The relationship is given in equation 64.

$$\tau_a^p = \frac{l_a}{T \cdot v(\rho_a^p)} \qquad \text{(eq 64)}$$

Where T is the period length. From this it can be derived that travel time is influenced by link length l_a, the speed-density function $v(\rho_a^p)$, and period length T.

Several tests are described in the next three sections to quantify the influence of link length, the influence of period length, and the combined influence of both.

5.2.1 Influence of Link Length

The accuracy of an assignment depends on the detail of the network, the speed-density function, and the variation in demand. A linear network can be represented by just one link or by several shorter links. It is easily demonstrated that there is a loss of detail with increasing link lengths. The traffic density on a linear network in a certain period is shown for several link lengths in figure 24. The number of links is denoted by n.

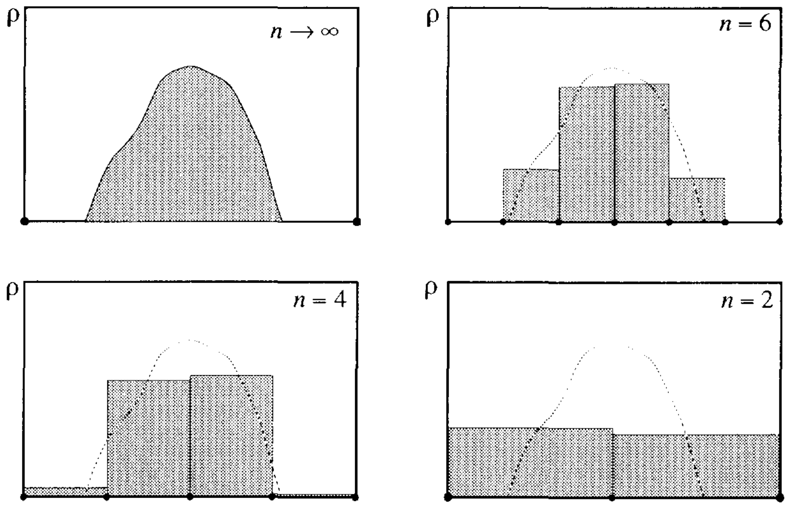


Figure 24: Comparison of a linear network constructed using a decreasing number of links n.

The loss of accuracy depends very much on the *shape* of the demand. When demand is constant, the number of links has no influence on accuracy, inaccuracy becomes worse with a frequently changing demand.

Although the loss of detail in figure 24 is rather significant, there is no loss of detail for down stream links, because each calculated trajectory maintains its correct path through time.

Consider the following example given in figure 25:

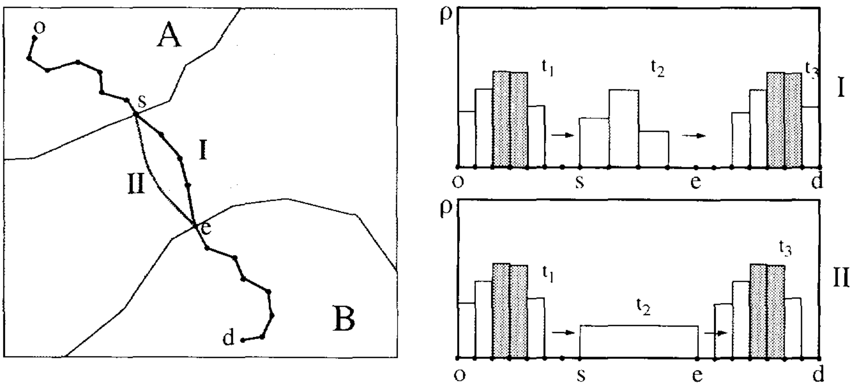


Figure 25: Demonstration of conservation of detail on downstream links; a bridge is represented by four links (I) versus one link (II), detail downstream is maintained, but a difference in travel time on the bridge occurs. The distance between s and e is for both situations equal.

EXAMPLE: On the left side of the figure, a path P_{se} connects two networks areas, A and B. The bridge, from *s* to *e*, is represented by four links for case I and by one link for case II. On the right side of the figure, two plots are given that represent the path from *o* to *d* along the x-axis and the density on each link on the y-axis. In each graph three different time periods are given, denoted by t_1, t_2 and t_3. In the upper graph (I) of figure 25, the bridge from s to e is represented by four links. At t_1 trips cause a certain density distribution on the links before the bridge. At t_2 the trips are on the bridge and at t_3 they have passed the bridge. In the lower graph (II) of figure 25, the bridge is represented with one link. The density distribution on the bridge changes significantly but past the bridge, the original distribution returns. The density distributions beyond the bridge remain unchanged in both cases. As will be explained on page 69, the loss of detail on the bridge can cause a difference in travel time, which results in an earlier arrival for case II.In accordance with the examples given in figures 24 and 25, inaccuracies in travel time increase with longer links and more variable demand. To quantify these effects, an experiment was carried out in which the travel time on a road of length *L* was compared using a different number of links and a triangle shaped demand.

Consider a path P_{se} from node s to node e of length L. The density distribution on this road is triangular. At the beginning of the road density is zero and increases towards $\alpha\rho^{max}$ in the middle of the road and then it decreases linearly back to zero, α is between zero and one. The situation is shown in figure 26. The road consists of n links of length $l = L/n$. Each link is equal, and speed on a link is a function of density. This speed-density function $v(\rho_a^p)$ is given in equation 21 on page 38.

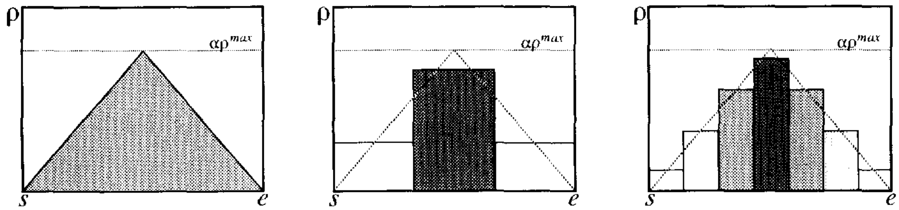


Figure 26: Traffic density on path P_{se}, for the continuous situation and for 3 and 7 links.

The travel time for path P_{se} is equal to the sum of travel times on the consecutive links. The travel time for a link a is derived using equation 21 to compute the speed $v(\rho)$ and equation 63 to determine the travel time t, using the link length l. It is assumed for simplicity that P_{se} is traversed within one period, and the superscript p will be omitted for the next equations.

The travel time for path P_{se} consisting of n links is the sum of the individual travel times, as given in equation 65:

$$t_{se}^{n} = \sum_{a \in P_{se}} \frac{l_a}{v(\rho_a)} = \sum_{a \in P_{se}} \frac{L/n}{v(\rho_a)} \qquad \text{(eq 65)}$$

The *exact* travel time for road P can be calculated by assuming that P is constructed of an infinite number of links. Assume P has n links, and $n \to \infty$, then equation 65 becomes an integral (eq 66):

$$t_{se}^{\infty} = \lim_{n \to \infty} t_{se}^{n} = \lim_{n \to \infty} \sum_{a \in P_{se}} \frac{L/n}{v(\rho_a)} = \int_0^L \frac{dx}{v(\rho(x))} \qquad \text{(eq 66)}$$

Equation 65 (with $n = 1,3,7,15$ *and* 31) and equation 66($n \to \infty$) are plotted in figure 27. The calculations are described in detail in appendix D

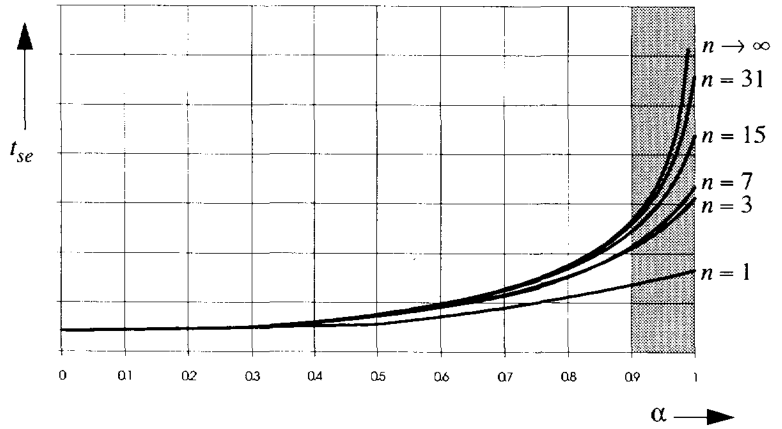


Figure 27: Travel times for path P_{se} consisting of 1,3,7,15, and 31 links, the exact travel time is labelled $n \to \infty$.

The line labelled $n \to \infty$ represents the exact travel time for road P, according to the assumptions[1]. All other lines represent the travel time for road P when the road is constructed with n links. Figure 27 shows that travel time is always undervalued.

The relative error in travel time can be calculated by using equations 65 and 66. The following quantitative criterion is used:

$$\Delta_n^\infty = (t_{se}^\infty - t_{se}^n) / t_{se}^\infty \qquad \text{(eq 67)}$$

Equation 67 is plotted in figure 28 using the same n as in figure 27

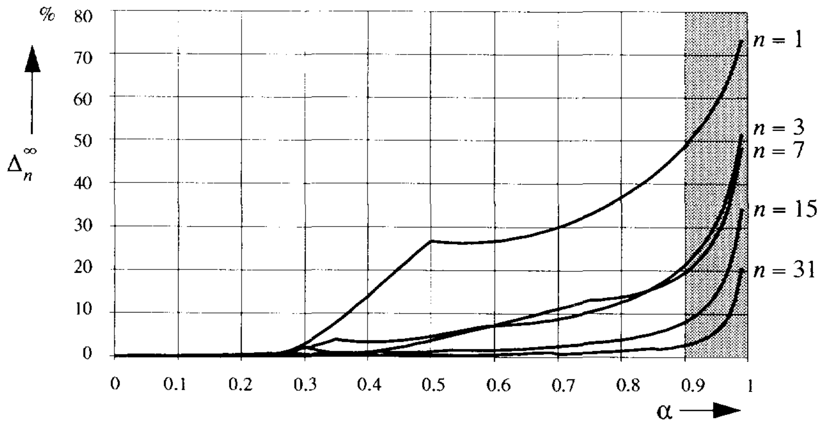


Figure 28: Relative error (%) in the travel times om P_{se} consisting of 1,3,7,15, and 31.

1. The assumed travel time function takes an infinite travel time at maximum density, which is unrealistic. It is because of this assumption that travel times differ considerably for $\alpha > 0.9$. In reality density hardly ever reaches such high values. In the 3DAS model such values are impossible due to the queuing algorithm.

Figure 28 shows that, when the density is below the critical density, the error is very small, the line shows a discontinuity for $n = 1$ at $\alpha = 0.5$. At this point density at the link exceeds the critical density. For $n = 3$ this occurs at $\alpha = 0.3$ for the link in the middle and at $\alpha = 0.75$ for the first and the third link. The discontinuity occurs so many times for $n = 31$ that it is hardly noticeable in figure 28.
The discontinuities can sometimes give strange results. In figure 28 for $n = 3$ and $n = 7$ the lines cross each other three times. This means that sometimes three links are more accurate then seven links. If the speed-density function were continuous, or the density distribution of a different shape, this would not occur.

Although the differences in travel time seem significant in figure 27 and 28, they are not likely to give serious problems with real networks. In general the detail of the network should be adopted to the *level of variance* in the demand. When there is a high level of variance in the demand, long links should be avoided. When the link length is able to capture the fluctuations in the demand there will be little loss of accuracy.
When link lengths vary considerably, it is advisable to divide long links into several short ones. In general the congested parts in the network with high densities ($\alpha > 0.5$) are the most important parts and should be modelled using short links for higher accuracy.

5.2.2 Influence of Period Length

An approach similar to that given in the last section can be taken to quantify the influence of period length. A total period of length T can be divided into several periods. The length of the period has an influence on the accuracy of the calculation.
Consider a link a with length L and a total time span T. This total time span can be divided into several periods with length T_{per}. The time to traverse link a depends on density ρ ($0 \leq \rho \leq \rho^{max}$) and the speed-density function $v(\rho_a)$. If the total time period T is represented by one period, then the travel time for a is equal to $L/v(\rho_a)$, see equation 63. When the number of periods within T are increased and it is assumed that all trips depart during the first half of the time span T, and the time to traverses link a is equal to $\frac{1}{2}T$, then the increase in travel time, denoted Δt, for link a is demonstrated in figure 29 for the trajectory assignment method. Three different period lengths are shown. The periods are given along the x-axis and Δt along the y-axis.

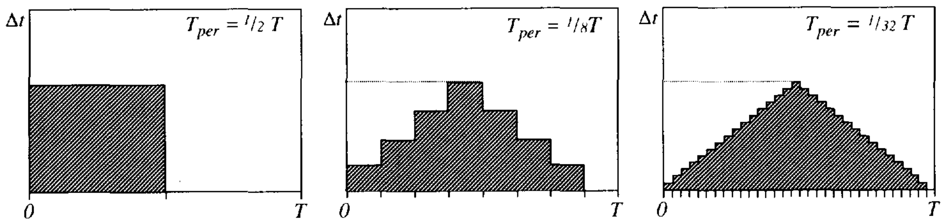


Figure 29: Influence of period length on travel time on one link using the trajectory assignment method, period length is equal to ${}^1/_2\,T$, ${}^1/_8\,T$ and ${}^1/_{32}\,T$.

The first plot shows that there is only an increase in travel time during the first half of the total time span. The second and the third plot show that the increase in travel time also affects the second half of the time span.

Figure 30 depicts the same example with the *surface* assignment.

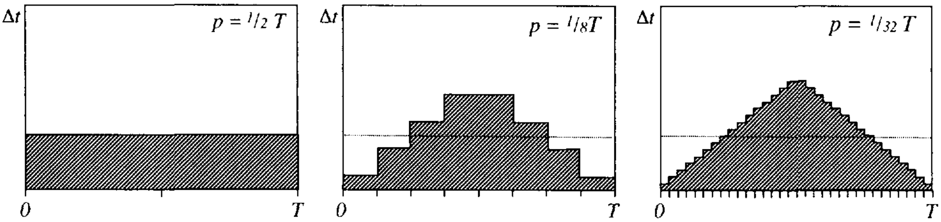


Figure 30: Influence of period length on travel time on one link using the surface assignment method, the period length is equal to $^1/_2\,T$, $^1/_8T$ and $^1/_{32}\,T$

Figure 30 shows that the increase in travel time is better distributed over the total time span using the surface assignment method. The change in travel time is constant for each individual period.

5.2.3 Combination of Link Length and Period Length

Another way of analysing the influence of link length and period length on assignment results is to look at the trajectories that are used to assign trips to the network.

Consider a path P_{se} consisting of identical links, and n trips departing in the first period from s to e. The trajectory assignment method assumes that all trips leave at the start of the period. Depending on experienced travel time, a certain trajectory will be followed. In principle three different situations are possible:

- travel time is exactly equal to the period length.
- travel time is larger then the period length.
- travel time is smaller then the period length.

The three possible trajectories are shown in figure 31 as solid lines. Trips that depart in the second and third period are shown as dashed lines.

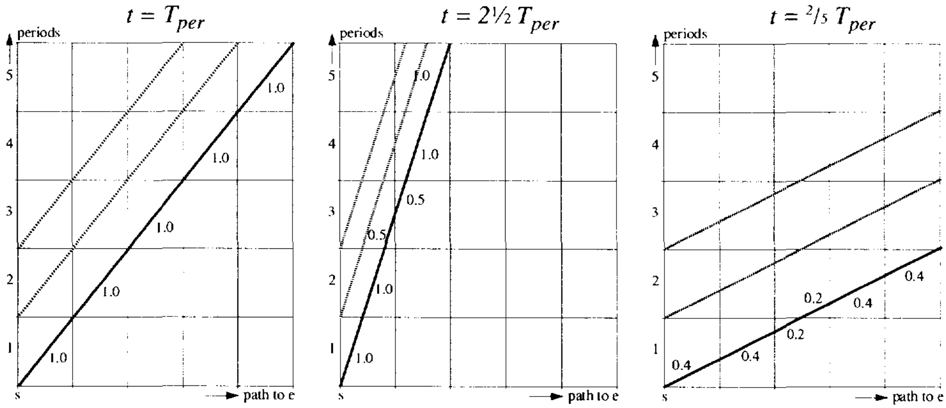


Figure 31: Trajectory assignment, three possible trajectories.

The numbers by the trajectories in the plots of figure 31 show the fraction of trips that are assigned to a link in a period. When $t = T_{per}$ the trips traverse the entire link in a period and the fraction assigned to each link is 1.0. When $t=2\frac{1}{2}\ T_{per}$ the trips cover only $^{2}/_{5}$ths of the link, which means that on this part of the link 100% of the trips are present. The trips are assigned to the entire link because of the assumption that traffic conditions are homogeneous along a link. This results in a density that is, for the first $^{2}/_{5}$ths of the link, 2½ times too low.
When $t = {}^{2}/_{5}\ T_{per}$ a different situation occurs. Only 40% of the trips are assigned to the first link, which is correct. At the third link, however, an error occurs because only 20% of the trips are assigned to this link in the first period. As 40% is actually present on the link, it covers only the first half of the link in the first period and the second half in the second period. As the conditions are homogeneous, the trips are spread over the entire length of the link, resulting in only 20% on the third link for each period.

The dashed lines in figure 31 show the trajectories of the trips which leave at the second and the third period. These trajectories compensate exactly for the errors just described. The situation for $t=2\frac{1}{2}\ T_{per}$ on the first link and the third period shows that the entire link is covered. The first $^{2}/_{5}$ths of the link are covered by the trajectory departing in the third period, the second $^{2}/_{5}$ths of the link are covered by the trajectory departing in the second period, and the last $^{1}/_{5}$th of the link is covered by the trajectory departing in the first period. In practice, with multiple OD-relations, several trajectories will cross a link in a period, thus correcting any misrepresentations.

5.3 Influence of the Speed-Density Function

Travel time functions or delay functions are an important part of assignment models, they express travel time on a link as a function of traffic flow, and they specify the effect of road capacity on travel times and route choice. Several of the graphs in the previous sections were influenced by the assumed speed-density function. Many different types of delay functions have been used in the past, and it would appear that for almost every major transport study, a new and different delay function is proposed, for a review see Branston 1976 [8] and Suh 1990 [87].
Two main approaches can be identified in the development of travel time functions; the *mathematical* approach and the *theoretical* approach:

- In the *mathematical* approach a simple mathematical function is devised to replicate the observed data. The function usually does *not* contain parameters which are directly based on network and/or link characteristics, such as signal settings.
- In the *theoretical* approach link characteristics are usually well represented and the function is normally based on queuing theory.

This section starts with a short review of travel time functions and discusses the use of a speed-density relationship versus a traveltime-flow relationship. The second section demonstrates how the relationship changes with different locations, different weather conditions, different points in time, and accidents. A suitable relationship is selected for 3DAS based on these sections. The section finishes with table 4 that gives an indication for speed-density relationships for particular types of road.

5.3.1 Travel Time Functions: A Short Review

A short overview of several common travel time functions is given in this section. First two functions are described based on a traveltime-flow relationship: the BPR function which adopts the mathematical approach, and the Davidson function, which adopts the theoretical approach. Secondly the disadvantages of a traveltime-flow relationship for dynamic traffic assignment are discussed. Thirdly two functions are described based on a speed-density relationship: the Smulders function, which adopts the mathematical approach, and a function introduced by B.Ran, which adopts the theoretical approach.

BPR function.

The most widely used function, based on the *mathematical* approach, is the BPR-function (Bureau of Public Roads 1964 [16]). Its polynomial form is easy to implement in computational procedures. The function is written as a traveltime-flow function and is given in equation 68,

$$t = t_0 \cdot \left[1 + \beta \cdot \left(\frac{q}{c}\right)^n\right] \qquad \text{(eq 68)}$$

where t is the travel time on a link, t_0 free flow travel time, q flow, and c capacity. β and n are parameters of the function. Different combinations of β and n are used, depending on the application.

Since 3DAS uses density as explanatory variable, the function had to be rewritten as speed-density function. Unfortunately it is impossible to derive a direct analytical transformation of the BPR function into a speed-density function. It is possible, however, [71] to derive an inverse speed-density function by using the fundamental hydrodynamic relationship describing the flow q in equation 69:

$$q = \rho \cdot v \qquad \text{(eq 69)}$$

Where q is flow, ρ density and v speed. The resulting density-speed function is given in equation 70:

$$\rho = \frac{c}{v} \cdot \left(\frac{1}{\beta} \cdot \frac{v_0 - v}{v}\right)^{\frac{1}{n}} \qquad \text{(eq 70)}$$

This function makes it possible to plot the function with its related functions (figure 32).

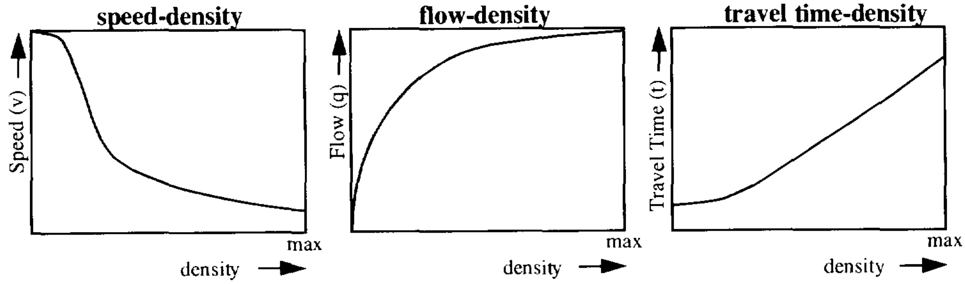


Figure 32: BPR and related functions

The BPR function has four parameters: c, n, β and v_0.

Davidson function

The Davidson function is a well known function based on the *theoretical* approach [21]. This function introduces a parameter for type of road and is given in equation 71:

$$t = t_0 \cdot \frac{1 - m\frac{q}{c}}{1 - \frac{q}{c}} \qquad \text{(eq 71)}$$

Where t_0 represents free flow travel time on a link, t average travel time on a link, m a model parameter expressing 'quality of service' (value between 0.8 an d 1.0), q flow, and c capacity.

Equation 71 can be rewritten as a speed-density function using the hydrodynamic relationship of equation 69 [71]. This results in equation 72:

$$v = \frac{c + (\rho v_0) - \sqrt{c + (\rho \cdot v_0)^2 - 4m\rho v_0 c}}{2m\rho} \qquad \text{(eq 72)}$$

Where v_0 represents free flow speed and ρ density.

The Davidson function has three parameters: t_0, m and c. The shape of the Davidson function and some related functions are given in figure 33.

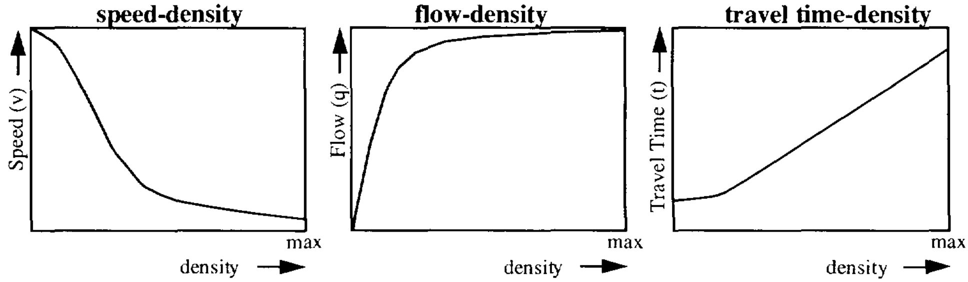


Figure 33: Davidson speed-density and related functions

Disadvantages of traveltime-flow relationships

Almost all the delay functions used for transport studies satisfy the following condition:

$f(x) = strictly\ increasing.$

This is a necessary condition for the assignment to converge to a unique solution.

It is a well known fact however, that the relationship between flow and travel time is *not* a strictly increasing function but a dual function. As density increases on a link, travel time and flow also increases, but at a certain critical density, flow *decreases* while travel time continues to increase. This effect does not appear in the flow-density relationship in figure 32 and 33. The flow still increases with increasing density. In reality the flow should decrease after a certain level of density is reached, because maximum density represents a no-motion traffic jam with flow zero.

The reasons for using a strictly increasing function are, apart for convergence reasons, given by Rothrock and Keefer in 1957 [78]. They pointed out that for data collected with a 6 minute interval there is indeed a dual relationship between flow and travel time. With an increasing sample interval however this effect becomes smaller, and with a sixty minute interval it vanishes almost completely. The use of a strictly increasing function for static assignment models seems plausible as they usually consider a total time interval of a peak hour to a complete day.

Rothrock and Keefer's experiment was replicated using the available data from the Amsterdam beltway, the first three graphs in figure 34 show a similar observation, although not as clearly as in the original tests done in 1957.

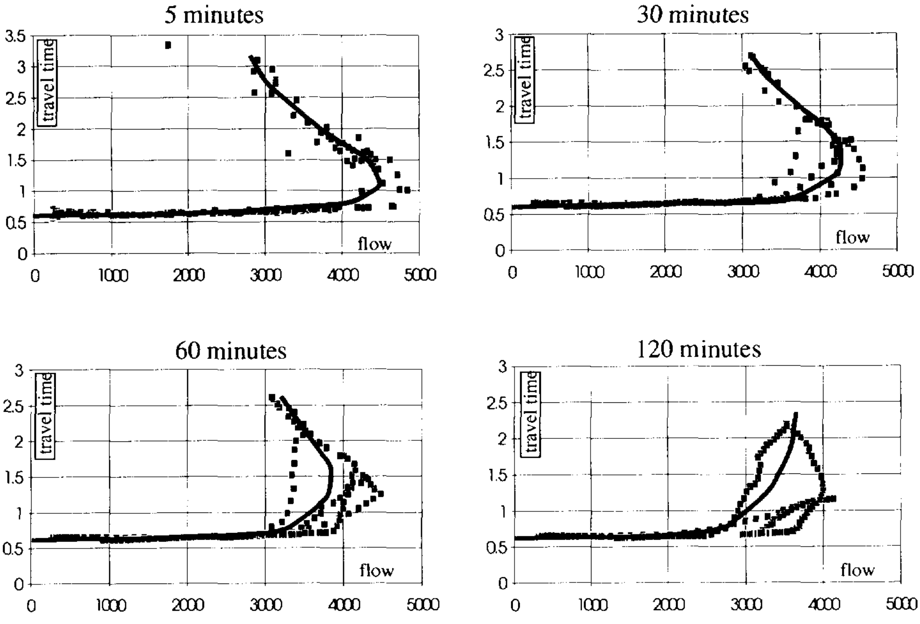


Figure 34: Relationship between flow and travel time with increasing sample interval, as moving average.

The reason the *turning back* of the relationship is still present with a sample interval of 60 minutes can be easily explained. Nowadays the rush hour lasts for more than one hour. Speeds drop at certain locations to less then 50 km/hr for a duration of two or three hours. When the sample interval is further increased to 120 minutes *turning back* is no longer observed, as can be seen in the last graph of figure 34.

With the introduction of dynamic assignment models however, justification of a long time span is no longer valid because dynamic models divide the total time interval into equal periods of typically five to fifteen minutes.

Another justification for using a strictly increasing function is given by Rose et. al. [77]. They distinguish *interrupted* and *uninterrupted* traffic flow conditions. Uninterrupted flow conditions exist where trips, traversing a roadway, are not impeded by any causes external to the traffic stream, such as signs or signals, although vehicles may be stationary due to causes internal to the traffic stream. Interrupted flow conditions exist where flow *is* impeded by external causes. Therefore, freeway systems provide uninterrupted flow conditions, while arterial roads are characterized as providing interrupted flow conditions. The relationship between travel time and flow takes a different form when it is necessary to account for the effect of flow interruptions. These relationships are shown in figure 35.

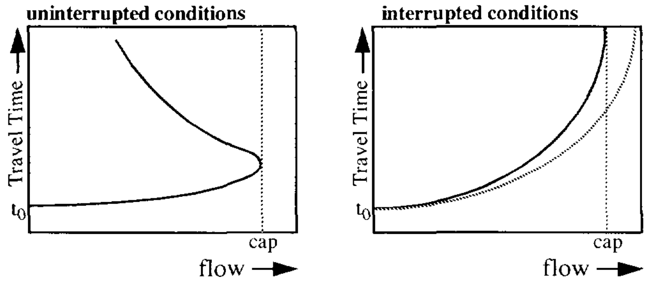


Figure 35: Flow-traveltime relationships for interrupted and uninterrupted conditions, the dashed line is usually used to overcome problems with infinite travel times.

Only when a network is entirely constructed of arterial roads is the use of a strictly increasing travel time function defensible. A strictly increasing travel time function is not justifiable for assignment studies with time spans shorter than one hour and dealing with freeways.

Many dynamic assignment models have recognized the problems with the traditional travel time functions and have adopted different travel time functions. The problem with duality is easily solved by using density on a link as explanatory variable instead of flow. The observed relationship between density and travel time *is* a strictly increasing relationship. By using this function flow and travel time increase with increasing density, and at a certain critical density flow decreases while travel time continues to increase. The basic relationships between density, flow and travel time are all satisfied.

Smulders function

The Smulders function [86] was used for the 3DAS model. This function was developed according to the *mathematical* approach. The formulation is written as a speed-density function and given in equation 73.

Speed v is given by two functions of the density ρ.

$$v(\rho) = \begin{cases} v^{max} \cdot \left(1 - \dfrac{\rho}{\rho^{max}}\right) & 0 < \rho < \rho^{crit} \\ v^{max} \rho^{crit} \cdot \left(\dfrac{1}{\rho} - \dfrac{1}{\rho^{max}}\right) & \rho^{crit} < \rho < \rho^{max} \end{cases} \qquad \text{(eq 73)}$$

Where v^{max} is the free flow speed, ρ^{crit} critical density and ρ^{max} maximum density. Maximum density represents a no-motion traffic jam. When the critical density is reached speeds drop more rapidly and the corresponding flow decreases.

The Smulders function and some related functions are displayed in figure 36.

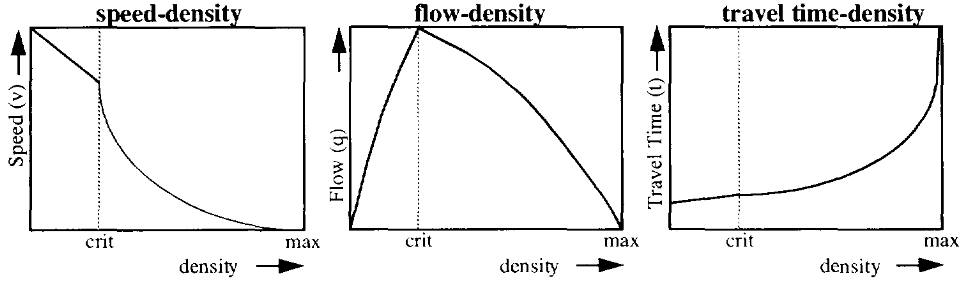


Figure 36: Smulders speed-density function and related functions

The difference between the Smulders function and the BPR and Davidson functions shows most clearly in the flow-density graph of figure 36. After a certain critical density is reached flow drops, representing the start of congestion due to decreasing flow.

To get more influence on the shape of this function two extra parameters were introduced for this function in 3DAS. The parameter α which influences the steepness of the first linear part of the function, and a parameter β which influences the curve of the second part of the function. The equation is given in equation 74:

$$v(\rho) = \begin{cases} v^{max} \cdot \left(1 - \dfrac{\alpha\rho}{\rho^{max}}\right) & 0 < \rho < \rho^{crit} \\ \phi \cdot \left(\dfrac{1}{\rho} - \dfrac{1}{\rho^{max}}\right)^{\beta} & \rho^{crit} < \rho < \rho^{max} \end{cases} \qquad \text{(eq 74)}$$

In total the modified Smulders function has five parameters: α, β, ρ^{max}, ρ^{crit} and v^{max} .

Ran's function

Bin Ran et. al. [73] have developed a travel time function according to the *theoretical* approach. Two different sets of functions are recommended. One function for longer time intervals (5-30 minutes) and one for shorter time intervals (1-5 minutes). In both parts the function is expressed as the sum of two components. The link is for this purpose divided into two parts, a flowing part and a queuing part. So travel time is the sum of a flow dependent cruising time D_1 for the first part of the link and a queuing delay D_2 for the second part of the link. See equation 75, in which τ represents the travel time for a link.

$$\tau = D_1 + D_2 \qquad \text{(eq 75)}$$

The function to express D_1 is a linear speed-density relationship, derived from Greenshields formula (1933). The function is rewritten to cope with the length of the queue which may be present on the second part of the link. See equation 76.

$$D_1 = 3600\left(\frac{l - x_2/e_m}{w}\right) \qquad \text{(eq 76)}$$

Where l is the length of the link and x_2 represents the number of vehicles present on the second (queuing) part of the link, e_m is maximum density and w the cruising speed for inflow on the link.

In the second part of equation 75, the term D_2, is expressed as the sum of two delay functions. See equation 77.

$$D_2 = d_1 + d_2 \qquad \text{(eq 77)}$$

Where d_l expresses a non-random delay due to signal cycle effects and d_2 expresses an overflow effect due to random arrival effects and oversaturation delays. A difference is made for long term and short term applications for d_2. For further details see Ran 1992 [73], as a discussion on queuing theory and signalled intersections is out of the scope of this dissertation.

The first part of Ran's travel time formula is visualized as a speed-density function in figure 37.

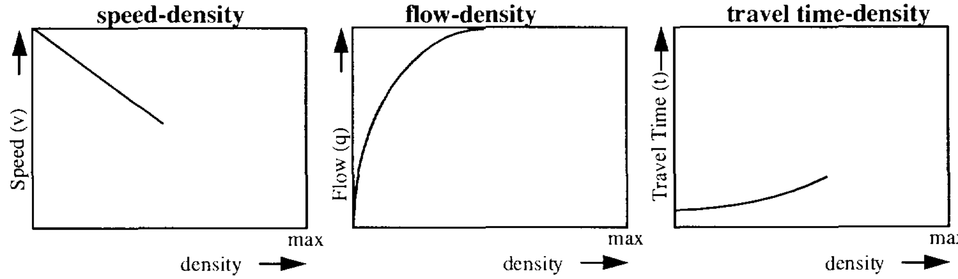


Figure 37: The flow dependent part of Ran's travel time function

Ran's function is meant primarily for arterial networks with signalled intersections.

5.3.2 Influences of Location, Weather, Time of Day and Incidents.

It can be concluded from the previous section that a *traveltime-flow* function cannot describe a congested state in a network, and that a *speed-density* function is better suited to dynamic traffic assignment. The correct shape of the speed-density function, however, is subject to much discussion. This section will describe how location, weather, time of day and incidents influence the shape of the speed-density relationship.

Data from the Amsterdam beltway was used for this research. One hundred locations were measured using induction loops during two weeks in December 1992. The data was aggregated to one minute. Weather conditions were retrieved from weather reports given for Schiphol International Airport. Four days in December 1992 were chosen from this data because these data sets contained few errors and there were several traffic jams during the days. The weather conditions for the days are listed in table 1.

Table 1: Weather conditions

Day	Wind (knots)	Rain (time)	Visibility	Sky
3 Dec. 1992 (Tuesday)	18	16:00 - 18:00 22:00 - 24:00	good > 10 km	cloudy
7 Dec. 1992 (Monday)	15	01:00 - 12:00 15:00 - 16:00	good > 10 km	half cloudy
8 Dec. 1992 (Tuesday)	5	-	04:00 - 07:00 fog 11:00 - 21:00 fog visibility 100-200m	cloudy
9 Dec. 1992 (Wednesday)	6	-	good > 10 km	cloudy

Ten locations were chosen from each day's data set. The locations were selected on variability of measurements during the day. To investigate the functions, measurements for all stages of traffic flow are most useful, a location with only free flow traffic is not useful for investigating the relationship between speed and density.

The locations were chosen to represent different situations, i.e. merge sections, off-ramps and tunnels. The same locations were used for each day. The chosen locations are listed in table 2. and are indicated by a loop number. These loop numbers can be found in appendix A.

Table 2: Induction loop locations

nr	loop nr	location	lanes	remarks
1	75	A16 Utrecht -Amsterdam	3	
2	232	A4 The Hague - Amsterdam	2	
3	108a	A8 Alkmaar - Amsterdam	2	entering the Coen tunnel
4	105	A8 Alkmaar - Amsterdam	3	
5	160	A4 NieuweMeer - Amstel	3	near on-ramp
6	157	A4 NieuweMeer - Amstel	4	merging area
7	204	A19 NieuweMeer	3	fly-over
8	200a	A19 Coenplein - NieuweMeer	2	exit ramp
9	196a	A19 Coenplein - NieuweMeer	1	exit ramp
10	195	A19 Coenplein - NieuweMeer	3	

This section will illustrate, how external factors, such as location, weather, time of day and incidents, influence the relationship between speed and density.

Influence of location

Four speed-density graphs are shown in figure 38 for four very different locations to illustrate the influence of location. In each graph the location is numbered according to table 2. The measurements were taken at a 5 minute interval for one complete day (24 hours).

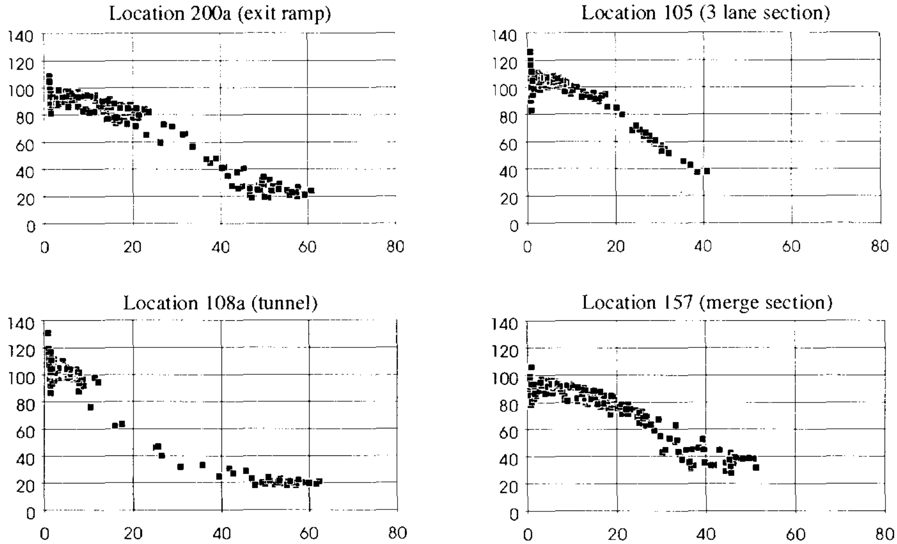


Figure 38: Speed-density graphs for the same day at four different locations, density has been corrected for the number of lanes.

The graphs in figure 38 show that location has significant influence on the relationship between speed and density. Location 200a is an exit with two lanes that shows a fluctuating relationship between speed and density. Location 105 is a 3 lane freeway section that has a normal relationship. Location 108a is a two lane section entering a tunnel, speed drops very quickly with increasing density. Location 157 is a four lane merging section, two freeways merge into one here. Maximum speed is low and the relationship is very flat.

Although the speed-density relationship is different at different locations, roughly the same pattern can be observed at each location.

Influence of weather

The Smulders function was fitted to the data set given in table 2 using a least squares method to illustrate the influence of weather. The resulting parameters of the calibrated functions are displayed in figure 39. Maximum speed, critical density, the beta parameter, and the resulting capacity are displayed for each location. The beta parameter is displayed as hundreds and critical density as percentage of maximum density, assumed to be 125 vehicles per lane.

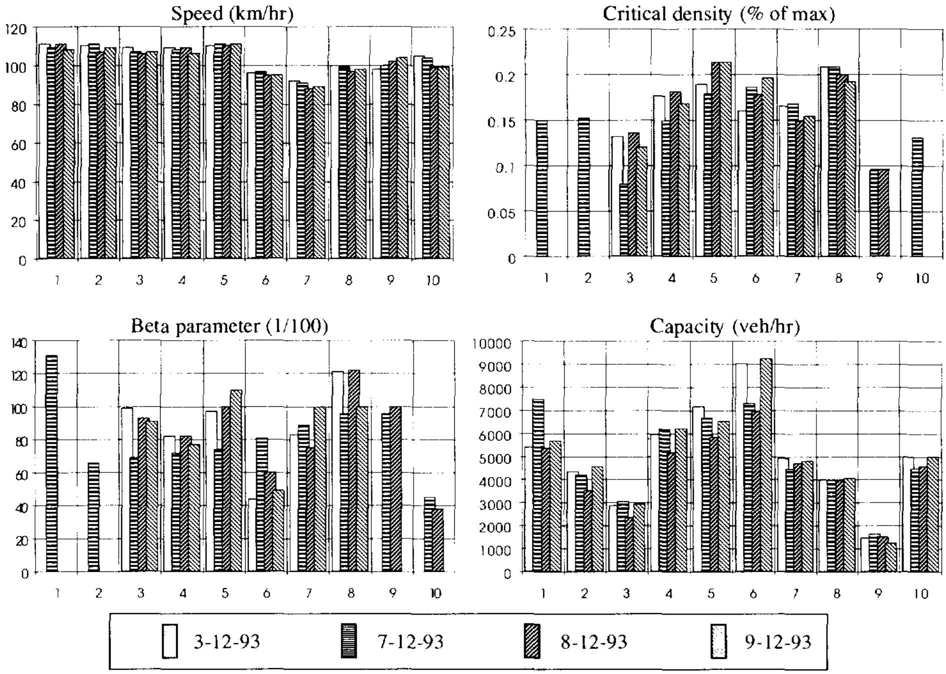


Figure 39: Maximum speed, critical density, β and capacity for each location on different days, locations 1,2,9 and 10 did not always have enough "low speed" data to give a reliable estimate of β and critical density, for those days the β value and critical density are omitted.

The graphs show little variation in maximum speed, however, the β parameters especially, which represents the speed-density relationship in a congested state, shows a lot of variation. Capacity is a result produced by all the other parameters. The last graph in figure 39 shows that, at some locations, there is a significant difference in capacity. In general data from 8-12-93 results in the lowest performance of the links. Since this was the day with the fog, this result is explainable.

Changes during the day

Until now the relationship between speed and density has been compared between different *days*. It is also important for traffic predictions to study the changes in the function *during* the day. These changes are probably related to weather circumstances and time of day, i.e. darkness.

The influence of weather is illustrated with the data set for the 8th of December 1993. It was a day with a lot of fog and bad visibility. There was a good visibility only between 07:00h and 11:00h in the morning.

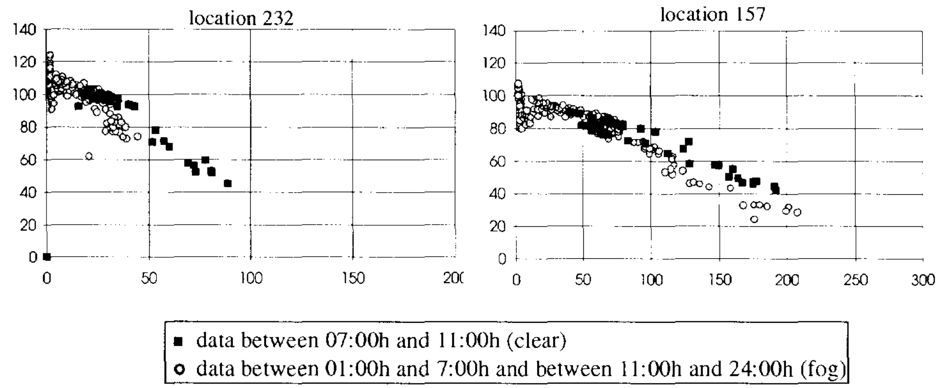


Figure 40: Within day influences for speed-density relationships

The graphs in figure 40 show data measured during the period with good visibility as black squares and during the fog with circles. Higher speeds were observed for both locations during the clear period.

When the data set is smoothed the changes in time for the speed-density relationship become more clear. The data is smoothed for 15 minutes in figure 41.

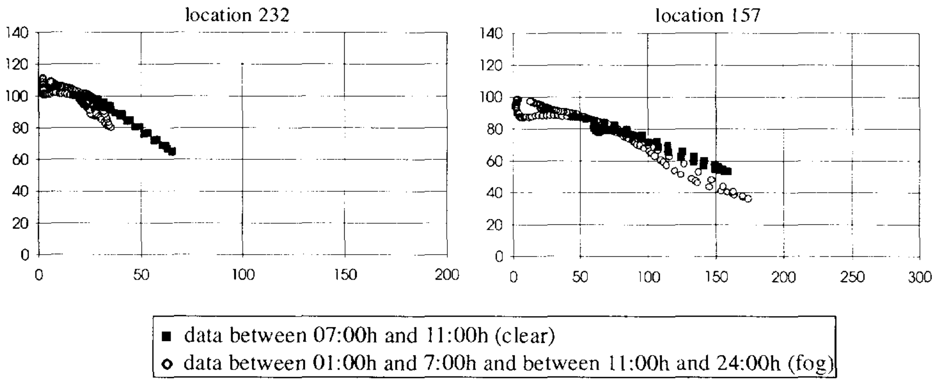


Figure 41: Within day influences for speed-density functions with aggregated data (moving average for 15 minutes).

The patterns show a slightly different relationship during the fog than during the clear period for both locations in figure 41.

The 3rd and the 7th of December were days with rain. The influence of rain seems to be not as clear as the influence of fog. Only very slight changes in the speed-density relationship were observed for these days. The influence depends, of course, on the amount of rainfall, which was not very high during these days, heavy rainfall may give a similar result to fog.

Influence of incidents

To investigate the effect of incidents on the speed-density function, additional information about incidents is required. No additional information was available with the Amsterdam beltway data. During data collection on the A13 between Delft and Rotterdam at the 23rd of March 1990 an accident happened while data was collected upstream and downstream of the accident. Figure 42 shows speed-density plots at two locations.

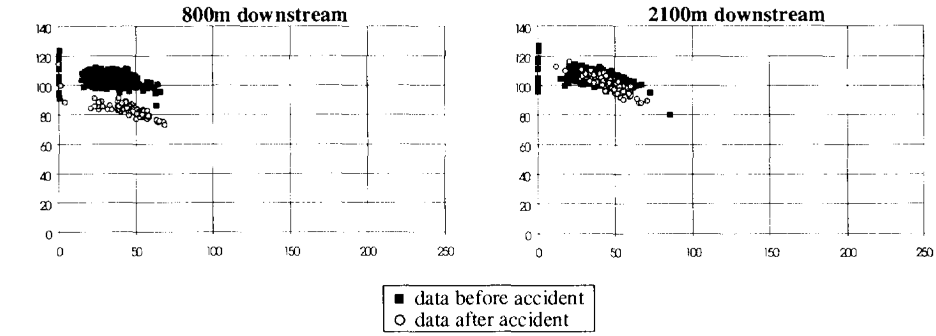


Figure 42: Speed-density relationship with an accident on the A13 in the Netherlands, measured on 23-3-90.

The figure shows the impact of the incident 800m downstream and 2100m downstream. At 800m the speed-density relationship changed significantly after the accident occurred. Cars are accelerating and a lower average speed is measured. No difference is measured at 2100m downstream.

5.3.3 Speed-Density Function used in 3DAS

The variability in the shape of the speed-density function was demonstrated in the previous sections. It seems very hard to define a generally applicable speed-density function. All kinds of external influences change the relationship between speed and density. If conditions are constant, however, the relationship between speed and density seems reproducible, and a certain basic shape is always recognizable. Making a distinction towards different circumstances is very hard, the influence of weather, for example, is very hard to quantify. All kinds of rainfall have to be distinguished, several levels of visibility, etc. To describe all these influences with a theoretical function requires a function with a large number of parameters, and it is for this reason that a mathematical approach for a speed-density function is preferred. The Smulders function, with the additional parameters, expresses a function that has a correct basic shape and can be fitted to various conditions by changing the parameters. This *modified* function seems to be best suited for the speed-density function in the 3DAS model and is described in equation 74.

Figure 43 shows how the parameters affect the shape of the speed-density relationship and the flow-density relationship. Speed at the critical density is changed, by changing the α parameter, which results in a changed capacity.

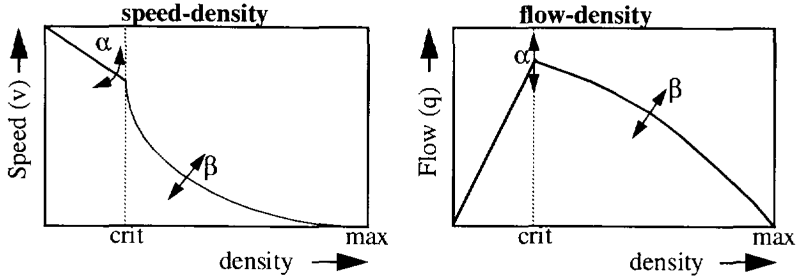


Figure 43: Influence of α and β on the speed-density relationship

The shape of the function is influenced by varying the α and β parameters. Changing these parameters influences the travel time directly. Figure 44 demonstrates this influence on the travel time for one link in one period.

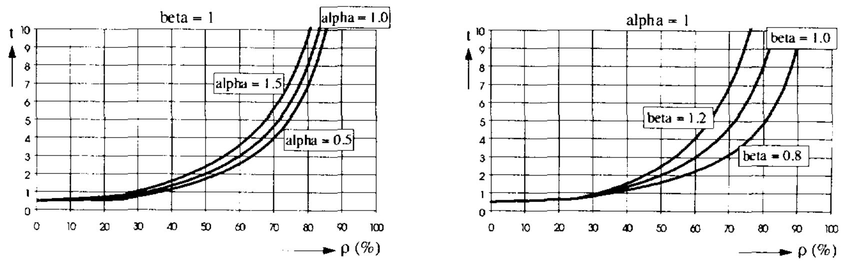


Figure 44: Influence of the parameters α and β on the travel time

Figure 44 shows that with high densities, the influence on travel time is very significant. Caution is needed when varying the α parameter, because this influences the *capacity* of a link directly. When the capacity is set to 100 percent at $\alpha = 1$, than the capacity decreases to 85.7 percent when $\alpha = 0.5$, and the capacity increases to 120 percent when $\alpha = 1.5$.

5.3.4 Generalizing the Influence of Location.

A more thorough study is required to generalize the result of the previous sections. More data is necessary, more precise weather information, and traffic information, i.e. accidents. If an attempt is made to generalize the results for different road types, a table can be made giving the parameters of the *modified Smulders function* (eq 74) for different road types. This table is not more then a rough guide, only data from the Amsterdam beltway system was used for this research (see appendix A).

Ten different *types* of roads are described in table 3. The types were chosen in such a way that obvious differences between speed-density relationships were eliminated.

Table 3: Road types

nr	Description
1	Three lane freeway section
2	Two lane freeway section
3	One lane on-ramp
4	One lane off-ramp
5	Three lane freeway with on-ramp nearby
6	Two lane freeway with on-ramp nearby
7	Two lane freeway change (e.g. fly-over)
8	One lane freeway change (e.g. fly-over)
9	3 lane freeway in tunnel
10	Four lane merging section (two, two lane freeways merge)

A maximum density of 100 vehicles per lane per kilometre was assumed for each location. Critical density, maximum speed and the parameters α and β were estimated, using the data set measured in Amsterdam. Data was used from 10 days with 100 screenlines per day and a sample interval of one minute. The 100 screenlines are classified according to the road types described in table 3. The data was measured on June 14,15,16, and 18 and December 1,3,6,7,8, and 9 1992.
The results are listed in table 4. The first column specifies the road type, the second column the number of screenlines in the data set of that type, the remaining columns specify the parameters of the Smulders function. For the locations column, the value between brackets represents the number of locations with congestion. The value for β could only be estimated for those locations. The other columns give the mean value and, between brackets, the standard deviation.

All the parameters were estimated using a least squares method. The value for the critical density is unreliable for those location with little or no congestion data. Table 4 shows a relatively large standard deviation for the α parameter. This indicates rather unstable behaviour for the speed-density function. A table is given in appendix B with the mean values and variance for each individual location on the Amsterdam beltway.

The impact of weather could not be performed due to a lack of accurate weather data. The impact of weather depends very much on the precise type of weather. Different types of rainfall should be distinguished and different types of visibility. All these influences plus the lack of data makes the quantification of weather impact on the speed-density relationship a precarious job.

Table 4: Rough parameter set for speed-density functions

road type	locations	crit dens	max speed	α	β
1	300 (46)	70.0 (7.9)	109.0 (6.6)	0.57 (0.38)	0.84 (0.15)
2	136 (43)	44.4 (8.3)	107.2 (4.6)	0.85 (0.5)	0.80 (0.20)
3	30 (2)	24.4 (2.17)	103.2 (3.0)	1.23 (0.42)	0.28 (0.03)
4	30 (7)	25.9 (9.08)	98.7 (4.9)	1.03 (0.51)	0.49 (0.18)
5	141 (21)	69.7 (7.46)	108.7 (8.3)	0.62 (0.37)	0.80 (0.17)
6	10 (0)	48.0 (0.09)	105.5 (2.72)	0.59 (0.31)	no data
7	140 (30)	48.1 (4.42)	99.7 (6.74)	0.51 (0.29)	0.78 (0.31)
8	30 (1)	23.6 (2.41)	105.5 (4.77)	0.57 (0.46)	1.13
9	20 (0)	72.0 (0.01)	107.9 (2.49)	0.44 (0.50)	no data
10	10 (5)	99.6 (6.11)	95.9 (2.01)	0.76 (0.10)	0.56 (0.11)

5.4 Comparing 3DAS with the Microscopic Simulation Model FOSIM

To validate the 3DAS model, it was tested in some simple simulated situations. If the model did not perform well in these simple situations, simulations of large networks would be useless.
Some basic tests were carried out to validate the evolution of traffic flow in the model, this is important for the proper behaviour of the model. No tests on route choice were included due to lack of comparative data.

The results of the following tests are reported:
- a 10 km roadway with a bottleneck
- a 10 km roadway with an on-ramp

The model is, in principle, not intended to reproduce each kilometre of roadway in detail, but it should be able to simulate correct in and outflow and travel time for the entire roadway.

These test situations were chosen for two reasons:

I. They can be found in any network and are crucial for traffic behaviour in a network. When a bottleneck situation is included in a larger network the model is capable of simulating the delay of traffic due to the bottleneck and the effects up and downstream within the network..

II. Test results for these situations are well known, which makes them easy to validate.

The test results were compared to a microscopic simulation model, FOSIM [94]. FOSIM is microscopic simulation model that evolved from INTRAS [100]. The model has been further developed at Delft University of Technology and specifically adapted to study merging behaviour [81][82]. It has been validated and calibrated for situations with bottlenecks, merging sections and on-ramps, and has been proved to represent the traffic realistically. It is possible to test 3DAS in several controlled and realistic situations using FOSIM.

The results of the microscopic simulation were aggregated to one minute data. The tests pushed the 3DAS model a bit beyond its limits, as an appropriate time period for network applications would be at least 5 minutes. An accuracy of one minute for entire networks is hardly feasible.

The different methods for pathfinding and assignment were tested and the parameter for queuing was calibrated during the comparison. The parameter for the merging algorithm was tested with the on-ramp network. The results of the comparison are reported in sections 5.4.1 and 5.4.2 and the results of the calibration are reported in section 5.5.

The tests show the capability of the 3DAS model to simulate some simple basic test situations. The 3DAS model is, however, intended for large networks and should not be confused with microscopic simulation models and their purposes. Whether the model behaves correctly for large networks is tested in chapter 7, where the model is applied to a part of the Washington D.C. road system, and in chapter 8, where the model is applied to the Amsterdam beltway.

5.4.1 A Bottleneck

A ten kilometre road with three lanes running into two lanes. The road is made of ten links of one kilometre length each. The first five links have three lanes and the last five links have two lanes. The simulated time is twenty minutes, with 40 periods of one half minute.

The network:

Figure 45: Linear test network with bottleneck at the sixth node

The trips travel from left to right. The number of cars entering the section for each minute is shown in table 5:

Table 5: Dynamic OD-values

minute	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
#cars	40	56	72	88	104	104	104	104	104	88	72	56	40	20	8	0	0	0	0	0

The graphs in figure 46 show the time dependent flow (veh/hr) for six of the ten links. The time periods are displayed along the x-axis. The fifth graph is the first link with two lanes. The grey line in each plot displays the FOSIM results and the black line displays the 3DAS results.

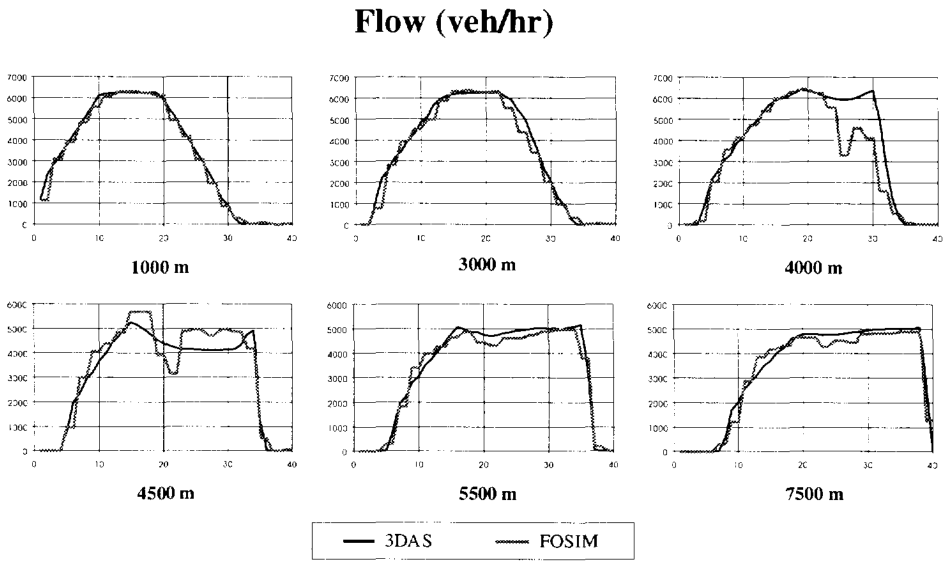


Figure 46: Flow on six links for FOSIM and for 3DAS

The six graphs in figure 47 show the corresponding speeds (km\hr) on the same links.

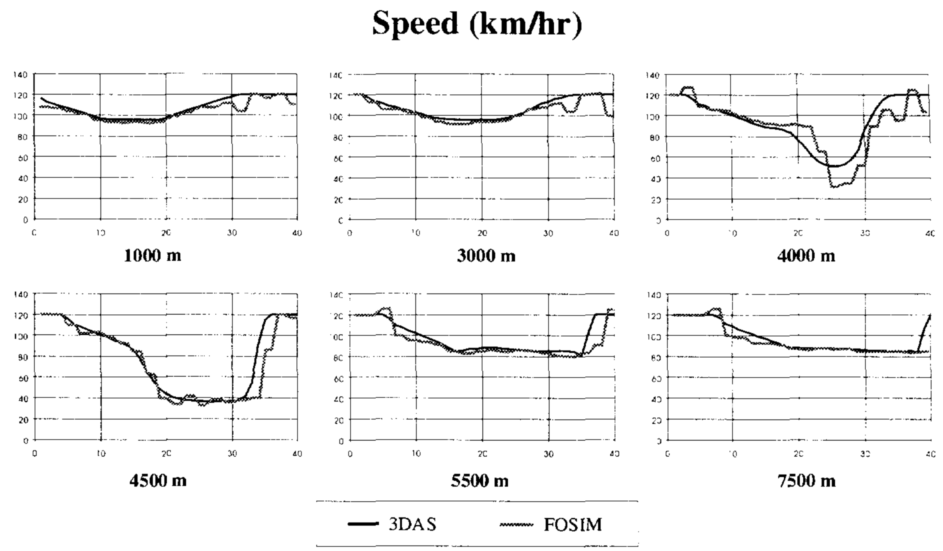


Figure 47: Speed on six links for FOSIM and for 3DAS

Comments:

The graphs show the flow (fig.46) and the speed (fig.47) at 1000m of trips as they enter the network. The same shape, shifted in time, is observed at 3000m. Due to queuing in the bottleneck (at 5000m), the flow at 4000m drops in the 24th period, the corresponding speed decreases to 20 km/hr. Up to the 28th period, the number of cars entering the bottleneck decreases and the flow increases again. After the 28th period, the number of cars entering the bottleneck decreases further and the flow also decreases, while the speed increases further. At 4500m the same process shifted in time is observed. Downstream from the bottleneck a steady flow of approximately 5000 veh/hr is observed. This flow is shifted in time in the last graph at 7500m.

The behaviour at the bottleneck is not reproduced in detail by the 3DAS model, but downstream from the bottleneck the results match well, the delay due to the bottleneck is reproduced. The queuing effects at 4000 m, however, are slightly different. It has to be realized that an exact reproduction of speed **and** flow is impossible. The 3DAS results represent the situation during the entire length of the link, while the FOSIM results were measured at one point. Further, the 3DAS model has a strict relationship between speed and flow, while FOSIM does not have this restriction. At 4500m, for example, the FOSIM results show a constant speed of ±40 km/hr between the 10th and the 20th period. The flow during this time, however, shows significant fluctuations. These fluctuations are impossible to reproduce using 3DAS, however, speed was reproduced.

5.4.2 An On-ramp

The second network is a ten kilometre road with two lanes throughout and with an on-ramp at the 5th kilometre. The road is made of ten links of one kilometre, the on-ramp starts at the 3rd kilometre and merges at the fifth kilometre. The time simulated is thirty minutes, with 60 periods of one half minute.

Situation:

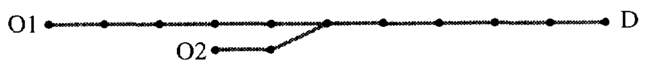


Figure 48: Test network with on-ramp at the sixth node.

Trips travel from left to right. The number of cars entering the section for each minute is shown in table 6:

Table 6: Dynamic OD-values at the main road and the on-ramp

minute	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
O1	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43
O2	12	12	12	12	12	12	12	12	12	18	26	34	39	38	34	29	23	23	14	2

The graphs in figure 49 show the time dependent flow (veh/hr) for eight of the twelve links. At 3500m and 4500m, the flow at the on-ramp is shown by the lower line. The black line displays FOSIM results and the grey line displays the 3DAS results.

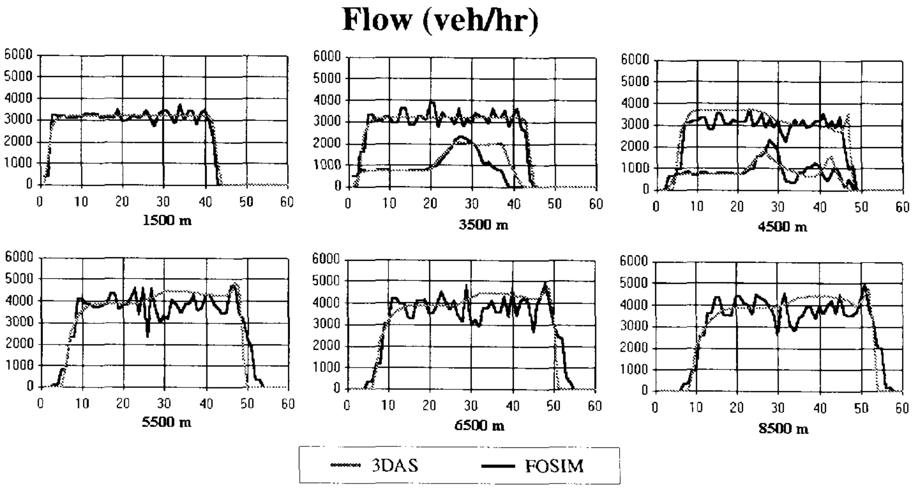


Figure 49: Flow at eight links for FOSIM and 3DAS

The six graphs in figure 50 show the corresponding speeds (km/hr) on the same links.

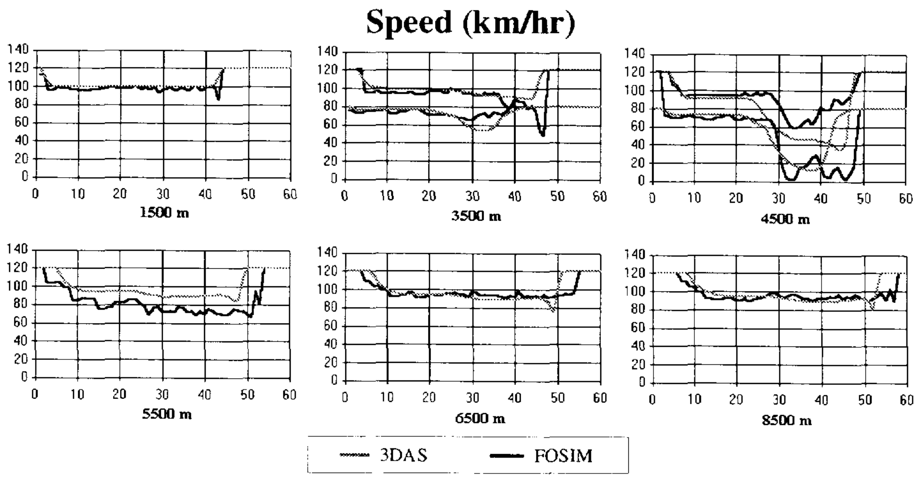


Figure 50: Speed at eight links for FOSIM and 3DAS

Comments:

The graph at 1500m shows the trips as they enter the network. A steady flow of 3000 veh/hr enters the main road of the network during a total time of 20 minutes (40 periods). The same flow can be observed shifted in time at 3500m. Trips entering at the on-ramp are shown in the lower plots in these graphs. A steady flow of 800 veh/hr enters the on-ramp for 20 periods, in the next 8 periods, the flow increases to 2200 veh/hr, and then it decreases to zero. At 4500m, the traffic has not merged yet and the same flow shifted in time is observed on the main road. At the on-ramp, however, queuing is observed at 4500m, and speeds are very low. At 5500m, the trips have merged and a flow of approximately 4000 veh/hr is observed. The same flow, shifted in time, with a slowly increasing speed is observed at 6500m and 8500m.

The results at the on-ramp are not reproduced in detail by the 3DAS model. FOSIM has more queuing at the on-ramp than on the main road. The 3DAS model has the same queues on the main road as at the on-ramp. This can be seen clearly at 4500 m. The downstream effect of an on-ramp is fairly well described.

5.5 Calibration Results of Comparison with FOSIM.

During the comparison with FOSIM, the different methods for pathfinding and assignment were tested and the parameters for queuing and merging were calibrated.

5.5.1 Pathfinding

The four pathfinding methods described in section 4.5.1 were tested with various networks. The two methods suitable for further validation were methods three and four. With these methods, trips departing later, cannot arrive earlier. Method four is theoretically the best and was expected to give the best performance. During the validation, however, some unexpected problems with method four were encountered. If there is a bottleneck at a certain link, trajectories departing from different departure periods tend to converge to the same trajectory, resulting in a high concentrated density. The observation is illustrated by figure 51. The figure displays the trajectories calculated in the last iteration of the model. The left side of figure 51 displays the trajectories as calculated by method three and the right side of the figure displays the trajectories as calculated by method four.

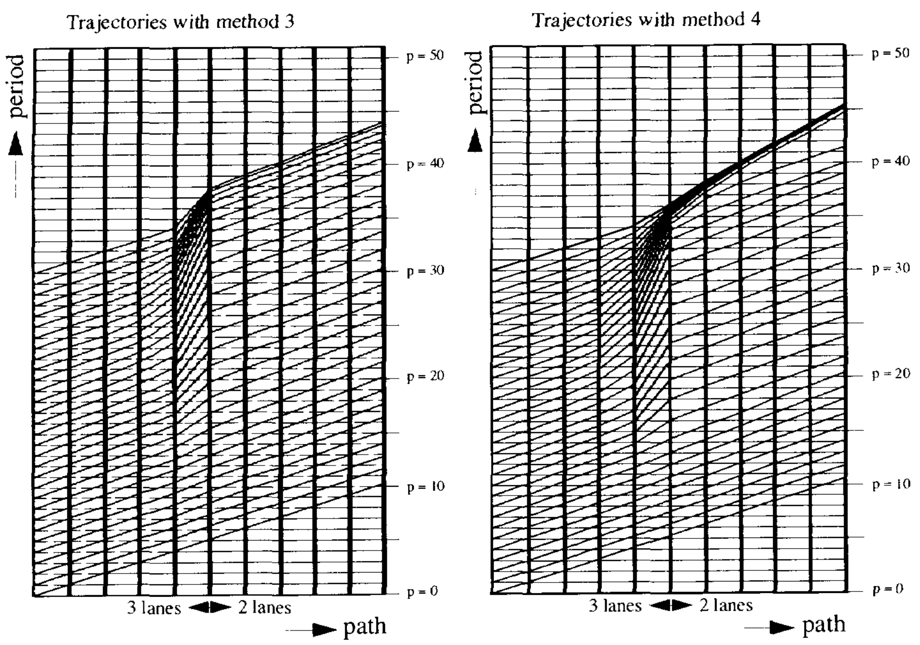


Figure 51: Trajectories (final iteration) with "bottleneck" network for method 3 and method 4.

The right side of figure 51 shows how the final trajectories converge to the same line. This is caused by the adaptation made when each trajectory crosses a time period. With decreasing travel time in the later periods, the trajectory becomes less steep and results in a highly concentrated set of trajectories in the final periods. With pathfinding method three, the adaptation in the trajectory is done using an increment of a period length, and thus results in better separated trajectories.

5.5.2 Assignment

The two methods for assignment described in section 4.5.2 were tested using various networks. Of the two methods described, the "surface" method was expected to give the best results. The results of the trajectory method depend very much on the network and the chosen period length. A situation in which the trajectory method performs poorly occurs with a bottleneck in the network. Figure 52 shows the flow on the seventh link in the network of figure 45. The link is situated downstream from the bottleneck.

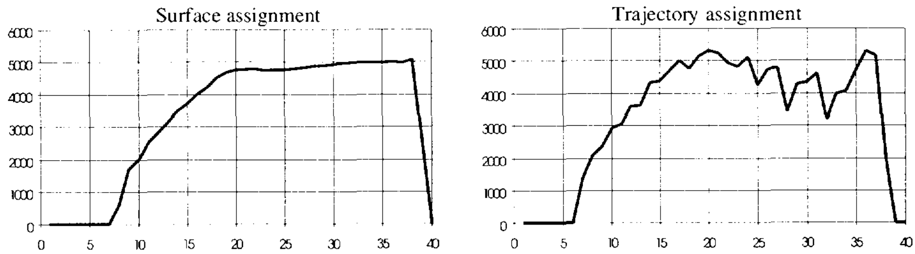


Figure 52: Flow on the seventh link of the bottleneck network assigned using the surface method and using the trajectory method.

The fluctuating pattern using the trajectory method can be explained by the fact that a trajectory can skip a link in a certain period. This situation can occur when a trajectory traverses link a in period p and the next trajectory traverses link a in period $p+2$. This results in no load in period $p+1$. The situation is demonstrated by an example in figure 53.

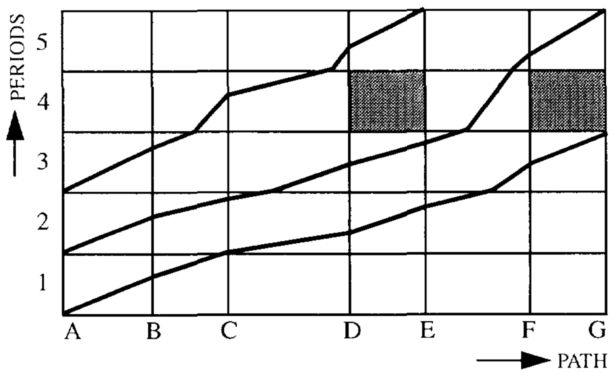


Figure 53: A trajectory in which certain periods are skipped, in this example nothing is assigned in period 4 to link D-E and to link F-G.

With the surface assignment method, trips are assigned between the trajectories. This results in a much smoother, better, result.

5.5.3 Queuing

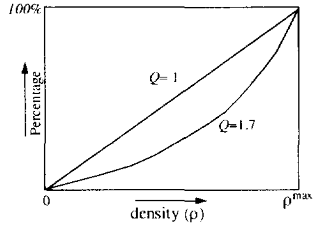
The queuing algorithm in 3DAS is explained in section 4.7. Two different methods to achieve a blocking back mechanism and to prevent density on a link from exceeding the maximum density are described. Both methods were tested with the *bottleneck* network. Based on these results, only the second method was tested using the *on-ramp* network.

Queuing method I starts to block back when maximum density is exceeded on a link. In the next iteration, all the trips will be assigned to the preceding link in the path. When using the Smulders speed-density function, this means that at maximum density, speed is equal to zero and travel time is infinite. This results in zero flow on the congested link and on all downstream links. This situation is not observed in FOSIM or in reality. Two different solutions to guarantee a certain outflow are possible:

I. Change the speed-density function, and maintain a certain minimum speed and minimum outflow at maximum density.
II. Do not allow the traffic to reach maximum density, but start to block back before maximum density is reached.

As described in section 4.7, the second option was chosen. The FOSIM results show that there is blocking back to preceding links before the maximum density is reached on the congested link. Method II prevents the congested link from reaching maximum density, and a consistent outflow is maintained. Outflow is controlled by the shape of the speed-density function and the parameter Q in equation 37. The equation and the shape of the corresponding function are repeated in equation 78.

$$F(\rho) = \left(\frac{\rho}{\rho^{max}}\right)^{Q} \qquad \text{(eq 78)}$$



Different values for the parameter Q were tested. The three graphs in figure 54 show the results with the bottleneck network of figure 45 using three different values for Q. The link preceding the bottleneck, the link with the bottleneck, and the link after the bottleneck are shown. The flow and the speed are shown and the three different values of Q are plotted in the same figure.

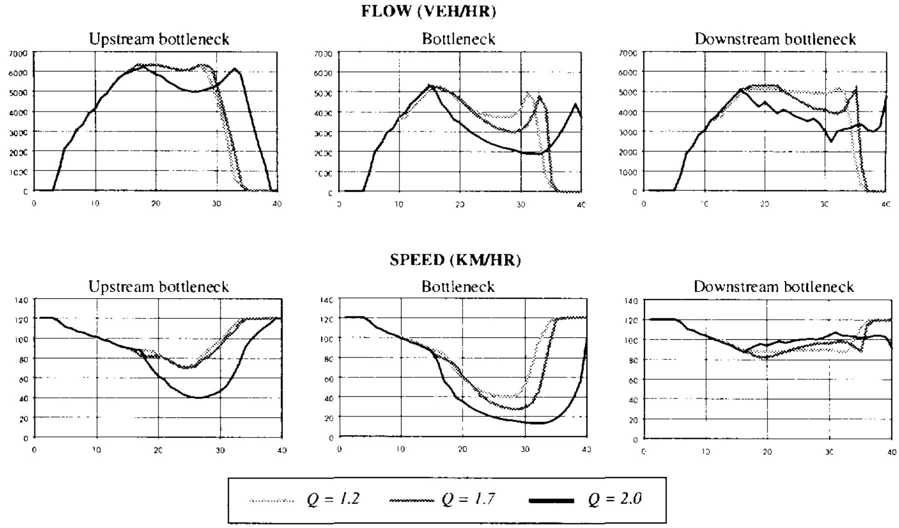


Figure 54: Three successive links with the load patterns in time for three different values of Q.

The best reproduction of the FOSIM results was achieved using a value for Q of 1.7. The influence of the parameter Q is very intuitive. A higher value ($Q > 2$) results in queuing with higher densities only which results in higher densities and a lower outflow, while a lower value ($Q < 1.7$) results in more queuing, lower densities and a higher outflow.

5.5.4 On-ramp Link Dependency (merging)

The merging algorithm was tested using the on-ramp network of FOSIM. As explained in section 4.7 the parameter M in equation 38, determines the influence of the on-ramp on the critical density. The best match with the FOSIM results was obtained with a value for M of *0.25*. The influence of the parameter is very intuitive. With higher values of M, greater influence of the on-ramp is observed; with lower values less influence is observed. The merging process needs more study. Data from the field are very sparse, and the influence of an on-ramp is complex. The FOSIM results showed more queuing at the on-ramp, while in reality the main stream seems to suffer more from merging traffic. The number of lanes plays an important role in this behaviour.

5.6 Conclusions

The consequences of various discretization phenomena, the influences on the speed-density function and the calibration of several parameters are demonstrated in this chapter.
The most obvious conclusion concerning link length is that longer links are less detailed and less accurate. Areas with an expected high level of congestion should be modelled by short links. It is most important that the definition of the network and the length of the periods are adapted to the fluctuations in demand. When demand fluctuates highly, shorter periods and a more detailed network are advisable. An important advantage of the 3DAS model is, however,

that detail does not get lost for downstream links, though an underestimation of travel time can occur on long links. This underestimation increases with higher densities.
To derive the travel time to traverse a link, based on the conditions on the link, a travel time function should be based on density, only with a speed-density relationship is the dynamic behaviour of a link described correctly. Using data from the Amsterdam beltway it is argued that there is no generally applicable speed-density relationship for all links. The relationship should at least be tailored to the location for planning applications (ATPS), because a merging section shows, for example a different relationship than an exit ramp. The relationship should also be adopted to current conditions, such as the influences of weather, time of day and accidents for management applications (ATMS). Using the Amsterdam beltway data it was showed that a *modified* Smulders speed-density function gives a good fit with the data, and is easily adapted to different situations.
The third and fourth sections described a comparison between the 3DAS model and the microscopic simulation model FOSIM. The 3DAS model is shown to be capable of reproducing some basic traffic situations. Although the results were not reproduced in detail, the effect of a bottleneck and an on-ramp *in a network* were reproduced. Several parameters of the model were calibrated during the comparison. The parameter for the queuing algorithm performed best with a value of $Q = 1.7$, and the parameter for the merging algorithm with a value of $M = 0.25$. Further the influence of the pathfinding method and assignment method were evaluated; path finding method 3 and the surface assignment method performed best.

CHAPTER 6

USING 3DAS FOR PLANNING AND MANAGEMENT

6.1 Introduction

Three different kinds of systems in relation to DTM can be distinguished, as has been stated in section 2.3. Firstly there is a need for systems to support *planning decisions* such as the construction of new roads or the placement of new instruments, secondly systems that provide information about *management decisions* such as incident management and rerouting measures, and thirdly there are various *control decisions* such as traffic signal settings. These three systems are called respectively: Advanced Traffic Planning Systems (ATPS), Advanced Traffic Management Systems (ATMS) and Advanced Traffic Control Systems (ATCS).
The main reason for making these divisions is the different objects and time horizons of the decision making process. ATPS is concerned with long term planning (days to years), ATMS with medium term management (few hours) and ATCS with short term control (few minutes). This division has direct consequences for the accuracy of the models, the required data and the level of aggregation. ATPS can do with rather crude data, aggregated over longer periods of time. Much more detailed and actual data is required for ATMS, with more accuracy and reliability. ATCS requires the most accurate and actual data, with a high level of detail. Further, ATMS still involves a person for decision making, while with ATCS no human interference is present, so the reliability of the system and the accuracy of the data is vital.

The dynamic assignment model 3DAS, described in chapters 4 and 5, models the traffic flow on a network with a medium level of detail. The conditions on a link are assumed to be homogeneous, and constant for the duration of each time period. The model does not discern individual vehicles or different lanes. This means that possible applications for 3DAS are to be found in ATPS and ATMS. A level of detail is required for ATCS which cannot be reached with the 3DAS model. Models that adopt microscopic simulation techniques are better suited for ATCS applications.

As has been explained in section 4.3, the basic input requirements for the 3DAS model are:

- A *network*, with links to represent the road segments and nodes to connect the links. Each link has its static parameters, such as length and maximum density.
- A *dynamic OD-matrix*, defining the travel demand for every OD-relation in the network, and for every departure period of the total time span.

To make the model adaptive to prevailing conditions, the model uses:
- *Speed-density functions*, which represent the *dynamic characteristics* for each link in the network, i.e. maximum speed and capacity.

The requirements for each of these three parts differ per application; for long term ATPS applications, the 3DAS results need to represent the average traffic distribution in the network. This means that the data must be based on average traffic demand and average conditions on a link, and that the results must be validated by a traffic distribution averaged over a longer period.
For ATMS applications, the 3DAS results must represent the actual traffic situation of that moment, including temporary disturbances such as incidents and roadworks. Now the data must be based on the actual measured conditions in the network, and the results should be compared with these actual conditions.

Section 6.2 will describe a framework for the use of 3DAS in ATPS applications while section 6.3 covers the use of 3DAS in ATMS applications. The chapter is concluded with a section about the hard and software requirements.

6.2 Using 3DAS for ATPS Applications

The main task of Advanced Traffic Planning Systems is to provide insight into the development of traffic for the long term, study the impact of various instruments and support decisions about infrastructural changes, the placement of changeable message signs and ramp metering installations, etc.
Dynamic traffic assignment models can give a better understanding of the consequences of congestion, the interaction between new infrastructure, new instruments, and congestion, and the environmental impact of traffic. Authorities are also demanding that the impacts of there policies are evaluated against objectives of increasing complexity. The application of dynamic traffic assignment models for long term planning studies, seems inevitable.
A typical planning application would study the influence of a new bridge, or an extra ramp meter on the level of congestion in the rush hour. Such a study provides insight into the effects of the measure taken, for example, how is the congestion alleviated or moved from one place to another.
Data requirements for such a study, using a dynamic assignment model, are higher than the data requirements for traditional studies, using static assignment models and, because the level of detail is higher using 3DAS, the data must also support this level of detail.

Using the 3DAS model in an ATPS application requires no additional new models. The model fits neatly into the traditional four step approach. The network is defined in a traditional manner, a static OD-matrix can be made dynamic with additional departure time pattern information, and speed-density functions can be based on average values. Figure 55 gives the relationships between the data sources and the model in a schematic structure. The requirements for the data in the framework of figure 55 are described in the next three sections. In chapter 7 the framework is applied to the road network of Washington DC, USA

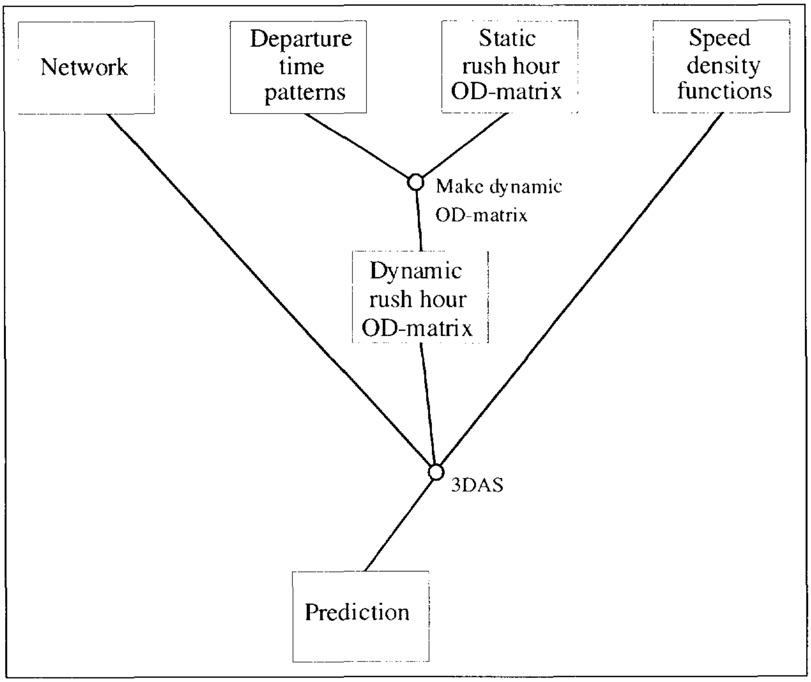


Figure 55: Framework for 3DAS in an Advanced Traffic Planning System

6.2.1 Network

The network is defined as described in section 4.3.1 and is basically similar to traditional static networks. Road characteristics, such as maximum speed, capacity and maximum density must be based on average values.

6.2.2 OD-matrix

For ATPS applications the OD-matrix, has to describe the average trip rates in the network. A time dependent OD-matrix is preferable because the 3DAS model is time dependent. By combining a *static* OD-matrix with departure patterns, an adequate approximation of the average time dependent demand can be achieved. Using the method described in section 4.3.2 a static OD-matrix can be made dynamic by using departure time patterns. A rush hour OD-matrix can be estimated by traditional methods and data sources, such as home interviews, license plate surveys and roadside interviews, and with additional data from induction loop detectors good quality static rush hour OD-matrices are possible.
The required departure time patterns can be derived from induction loop information. An impression of the time dependent flow is obtained, by measuring the flow as time series at several points in the network. These flow patterns can serve as departure time patterns.

6.2.3 Speed-Density Functions

The speed-density functions describe the dynamic behaviour of a link. The function can be based on the number of lanes and the type of the road (tunnel, on ramp, arterial, etc.). As an ATPS application requires the average behaviour of a link, it is not necessary to adapt the functions to weather circumstances, accidents, etc. Table 4 on page 87 gives a good starting point for the speed-density functions.

6.3 Using 3DAS for ATMS Applications

One of the goals of ATMS is to prevent congestion, so rather than analyse and react to information, forecasts are required to take action prior to degradation of the traffic flows. Thus, one of the tasks of Advanced Traffic Management Systems (ATMS) is to give a prediction of the traffic conditions for the medium term. Based on this prediction it is possible for a traffic operator to make decisions concerning rerouting measures, travel advice and ramp metering settings. Such a prediction system must respond to accidents, weather changes, road works, etc., and input data must be based on *actual measured* data. To react promptly to the prevailing conditions, new predictions must be made frequently. When the 3DAS model forms the basis of such a prediction system, two extra modules are required to pre-process the on-line data. One module to estimate OD-matrices, and one module to update the speed-density functions. The framework is given in figure 56.

Different data sources are possible within the ATMS, such as historical data, probe vehicle data, and other data sources (police, mobile phones, etc.). The modular structure makes it possible for each module to work independently, and decide for itself which data sources to use. At this time, the main source for data in the Netherlands are induction loops. The Netherlands has a nation wide data collecting system called the monitoring system [3][37]. This system collects traffic data using induction loops and equipment to measure the road and weather conditions. The Motorway Control & Signalling System (MCSS) system is a part of the monitoring system. The MCSS system on the beltway of Amsterdam is explained briefly in Appendix A.
Each module in this system is described briefly in the next sections. An application of 3DAS as an ATMS is described in chapter 8.

6.3.1 Network

For ATMS applications, the network has to be dynamic. Although most changes in the network are covered by the speed-density functions, it is also necessary to adapt the network to changing conditions, such as road works and incidents. The updates to the network can be done manually or automatically. Using probe vehicles it is possible to detect changes in the network automatically. The state of the network can be monitored by collecting position data from the probe vehicles, and accidents, closed roads and even new roads can be detected [99].

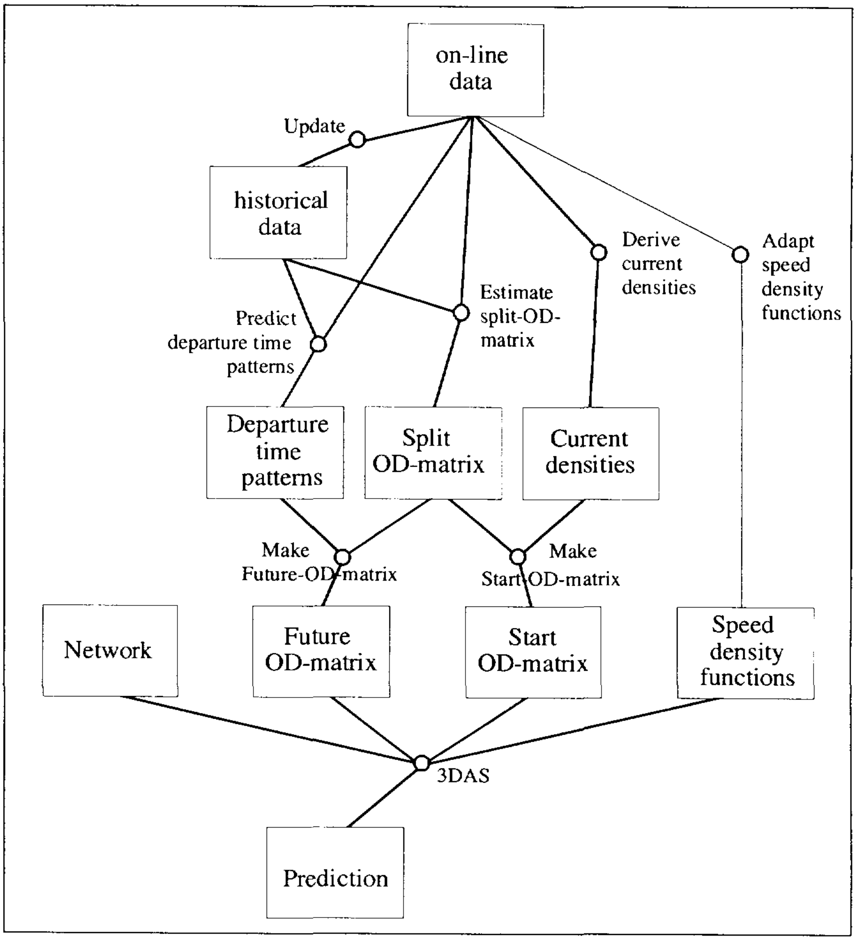


Figure 56: Framework for 3DAS in an Advanced Traffic Management System

6.3.2 OD-matrix

As can be seen in figure 56, estimation of the OD-matrix involves several modules which results in two OD-matrices as input for the 3DAS model. For an ATMS application the evolvement of traffic has to be calculated frequently, so the 3DAS model should be able to start with a loaded network. Each time a new calculation starts, the current flow on the network has to be used as a start situation. To achieve this two different OD-matrices are needed for the assignment. These two different OD-matrices will be referred to as *the start OD-matrix* and *the future OD-matrix*.

Start OD-matrix:

To reproduce the current status of the traffic on the network an OD-matrix is needed with every node in the network as an origin. This matrix is 2-dimensional because it is defined for only one departure period. This matrix will fill the entire network instantaneously and reproduce the current measured status of the network.

Future OD-matrix:
To calculate the traffic in the network an OD-matrix is required that defines the traffic that enters the network for the future periods. This matrix has to be defined only for the origin nodes. The matrix is 3-dimensional because the traffic demand is time variable.

The *start OD-matrix* contains information only for the first period of the calculation and the *future OD-matrix* contains information for the remaining (future) periods.

In figure 56 five different processes are involved to estimate the OD-matrices. At Delft University interrelated research is being done towards the estimation of OD-matrices for dynamic traffic management purposes [104][105]. As this research is aimed at the estimation of OD-*split*-matrices, additional information is required to estimate the current volumes at each link, and the future volumes at the origins. Using the current volumes and the OD-split matrix it is possible to determine the start OD-matrix, and using the future volumes and the OD-split matrix it is possible to determine the future OD-matrix. *Future volumes at origins* is usually referred to as *departure time patterns* or *departure time functions*, for this reason, two additional modules are necessary to calculate the OD-matrices. A module to derive the *current densities*, and a module to estimate the *departure time patterns*. The *start-OD-matrix* and *future-OD-matrix* are easily calculated using these sources of information. All the processes are explained briefly in the next five sections.

6.3.2.1 Derive Current Densities

This module determines the current state of the network by estimating the current density for each link, based on the induction loop data. The induction loops on the Amsterdam beltway cover almost every link, and using this data the current state in the network is easily derived. For those links not covered by induction loops, density is estimated using up- and downstream loops.

6.3.2.2 Estimate OD-Split Matrix

The OD-split matrix contains, for each origin, the percentage of the total number of trips that depart to all possible destinations. Research towards the design of this module is an independent project [104][105]. The research is aimed at the development of an estimation method for time dependent OD-matrices, preferably using low cost data, like induction loop data. Several different methods are under development.

One of these methods uses a *dynamized* model of trip generation based on existing gravity type models:

$$q_{rs}^{d} = X_{r}^{d} \cdot Q_{s}^{d + w_{rs}^{d}} \cdot F(w_{rs}^{d}) \qquad \text{(eq 79)}$$

with:

q_{rs}^{d}	: Number of trips from zone r to zone s departing in period d
X_{r}^{d}	: Production of zone r in period d
$Q_{s}^{d+w_{rs}^{d}}$	: Attraction of zone s in period $d+w_{rs}^{d}$
$F(w_{rs}^{d})$	: Value of distribution function for travel time w_{rs}^{d}
w_{rs}^{d}	: Travel time for zone r to zone s for departure period d

The parameters of the model, X, Q and F, are estimated using a least squares minimization. The difference between the observed flow on link a in period p, f_a^p, and the calculated flow is minimized by:

$$\underset{X,\,Q,\,F}{\text{Min}} \sum_{a,\,p} \left(f_a^p - \sum_{rsd} \tau_{rsa}^{pd} q_{rs}^{d} \right)^2 \qquad \text{(eq 80)}$$

where q_{rs}^{d} is given by equation 79. Dynamic assignment techniques are required to estimate the values w_{rs}^{d} and assignment fractions τ_{rsa}^{pd}, which represents the proportion of the flow q_{rs}^{d} traversing link a in period p. As the value of τ_{rsa}^{pd} is estimated by assigning the matrix to a network, and the matrix is estimated using the values of τ_{rsa}^{pd}, a fixed point problem exists. To solve this problem, an iterative solution technique is required. In each iteration an OD-estimation and an assignment is performed.
Several additional constraints have been added to this method. For further details about the estimation of time dependent OD-matrices, see [104] and [105]. Although the model estimates a dynamic OD-matrix, only the split values of the resulting matrix are used. Continuing research at Delft University of Technology towards the estimation of OD-matrices is focused on OD-split matrices.

6.3.2.3 Estimate Departure Time Functions:

Information is required about the number of trips that depart at each origin, because the OD-module only calculates an OD-*split* matrix. This departure information is needed for each period, and is usually referred to as the *departure time function.* A departure time function specifies the number of trips that depart during each period. These functions have to be estimated.
The departure time function for each origin can be derived using induction loop data. Experiments done with the Amsterdam data shows that the departure time pattern of a certain location shows very little change for different days.
Figure 57 shows the departure time patterns for four different on ramps in the Amsterdam network. Each graph displays a different location. The time (hours) is represented along the x-axis and the flow (veh/hr) is given along the y-axes.

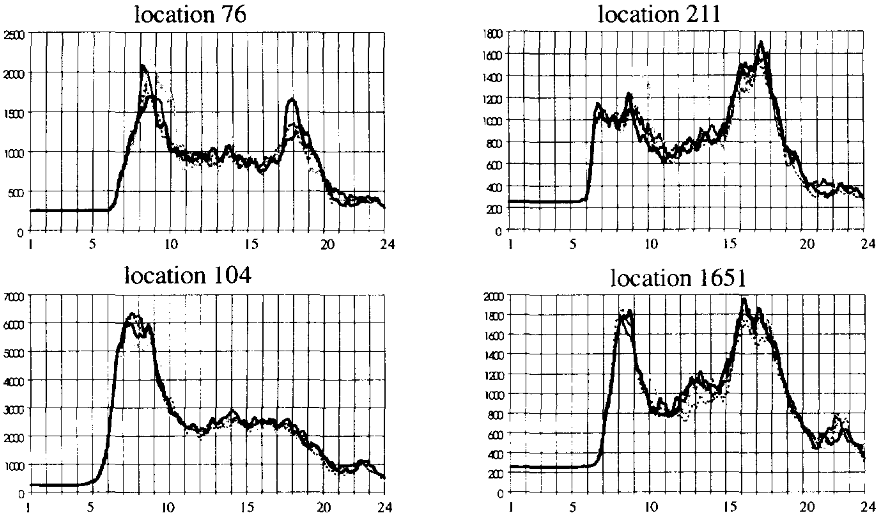


Figure 57: Departure time patterns for four locations and for four different days (December 3,7, and June 5,16 1992). See Appendix A for a map with the locations.

Although there are some differences between days, the *basic* shape of the departure time pattern at a certain location is similar and can be used for different days. For different locations however the patterns are quite different. This indicates that each location has a steady departure time pattern for weekdays. Weekends, however, show a different pattern, and should be handled separately.

It seems obvious to use historic data to predict the departure time pattern. In literature several methods are known to deal with time series. The most widely used technique in transport science is the Box-Jenkins method [10], especially its simpler variant, the Auto Regressive Integrated Moving Average (ARIMA) model. Some approaches describe techniques for estimating traffic volumes for longer term planning applications [60], others for shorter term IVHS applications. An application of the ARIMA model combines volume patterns from different days, and predicts a volume pattern by calculating the optimal combination of these patterns.

Several methods have been tested for the 3DAS model. The flow pattern is measured for each origin (on ramp) in the Amsterdam network and stored in a database. The pattern is measured for several days and averaged and smoothed. The resulting *historical* departure time pattern forms the default for each origin. Experiments with the Amsterdam data showed that, in most cases, this default pattern describes the departure time pattern fairly well. When the departure time pattern deviates from the default pattern, it seems to be shifted or stretched in time, or has a higher or lower amplitude. The basic shape, however, is still applicable.
It is possible to adapt the historical pattern to the current pattern by deforming the default pattern. There are several methods for adapting the pattern. The pattern can be shifted along the time axis, scaled along the time axis, shifted along the flow axis and scaled along the flow axis. Different combinations were tested to predict the departure time pattern for the next one

to two hours. The results showed that the combination of shifting along the time axis and scaling along the flow axis yielded the best results.

The method, to adapt the historical pattern to the current departure time pattern, minimizes a discounted least squares error between the historical and the current departure time pattern. When the historical departure time pattern is described by $H(t)$ and the current pattern by $C(t)$, than the following summation is minimized.

$$\text{Min } z(\lambda,\mu) = \sum_{t \le \tau} w(t) \cdot [\lambda \cdot H(t+\mu) - C(t)]^2 \qquad \text{(eq 81)}$$

Where τ is the current time and λ denotes the scaling factor and μ the shift in time. The square error is weighted for different t by the function w. The most recent t have more weight than older t. Typically the function w is described by $\alpha^{t-\tau}$ where $\alpha > 1$.

This method has been tested at several on ramps in Amsterdam. In general very little shifting and scaling was necessary to give a good prediction of the departure time pattern for the next 30 to 60 minutes.

To illustrate the method the following experiment was done. The data measured on December, 1st (Tu), 3rd (Th), 7th (Mo), 8th (Tu) and 9th (We), were combined, averaged and smoothed for on ramp number 1651. The resulting departure time pattern represents the historical departure time pattern for on ramp 1651. The days used for this database were all week days, because the weekends show a very different pattern. The only day not present in the database is a Friday. On a Friday a slightly different pattern is expected, with an earlier starting evening rush hour and a lower peak. When the historical pattern of the above mentioned week days is used to predict the Friday pattern, the pattern has to shift forward in time and be scaled down. The graph in figure 58 shows the historical pattern and the measured pattern in one graph.

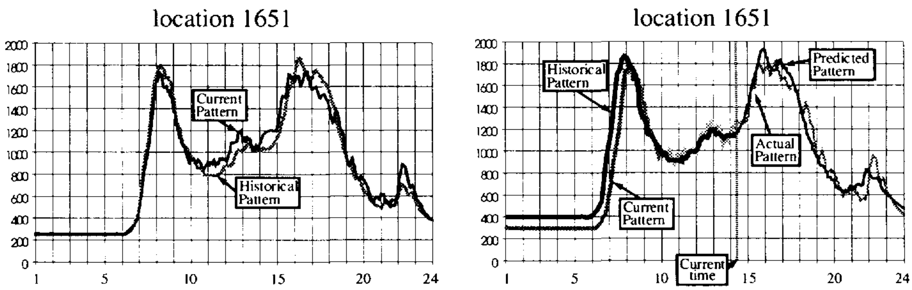


Figure 58: Historical departure time pattern (December 1,3,7,8,9 1992) and measured departure time pattern (December 4, 1992). Right side figure with curve fitting.

The left graph in figure 58 shows the historical pattern and the pattern as it was actually measured on Friday, December 4th, 1992. The figure shows that the morning rush hour matches fairly well without any scaling or shifting. The evening rush hour starts slightly earlier. In the figure on the right, the current time is assumed to be 14:15h. This means that the current pattern, now the grey line, is only known until 14:15h and thereafter shown as thinner after

14:15h, indicating that the pattern is unknown from this point. The line is shown anyway to give an impression of the performance of the method. The first one to two hours after 14:15h must be predicted. The prediction is done by fitting the historical curve, the solid line, with the measured curve, the grey line. The measured curve is only known until 14:15h. The solid line after the vertical dotted line (14:15h) represents the predicted pattern. Every time new data becomes available, the dotted line shifts to the right, and the historical pattern is fitted again to the measured pattern, which now includes the new data.

The method of line fitting was tested using several on ramps at the Amsterdam beltway. In general similar results were achieved to those described in the example of figure 58. Results might be improved by enhancing the database. Different days show different departure time patterns. The weekends especially are completely different from weekdays, but, as the example shows, Fridays also show slightly different departure time patterns.
The constant stream of data makes it fairly simple to maintain a large database with departure time patterns, differentiated for different days and different weather conditions. When this database is kept up to date continuously, it is a valuable and fairly accurate source of departure time information.

6.3.2.4 Calculate Start OD-matrix:

This module performs a matrix multiplication. The start OD-matrix is calculated based on the first period of the split-matrix and the current densities. The start OD-matrix represents the information to reproduce the current state of the network instantaneously.

6.3.2.5 Calculate Future OD-matrix:

This module calculates the Future OD-matrix, representing the traffic that will enter the network for the coming periods. This matrix is calculated for each origin using the split matrix and the departure time functions. This module performs a matrix multiplication.

6.3.3 Speed-Density Functions

Each type of link in the network can have its own speed-density function. The capacity, speed and flow of a link are influenced by the parameters of the speed-density function. So the effects of accidents, weather changes, roadworks, etc. are incorporated in these functions.
The correct shape of a speed-density function has been, and still is, subject to many discussions. It is not possible to give an average speed-density function that is applicable for all types of roads. As shown in section 5.3.2, even for one link a speed-density function can change due to weather changes, percentage of trucks, time of day, etc. All these influences can be incorporated in the delay function. Section 5.3.2 demonstrated that a general applicable function is not satisfactory, and that the parameters of a separate function for different road types show a rather significant variance. A good strategy might be to adapt the function continuously to the current measured situation. The function used in 3DAS is the modified Smulders function. This function has four parameters, α, β, v^{max}, and ρ^{crit}.
The parameters of the speed-density function can be estimated for each link in the network using historical data. This historical function is the default function for that link. The relationship between speed and density is measured by using the real-time data from induction loops. This data is used to adapt the function to current conditions. In principle it should be possible

to incorporate incidents and weather changes in the model solely by adapting the speed-density functions to the real-time data.
In theory it is possible that a sudden change at a certain location is detected, because the speed-density function suddenly changes rapidly for that location. Which means that the model is responding to a sudden change, without any knowledge of the cause of the disturbance (weather change or an incident).

The function is fitted to the measurements using a discounted least square method. The function is updated for every new prediction. The last measurement has more influence than older measurements. The starting point for each minimization is the default (historical) function. The function z, (eq 82) is minimized using a steepest descent method.

$$Min\ z(f) = \sum_{t \le \tau} w_t q_t [f(\rho_t) - v_t]^2 \qquad \text{(eq 82)}$$

Where w_t represents the weight for the measurement at time t. Recent data has more weight than older data. Typically the function w is described by μ^{t-P} with $\mu < 1$.

The flow at time t is represented by q_t, $f(\rho_t)$ represents the speed according to the speed-density function value at ρ_t. The matching measured speed at time t is represented by v_t. The discounted least square is summarized for all periods proceeding the current period τ. The parameters to be solved are the parameters of the modified Smulders function, α, β, v^{max} and ρ^{crit}.

6.4 Hardware and Software Requirements

This section describes the hardware and software requirements for the system. The section is divided into three parts. The first part describes some relevant aspects of the MCSS monitoring system and the consequences. The second part describes the software and the third part the hardware requirements for real time applications.

6.4.1 Induction Loops

The induction loop data used in this chapter was derived from the MCSS system in Amsterdam. This data is used primarily for this signalling system and was not specifically designed for more advanced systems. The aggregation of the data takes place on road side computers to alleviate communication lines. Speed is aggregated using the normal average, and flow is converted to hour flow and stored in a variable with a unit of fifty cars per hour. This means that flow is actually stored as $^5/_6$ cars/minute. This awkward conversion was made to store the number in only one byte and still use hour flow. Further, there seems to be a minimal flow of 250 vehicles per hour. The reason for this is historical and was probably done to provide an indication that the loop is still operational.
Although the MCSS system is very successful as a signalling system (see section 2.4.1), it is poorly designed as a data collecting system. Using a harmonic averaged speed would improve the results, and using flow per minute would be more logical and more accurate and fit easily in one byte. Further, information concerning the length of vehicles is important for calculating density on a road segment.

It is possible to collect data at a lower level by inserting special hardware units in the road side computers. The lane, time (ms), speed, and vehicle length are stored on a tape for each individual passage. This data is not on-line available as the tapes have to be collected at the end of the day. Using this data, however, it is possible to demonstrate the impact of different aggregation methods. Three different aggregation methods can be demonstrated. The speed of car i is denoted by v_i, the length of car i by l_i, the time period by T, and the number of cars during the time period by n. The average speed for time period T is denoted by u, the flow by q, and the density by ρ.

- The first method uses the MCSS method, using the awkward representation of flow, and normal averaged speed. The density is calculated by dividing the flow by the speed.

$$
\begin{aligned}
q &= \mathrm{int}\left[\frac{(n \cdot 60)}{50}\right] \cdot 50 \\
u &= \frac{1}{n}\sum_{i \in T} v_i \\
\rho &= \frac{q}{u}
\end{aligned}
\qquad \text{(eq 83)}
$$

- The second method uses cars per minute for the flow, and the harmonic speed. The density is calculated by dividing the flow by the speed

$$
\begin{aligned}
q &= n \cdot 60 \\
u &= n / \sum_{i \in T} \frac{1}{v_i} \\
\rho &= \frac{q}{u}
\end{aligned}
\qquad \text{(eq 84)}
$$

- The third method is the most sophisticated method and uses the lengths of the passing vehicles. This method has been described by Polak 1994 [70]. This method does not use cars as the unit but *meter of cars*. Polak introduces a new parameter called *speed per meter of vehicle*, and a parameter called *production*, representing the sum of the lengths of the vehicles passing during the aggregation period (here one minute) divided by the length of the aggregation period.

$$
\begin{aligned}
q &= n \cdot 60 \\
u &= \sum_{i \in T} l_i / \sum_{i \in T} \frac{l_i}{v_i} \\
\rho &= \frac{\sum_{i \in T} l_i}{T} \cdot \frac{1}{u} \cdot \rho^{max}
\end{aligned}
\qquad \text{(eq 85)}
$$

The parameter ρ^{max} represents the maximum density

Figure 59 shows the three methods in three different plots for the same location. The measurements are displayed as a speed-density diagram.

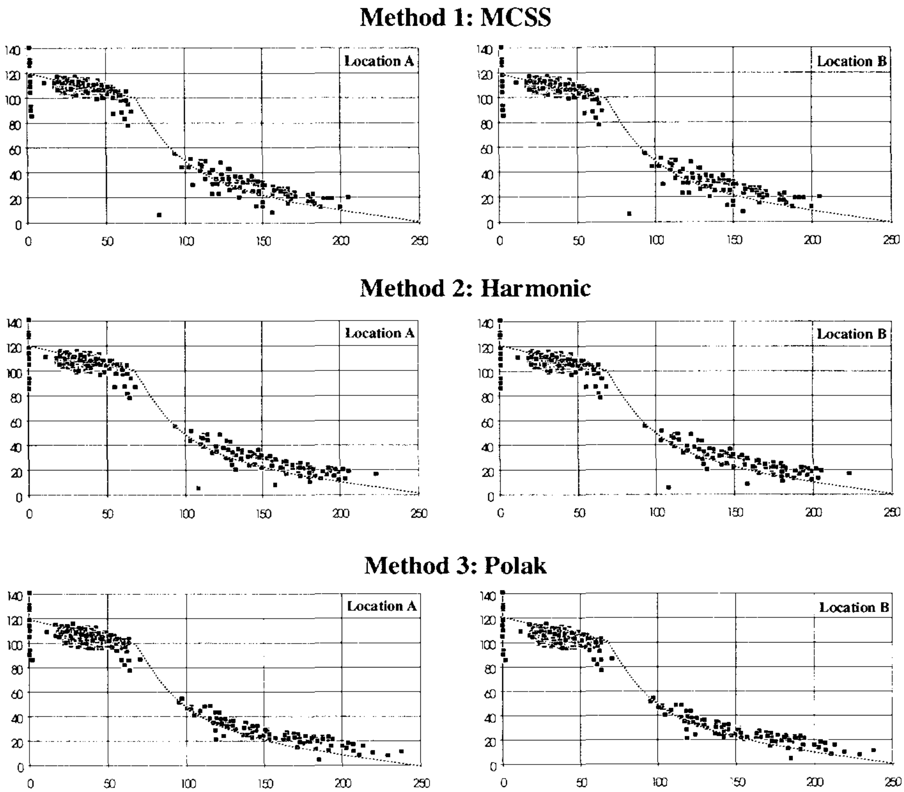


Figure 59: Three different aggregation methods to calculate density and speed from induction loop data at two different locations, location A on the left side of the figure and location B on the right side, the locations are both on the A13 in the Netherlands, measured at 23th of March 1990.

The plots show clearly that the relationship between density and speed improves with a better aggregation method. The most relevant improvement is caused by using a harmonic averaged speed instead of a normal averaged speed. The third method improves the situation only slightly, but shows the largest improvement when there are many trucks in the flow. The method is still experimental but shows great promise for improving the quality of induction loop data.

6.4.2 Software and Visualization

The 3DAS model is implemented on a graphic workstation. The software makes testing and validation of the model possible. The program is implemented in the UNIX[1] operating system and the X-window graphic environment, and makes extensive use of graphic presentations.

The use of graphics is not just to present 'nice pictures' it provides very important benefits. Graphic presentation of dynamic assignment is a basic requirement for analysing the results. The graphics provide a lot of insight into the behaviour of the model and unexpected results frequently become clear because of the visualization. This section is therefore not meant to describe the software developed, but to emphasize the importance of visualization in dynamic traffic assignment and to present some ideas for visualization.

The advantages of interactive computer graphics have become evident in several fields of research, and are nicely stated by Foley et al. [23]:

> *"Graphics provides one of the most natural means of communicating with a computer, since our highly developed 2D and 3D pattern-recognition abilities allow us to perceive and process pictorial data rapidly and efficiently. In many design, implementation, and construction processes today, the information pictures can give is virtually indispensable. Scientific visualization became an important field in the late 1980s, when scientists and engineers realized that they could not interpret the prodigious quantities of data produced in supercomputer runs without summarizing the data and highlighting trend and phenomena in various kinds of graphic representations."*

The 3DAS software package is window based and every visualization is presented in a separate window. Six aspects of the model are visualized in six windows described below:

- The *main window*. This window shows the network as nodes and links. Different types of links are presented in different colours. Links can be selected and attributes can be changed. Zooming and panning is provided. The main window comes with several small pop-up windows to select several graphic settings and to select settings of the model, for example, the pathfinding method and the assignment method. Several previous calculations can be saved in memory and/or to a file, which makes it possible to compare different calculations.
- The *departure function window*. This window shows the departure function for a selected origin. The function can be edited interactively. A new function can be loaded from a file and saved to a file.
- The *travel time functions window*. This window shows the travel time function for the selected link. The parameters of the function can be changed interactively, and different types of functions can be chosen.
- The *link information window*. This window shows the attributes of the selected link in time. Speed, flow, density or travel time is displayed for all periods. It is possible to show previously saved calculations in the same window simultaneously.
- The *trajectory window*. This window shows the trajectories of a previously specified path for several departure periods. The trajectories are saved for each iteration, and the convergence towards a certain trajectory can be observed.
- The *movie window*. This window presents the flow in the network in a movie-type fashion. Density on a link is shown using different colours, usually a colour palette is used, which presents a low density as light blue. With increasing density, the blue darkens, until critical density is reached. Critical density occurs at maximum flow, the colour then changes to light orange, which darkens to dark red with increasing density. Light orange represents the start of congestion and dark red represents maximum density with a very low speed.

Examples of these windows are given throughout this dissertation.

---

1. UNIX is a trademark of AT&T Bell Laboratories.

6.4.3 Hardware

A fast computing time is essential for predictions, a prediction is clearly useless when the computing time exceeds the required prediction time.
Each module described in this chapter can operate independently. This makes it possible to divide these tasks over different computers. The 3DAS prediction, however, will take the largest computational effort, and from the start of this study it was clear that for real-time applications significant computational power would be required. At the faculty of Applied Physics of Delft University interrelated research is being carried out towards the design of special purpose hardware for certain classes of problems [29]. Path finding in large networks is part of such a class. Exploiting the inherent parallelism in determining and assigning shortest path trees, it turns out that a parallel architecture of p processing units is an efficient way to speed up the assignment computations. An almost linear speed-up is possible by adding a fast transmission facility, independent from p.

The above mentioned parallel computer has been constructed and is called the Linear Processor Array (LPA). It consist of 16 processor boards, that are linearly inter-connected. Each board contains an Intel i860 RISC processor. Communication between two boards takes place using FIFO's. These devices have the functionality of a buffer and allow asynchronous communication with a speed of 160 Mbytes per second. Each board can transfer information to its neighbours independently. This allows fast transmission. The LPA is connected to a host computer which provides a user friendly work environment and access to the LPA for everyone who is connected to the Internet, and has permission to use it.

During the development of the 3DAS model, possible implementation of the algorithm on the LPA was kept in mind, and towards the end of this research the special purpose computer became operational. The 3DAS model was implemented successfully on the LPA, and several test have been done. The Amsterdam case described in chapter 8 reports on the performance of the LPA.
Only minor changes in the 3DAS software were necessary to run the program on the LPA. The problem is split between the boards of the LPA by dividing the OD-matrix. The total number of OD-relations is divided into equal groups and each board calculates the shortest paths and the assignment for its group. This is realized by starting the 3DAS program on each board. Based on the board number and the total number of boards each program knows which OD-relations it has to compute. When all boards are done, the results (load per link per period) are broadcasted through the FIFO's to one board, which adds all the loads *while* they arrive. The final result of this iteration becomes available at this board, which broadcasts this new situation to all the other boards. Based on the newly arrived loads, each board starts the next iteration. The only overhead existing in this scheme is the transmission of the data. Transmission time is negligible because of the very fast FIFO's, which makes the linear speed-up possible. The LPA is described in detail in [29].

6.5 Conclusions

This chapter described a structure to use the 3DAS model in an ATPS application and in an ATMS application. Little extra effort is required to use the 3DAS model for ATPS applications. Traditional *static* rush hour OD-matrices can be used, and made dynamic using departure time patterns. Several extra modules were required for ATMS applications to provide the

3DAS model with input and to make the model adaptive to prevailing conditions. One of the modules, the OD-split matrix module, is a research project by itself and is described elsewhere. Some ideas and test results concerning the departure time patterns module, and the speed-density function module were described in this chapter.
The departure time patterns module uses historical data to predict future patterns. Empirical research showed that departure patterns are fairly consistent each day. This gives the opportunity to use the historical pattern as a basis for prediction. Good results were achieved with the on ramps in Amsterdam by fitting the historical pattern to the current measured pattern.
The section about speed-density functions presented a method to incorporate external influences in the model by adapting the speed-density functions. It is possible to adapt these functions with data from induction loops. It is not necessary to be aware of all types of weather circumstances or incidents. A suitable function, which is easily adaptable, is introduced as the 3DAS function. This function is written as a speed-density function and has a strictly increasing relationship between density and travel time. The matching flow increases with increasing density until a certain critical density is reached, then the flow decreases again, representing congestion.
The last section of this chapter showed that the relationship between speed and density could be further improved by using better aggregation methods for induction loop data. A short impression was given of the software and the research towards a special purpose computer to apply the model in real time situations. A fast computing time is essential for short term predictions. Such a prediction is clearly useless when the computing time exceeds the required prediction time. Research towards special purpose hardware should solve any computational problems.

The proposed ATPS structure is tested in the Washington Metropolitan area in chapter 7. The proposed ATMS structure is tested, including the LPA, on the Amsterdam beltway in chapter 8.

CHAPTER 7

APPLICATION OF 3DAS AS ATPS WASHINGTON

7.1 Introduction

This chapter describes a research project in which the dynamic assignment model, 3DAS, is used as an Advanced Traffic Planning System (ATPS) [76]. The study had two objectives. The primary objective was to find answers to the following three questions:

I. Can *dynamic* assignment be used for planning purposes?

II. Does *dynamic* assignment have an advantage above *static* assignment?

III. Is *dynamic* assignment a useful tool to investigate the effects of Dynamic Traffic Management (DTM)?

The secondary objective was to gain insight into the possibilities and problems of using 3DAS for large networks.

The model was tested using data from the south western part of Washington D.C. in the USA. This area was chosen because it offers a heavily congested urban network, with several rerouting possibilities. Within the area there are several ramp metering installations in operation, and parts of the freeways are monitored by induction loops.
The data used in this research was obtained from the Virginia Department of Transportation (VDOT), and the Metropolitan Washington Council of Governments (COG). A small portion of the study area is monitored by induction loops, from which one minute data was used to derive the departure functions and to validate the calculation results.

The study was only intended to be an example of using dynamic assignment as a planning tool. The results are not suitable for any serious planning decisions, due to a lack of data, and the short study time. The results do however permit us to decide on the *usefulness* of 3DAS for planning purposes.

One static scenario and three dynamic scenario's were calculated; a morning rush hour, a scenario with several ramp metering installations, and a scenario with an incident. The results produced by the model for these scenario's are reported below.

7.2 Research Approach

The following research approach was set up in accordance with the objectives. The first objective consisted of three questions:

Can dynamic assignment be used for planning?

Dynamic assignment can be used for planning if, given a network and a traffic demand, it can predict a correct distribution of traffic flow. As for long term purposes traffic demand will represent average demand, the expected traffic *distribution* will also be an average. This is in contrast to ATMS applications where the results should be based on the actual situation of the moment, including temporary disturbances such as incidents or roadworks. The average traffic demand and a measured traffic distribution, averaged over a longer period, are required to validate the 3DAS model.

Does *dynamic* assignment have an advantage over *static* assignment?

There are several (well known) problems with static assignment models. A static assignment model:

- gives incorrect results for congestion. As traffic is assigned along the complete route, a car can contribute to more than one traffic jam at the same time.
- cannot show correctly the effects of variable traffic demand.
- cannot show correctly the effects of temporal disturbances, such as road works or accidents.
- cannot predict queue lengths, and how a growing queue can limit the capacity of upstream junctions.

The research project looked at whether dynamic assignment can solve these inconsistencies, and how it might improve decision making for planning.

Is dynamic assignment a useful tool for investigating the effects of Dynamic Traffic Management?

3DAS was extended to model several DTM instruments, such as ramp metering, rerouting and tidal flow. Two scenarios are executed to answer this question. The first scenario considered several ramp metering installations, and the second scenario considered an accident on one of the freeways. The effects of diversion measures are reported for the second scenario.

Since the network used for this study is fairly large, the secondary objective of this research, which was to gain insight into the possibilities and problems associated with dynamic assignment applied on larger networks, was also satisfied with the above research approach.

7.3 Data

The study area covers the eastern part (Virginia part) of the beltway in Washington D.C. The major interstates are the I-95, I-395, I-66 and the I-495, a large part of the arterial network is also included.

7.3.1 Network

Figure 60 represents the network used for this study. The network consists of 857 nodes and 2086 links. There are 180 zones. Most freeway intersections are represented in a fairly detailed way. An example of two of these intersection details is given in figure 60. Each line in the figure depicts a separate one directional road consisting of one to four lanes. Some of the longer link were divided into two shorter links. In general no problems were expected with the underestimation of travel time on long links (see section 5.2.1), because demand is represented by a rather smooth pattern.

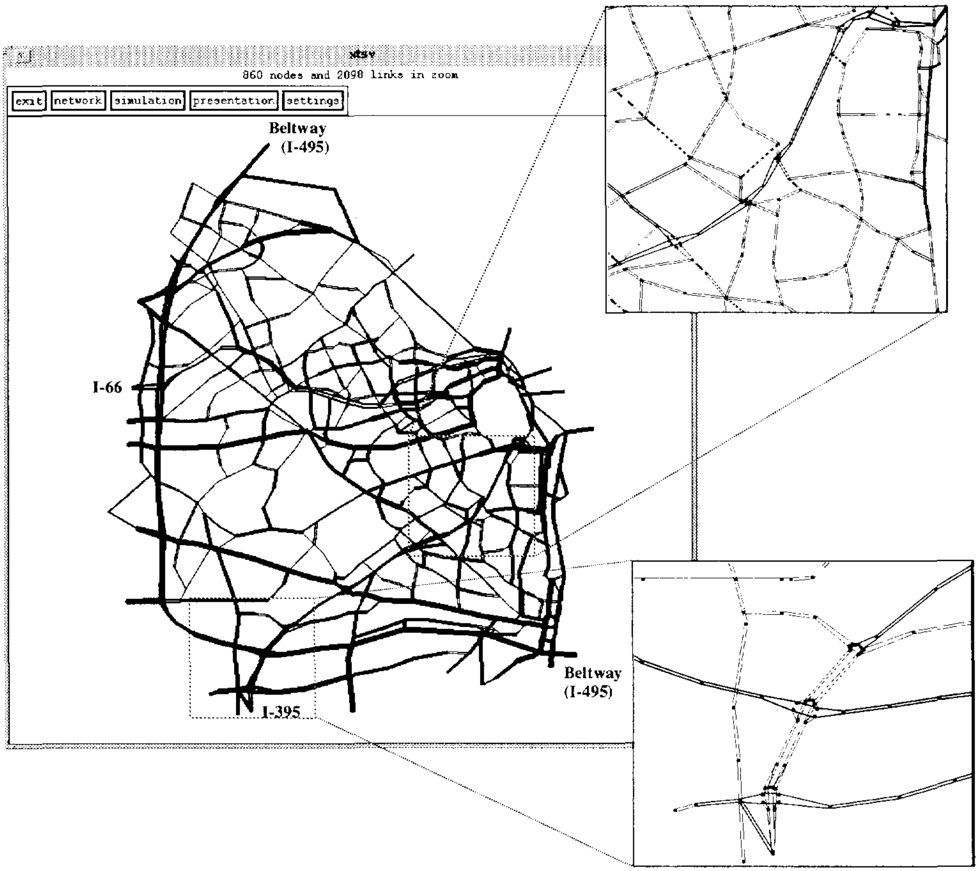


Figure 60: The study network

The 2086 links were divided into 13 types, each type represents a certain road type, all the links in one type have the same attributes, these are given in Table 7. The attribute for capacity is not given but derived from maximum density, maximum speed, and the speed-density function.

Table 7: Network Characteristics

Type	Description	Max Speed (km/hr)	Max Dens (veh/km)	Capacity (veh/hr)
1	Beltway, 4 lanes	100	440	7250
2	3 lanes part of the I-395	100	330	5800
3	2 lanes part of I-395 and 2 lanes part of I-66	100	224	4390
4	3 lanes part of the I-66	100	330	5800
5	4 lanes part of the I-395	100	440	7250
6	freeway intersection (right turn)	70	110	1460
7	freeway intersection (left turn = circle)	40	115	850
8	bridges in centre over the river Potomac	50	224	2200
9	on and off ramps	50	115	1060
10	Unused road type	n.a.	n.a.	n.a.
11	1 lane arterials	35	112	730
12	2 lanes arterials	60	224	1750
13	3 lanes arterials	65	336	2740
14	4 lanes arterials	65	448	4000

7.3.2 OD-matrix

The network is not accompanied by a matching *dynamic* OD-matrix, so the OD-matrix must be constructed from other data sources.

The best OD-matrix available was a (static) 24-hour matrix that covered a much larger area. This OD-matrix was a result of a study carries out by the *Metropolitan Washington Council of Governments* (COG). The OD-matrix for the study area was extracted from this OD-matrix. Departure functions were used to make the OD-matrix dynamic. A departure function gives, for each OD-pair, the portions of the amount of traffic departing in each period. These departure functions can be estimated and calibrated using link volume data.

The large network covered 5983 nodes and 18104 links.

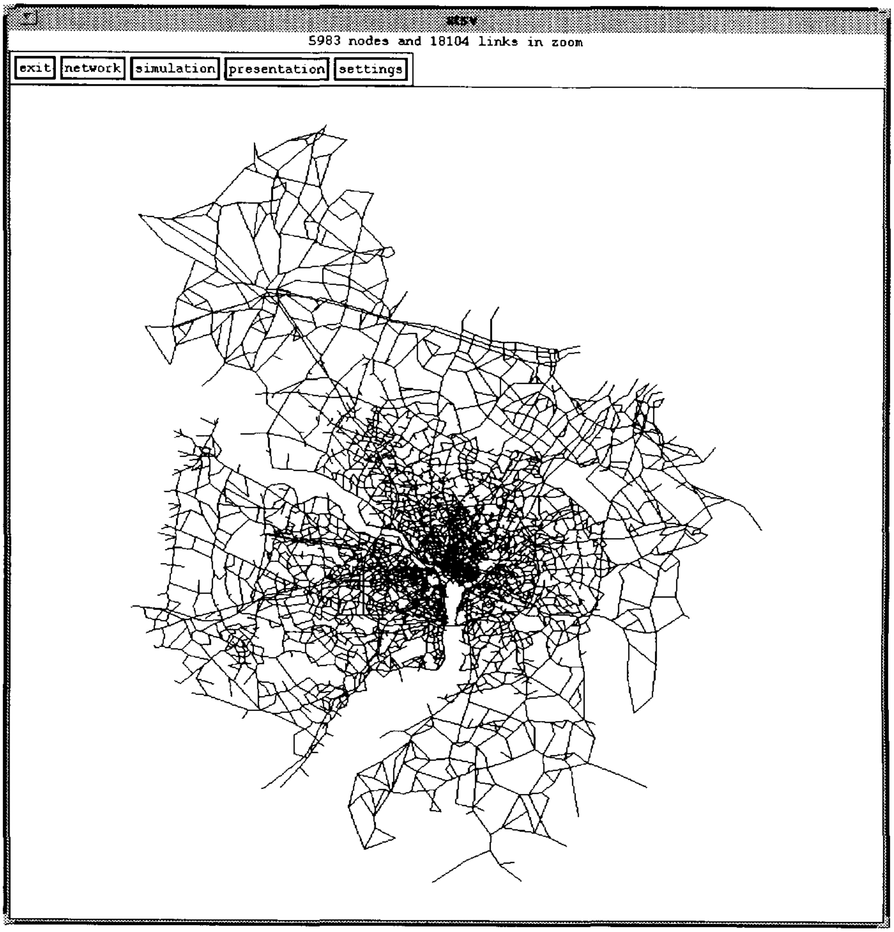


Figure 61: The large network

The COG study [1] was carried out taking 1990 as the base year, and covered 293 districts (1478 zones) covering the following jurisdictions:

- District of Colombia, (DC)
- Montgomery County, MD
- Prince George's County, VA
- Arlington County, VA
- Alexandria, VA
- Fairfax County, VA
- Loudoun County, VA
- Prince William County, VA
- Frederick County, MD
- Howard County, MD
- Ann Arundel County, MD
- Charles County, MD

VA = Virginia, MD = Maryland.

The model used by COG for trip generation, distribution and mode choice was a gravity model that calculated at district level. The districts were then "split" into zones via land-use factors. These land-use factors were based on households and groups-quarter population for productions, and they were based on office, retail, industrial and other employment for attractions. The resulting OD-matrix had 1478 zones.

The network used for this research (figure 60) is only a small part of this large network, so the OD-matrix for the smaller network (180 zones) had to be derived from the big OD-matrix (1478 zones). Zooming in, the study network becomes recognizable. All trips made *within* the study network were easily derived. All trips entering, leaving or passing through the study network were derived using a selected link analysis. Links marked with a dot in figure 62 indicate the selected links.

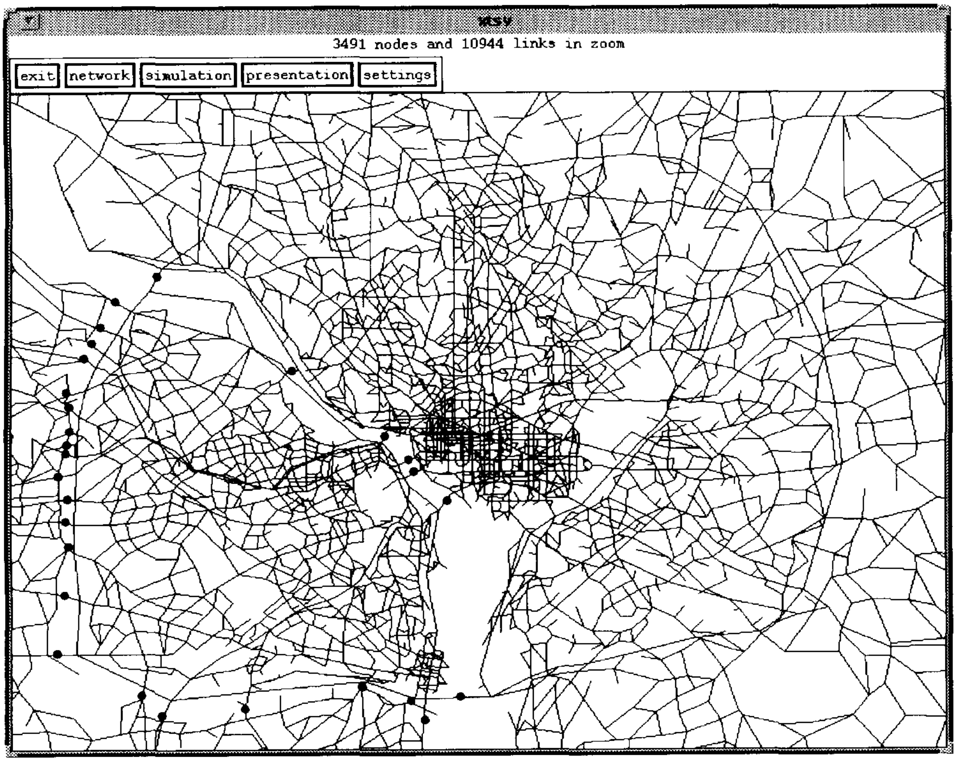


Figure 62: Zoomed in part of the large network

To perform the selected link analysis, the OD-matrix was assigned to the network using a static "All-or-Nothing" assignment. Shortest pathfinding was done using the *calculated* speeds in the network. These *calculated* speeds were derived from the static assignment produced by the Washington Council of Governments. Origin and destination were stored for all OD-pairs crossing the selected links. The entering and exiting links then become new origins and destinations, and the trips are summed. Using this method all the entering and exiting traffic can be aggregated to the links with which it exits or enters the subnetwork.
Deriving an OD-matrix for the subnetwork using this method had one major drawback. No alternative routes were chosen for OD-pairs, because an All-or-Nothing assignment is used. To minimize the effects of this problem some links were added to the subnetwork to allow a diversion for some origins to different links to enter the network.

7.3.3 Induction Loop Data

The Northern Virginia Traffic Control Centre controls a part of the freeway system in Northern Virginia. The freeways covered are the I-66 and the I-395, and these are equipped with several hundreds of induction loops. One minute data of a fixed portion of these induction loops can be downloaded onto a data tape. Unfortunately the Traffic Control Center is not yet fully equipped, and the downloading of data from induction loops is therefore not easy. Only one tape (one day) was available for this research. Although the traffic patterns for this one day are not sufficient to derive any statistical information, it was the best data available.
The tape used for this study contained data measured on Monday the 7th of December 1992 from 18.00h until 11.00h the following day. The number of passed vehicles are registered and downloaded for each minute onto a tape, and this tape was made available for research purposes. Unfortunately the tape contains many errors and at several locations only one of the three or four lanes was read. Nevertheless the data from several locations were useful for this research.
The following graphs give an impression of the traffic pattern at several locations on the I-66 and the I-395. The x-axis shows the time in hours. Registration started at 18:00h and lasted until 11:00h the next day. The y-axis shows the flow in vehicles per hour. Two graphs are shown for each freeway and each direction. The first graph is located somewhere at the beginning of the freeway, the second graph is located somewhere near the end of the freeway. The locations of the induction loops are displayed in figure 67 for the I-66 and in figure 68 for the I-395.

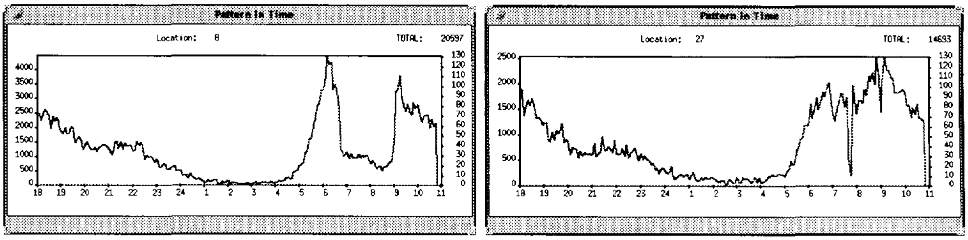


Figure 63: The traffic flow from 18:00h until 11:00h (next day) for locations 8 and 27 on the I-66 Eastbound.

The graphs in figure 63 show clearly that the rush hour started at ±05.00h. At location 8 flow increased in approximately one hour to 4500 veh/hr. At ±06:00h some kind of congestion occurred and flow dropped rapidly (possibly an incident). After ±09:00h flow increased again. The end of the rush hour was at approximately 11:00h. At location 27, which is further downstream the I-66, the flow increased to ±2500 veh/hr.

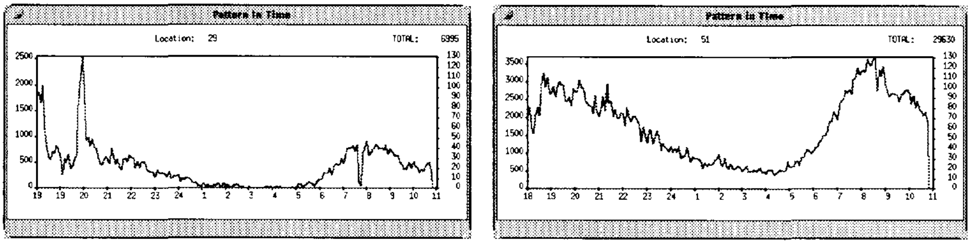


Figure 64: The traffic flow from 18:00h until 11:00h (next day) for locations 29 and 51 on the I-66 Westbound

The two graphs in figure 64 show that flow on the I-66 westbound was significantly lower, and that no congestion occurred in this direction. At location 29 flow increased to ±2000 veh/hr. At location 51 flow increased to ±3500 veh/hr.

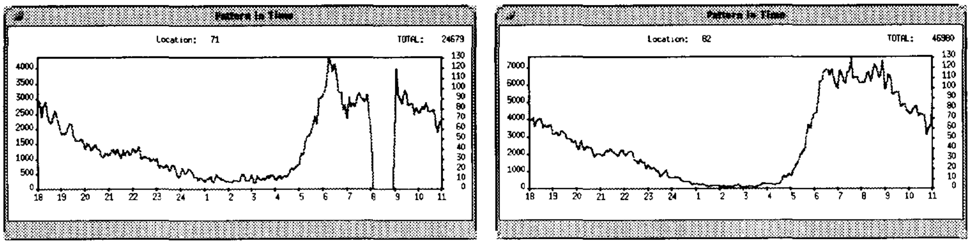


Figure 65: The traffic flow from 18:00h until 11:00h (next day) for locations 71 and 82 on the I-395 Northbound

Roughly the same pattern is found for the I-395 (figure 65) as for the I-66. The rush hour starts at 05.00h. At location 71 flow decreased after ±06:00h due to congestion. Between 08:00h and 09:00h there was probably an error in detection. Further downstream the I-395 at location 82 no congestion occurred. Flow increased to 7000 veh/hr (4 lane section).

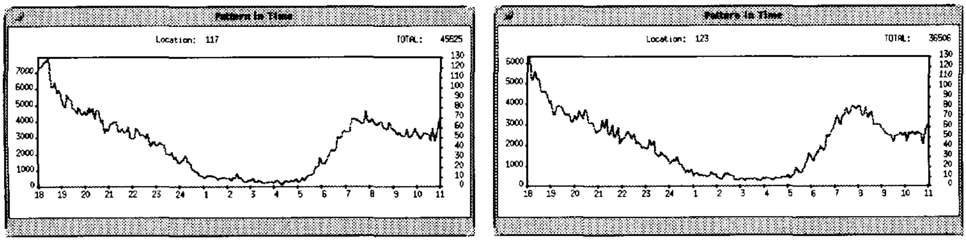


Figure 66: The traffic flow from 18:00h until 11:00h (next day) for locations 117 and 123 on the I-395 Southbound

On the I-395 Southbound in figure 66 there was no congestion. At location 117 flow increased to ±4000 veh/hr, and also at location 123 to ±4000 veh/hr.

The graphs in figure 65 show higher flows in the morning rush hour (08:00h) than in the evening rush hour (18:00h). Figure 66 shows the opposite, higher flows in the evening rush hour than in the morning rush hour. This may be explained by commuting traffic, entering the city in the morning and leaving in the evening. The I-66 (figure 65 & 66) shows the same pattern, although not as clearly.

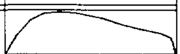
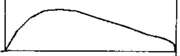
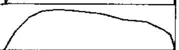
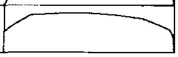
It was decided to calculate a morning rush hour from 05.00h to 11.00h using the induction loop data. This time period captures the total morning rush hour, and the graphs showed that before 05:00h the network is still reasonably empty. This has the advantage that the calculations can be started with an empty network.

7.3.4 Departure Time Functions

To use a static OD-matrix as a substitute for a dynamic OD-matrix, departure time functions are required. A departure time function is a discrete function, which determines, for each period, the percentage of the OD-value that departs during that period. Induction loop data can be used to derive these departure time functions.
One departure function for all OD-pairs will not give a realistic representation of the dynamic OD-matrix for a rush hour. The departure functions of individual OD-pairs can be quite different. The graphs in section 7.3.3 show that traffic on the I-66 travelling westward (figure 64), and traffic on the I-395 travelling southward (figure 66), have a lower volume in the morning rush hour and a higher volume in the evening rush hour, and depart according to a different departure time function. This requires at least a different departure time function for trips going into Washington and traffic leaving Washington.
The OD matrix was split into four major trip types for this study, and for each type a different departure function was used. The different types are described in Table 8.

In total the 24-hour OD-matrix contains 293930 trips. It can be derived from the induction loop data that during the rush hour, from 05:00h until 11:00h, approximately 35 percent of the trips specified by the matrix departs. The total surface of the departure time function determines the percentage of the OD-matrix. So for each type the total percentage of trips that departs for that type is influenced by the total surface of the function.
Table 8 gives the total amount (trips) and the surface of the departure time function (%) and the form of the function.

Table 8: Departure Tables

Type	Description	Trips	%	Form
1	Entering Washington or passing through	88625	36.9	
2	Entering the Washington on the freeways (I-66 from the east and I-395 from the south)	88527	13.9	
3	Exiting Washington on the freeways (I-66, I-395)	19625	33.4	
4	Internal traffic within the network	97153	35.0	

For type 1, the total number of trips entering Washington or passing through was 88625 trips for the total 24-hour period. From this total 36.9% was assigned to the rush hour period. The trips depart from their origin according to the departure time function given in table 8.
Table 8 shows a low percentage for the trips entering the network on the I-66 (east) and the I-395 (south), this is due to a large overflow on these freeways. Using a percentage as low as 13.9% a reasonable reproduction of the induction loop data can be achieved. This overspecification of the flow in the OD-matrix was partly due to the fact that an All-or-Nothing assignment has been used to derive the OD-matrix, and partly to possible overspecification in the original matrix.

Using only four different departure functions will give a rather rough reproduction of the traffic patterns. This choice was made because of the lack of data to derive more departure functions and a lack of data for evaluating the results.

7.4 Modelling and Calibration

It was decided, based on the data available, to calculate a morning rush hour from 05:00h until 11:00h, in 24 periods of 15 minutes. The total time span of six hours captures the total morning rush hour. The period length of 15 minutes was chosen for practical reasons to keep the calculation time within bounds. A period length shorter than 5 minutes is not to be recommended because it suggests detail which cannot be reached given the available data, while a period length longer than 20 minutes dissipates the internal dynamics of the traffic assignment too much.

To satisfy the objectives of this research four scenario's were considered.

I. The first scenario was designed to achieve a reasonable reproduction of the morning rush hour. The departure functions were calibrated using induction loop data, and the resulting flows were compared to the induction loop data. It was possible to reproduce the induction loop data at the beginning of a route by adapting the departure functions. A good result was obtained when the flow pattern further down stream that route still matched the induction loop data. The flow pattern could be tested at several locations on: the time varying form, the average height of the flow, and the moments of sudden changes in the flow. As however, we have only one day of induction loop data and no information about weather or incidents, the data does not represent an average flow pattern, thus only a rough reproduction of volume patterns can be expected.

II. The second scenario was a static equilibrium assignment, comparing the results to the dynamic assignment. The advantages and disadvantages of time variation were studied.

III. The third scenario introduced rampmetering on all ramps on the I-66 Eastbound and all ramps at the I-395 Northbound. Influences on queue length, travel time, and diversion behaviour were investigated.

IV. The fourth scenario introduced an accident on one of the freeways (I-66). Two different situations were calculated for this scenario. In the first situation the drivers were unaware of the accident. This was simulated using initial travel times for the section with the accident. In the second situation the drivers were assumed to be fully informed, an equilibrium assignment was used here.

The third and the fourth scenario investigated the possibilities of dynamic assignment for Dynamic Traffic Management. The input data used for these scenarios was the same as for the morning rush hour scenario (I). The departure time functions and the OD-matrix were not changed.

7.5 Hardware & Software

The model is implemented as an X-window program under the UNIX operating system, and is described in section 6.4.2. Several different computers were used to run the program. Either a Silicon Graphics 320VGX or INDIGO, an IBM RS6000, or SUN Sparc2, whatever computer was available at the "Laboratory for Scientific Visualization" at Virginia Tech.

On the Silicon Graphics 320VGX one iteration took approximately 5 minutes. All the OD pairs are assigned to their time dependent shortest paths in one iteration. For this study with 180 zones and 24 departure periods, this results in 28835 OD-relations per period ($180^2 = 32400$; not all OD-pairs exist!). For 24 periods these are 692040 OD-relations.

Large arrays of numbers on paper are difficult to interpret, so visualization of the results is very important. The 3DAS software displays several results in graphics. The pattern in time can be investigated for each link, and to get an overall impression of the traffic flows, build up of traffic jams, etc., the results can be displayed in a movie like fashion. Errors in the input or other anomalities are easily found using a good graphic representation.

7.6 Results

The *dynamic* morning rush hour scenario is described in section 7.6.1, the *static* morning rush hour scenario in section 7.6.2. The capabilities of the 3DAS model to cope with DTM instruments and changing conditions, such as an accident, are demonstrated in sections 7.6.3 and 7.6.4.

7.6.1 Morning Rush hour Scenario

A dynamic assignment was carried out based on the OD-matrix, the departure functions and the network attributes . Heavy congestion was found on the I-66 and the I-395 going into Washington. Small congestions were found at several locations on the beltway, and on certain arterials.
The "movie" representation shows quite clearly where the congestion starts, and how it evolves. The flow patterns at two locations along the I-66 (figure 67) and along the I-395 (figure 68) are shown, to give an impression of the results.

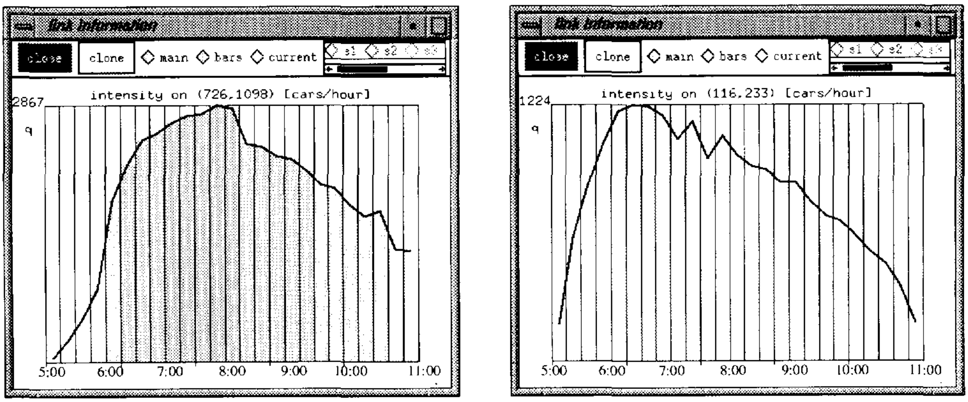
The graphs in figures 67 and 68 show the flow at four different locations on the I-66 and the I-395. The x-axis represents time and the y-axis shows the flow. Each bar represents a time period of 15 minutes. The height of the bars measures the flow, while the shade of the bars indicates the density. Light grey represents a low density and dark grey represents a high density. The difference between a low flow caused by a high density (dark grey) and a low flow caused by a low density (light grey) can be discriminated using this representation.

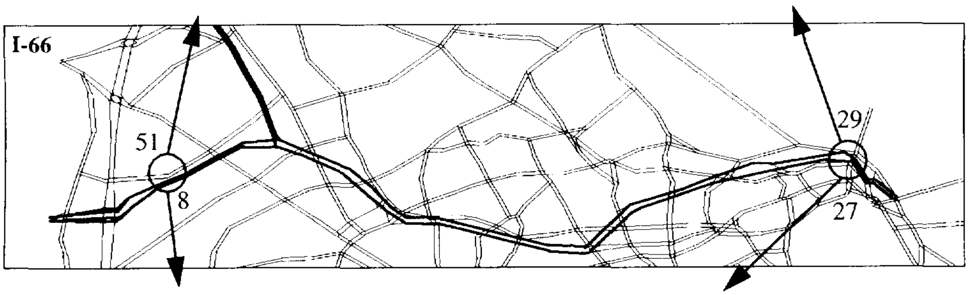
The graphs in figure 67 represent the same locations on the I-66 as used for the induction loop graphs in figures 63 and 64 in section 7.3.3; for figure 68, which represents the I-395, the matching induction loop graphs are figures 65 and 66.

Comparing the graphs in figures 67 and 68 with the induction loop graphs (figures 63,64,65 and 66) shows that a reasonable reproduction of the traffic distribution is possible. On the I-66 eastbound however, the induction loop data shows heavy congestion, with a low flow (almost zero); based on the low flow downstream, we may assume that there was probably some kind of incident during that day. A much higher flow downstream was found in the simulation. If there really was an incident the differences between the model and the induction loop data are explainable. The flow patterns on the freeways were compared with the induction loop data for more places than shown in figures 67 and 68.
Since there was no induction loop data available for the arterial system, it was not possible to validate these sections.

.

I-66 Westbound





I-66 Eastbound

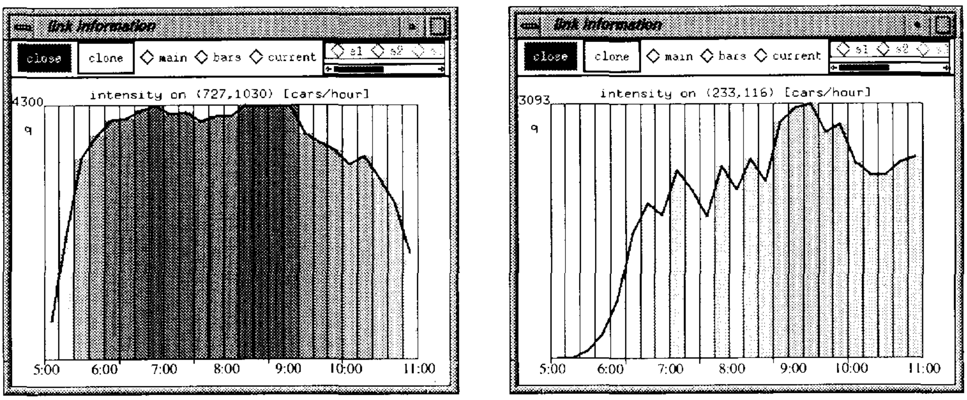


Figure 67: Flow(intensity) calculated using 3DAS for the rush hour from 05:00h to 11:00h at 4 locations (51,8,29,27) on the I-66.

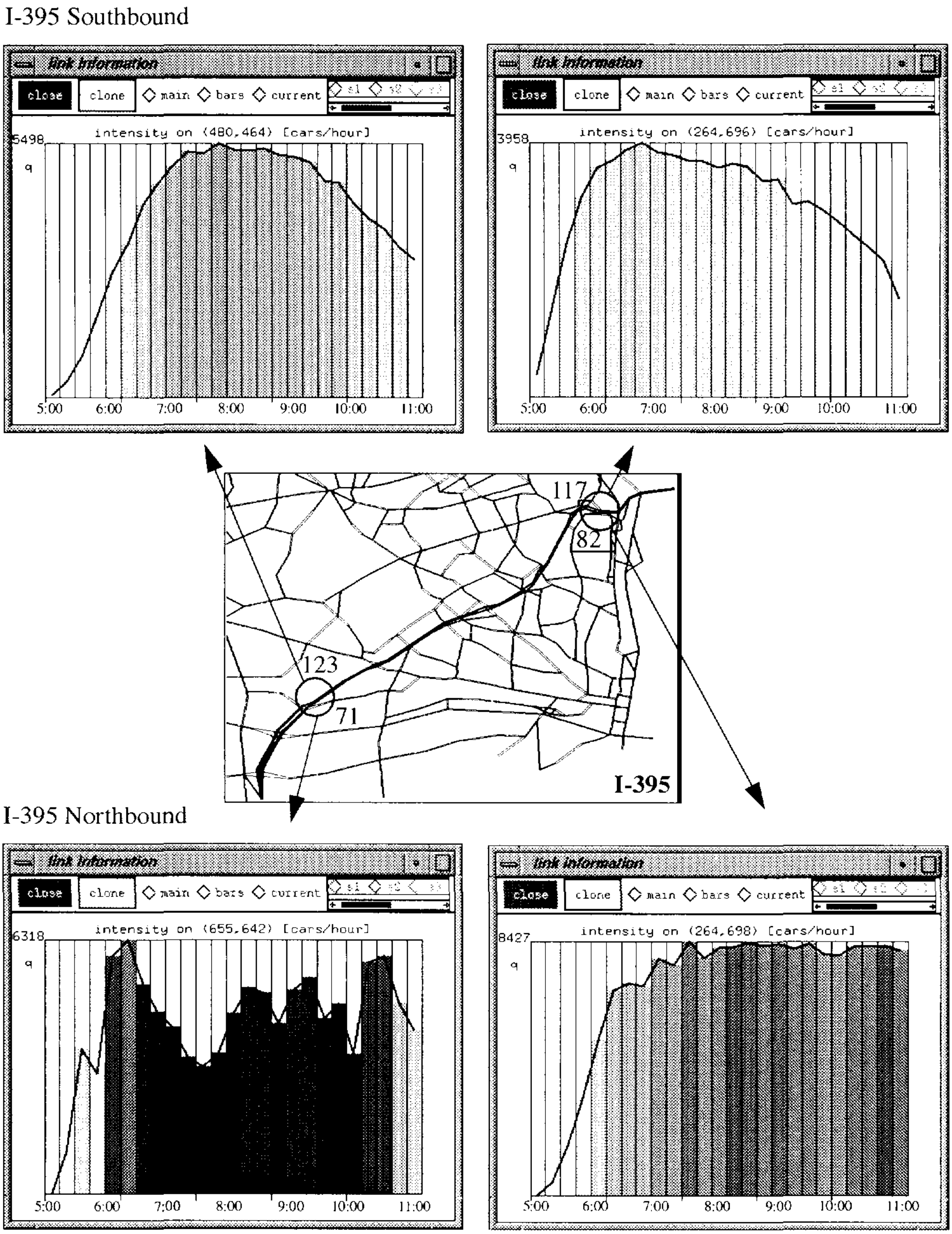


Figure 68: Flow (intensity) calculated using 3DAS for the rush hour from 05:00h to 11:00h at 4 locations (123,71,117,82) at the I-395:

7.6.2 Static Assignment Scenario

The 3DAS model was compared to a static user equilibrium assignment. The results show heavy overload for the static assignment on the entire I-66 and I-395. The results do not show where the congestion starts, i.e. where the bottleneck is located, and sometimes downstream occurrences of overflow do not exist in reality due to upstream congestion. The static assignment result is displayed in figure 69. In colours the flow is represented. Flow is represented in shades of grey, dark grey represents a high flow..

Figure 69: The static assignment scenario

For example, on the I-66 the *dynamic* assignment shows congestion halfway along the freeway, with a free flow situation downstream. The *static* assignment shows an incorrect congestion situation downstream as well. This situation could give rise to wrong planning decisions. Obviously the static assignment does not give any insight into the *development* of the rush hour as it evolves, and it does not give information on the length of traffic jams and how they evolve. On the positive side, the results of a static assignment are much easier to interpret. The amount of data produced by a static assignment is much less, and it therefore takes a shorter time to evaluate the results.

7.6.3 Dynamic Traffic Management Scenario

To test whether DTM strategies can be investigated using 3DAS, a scenario with rampmetering was created. Rampmetering was introduced for all on-ramps on the I-66 eastbound, and for all on-ramps on the I-395 northbound. Ramp metering was implemented as a simple maximum flow limit for all on-ramps. Since the shape of the speed-density functions dictates maximum flow (i.e. capacity), the maximum flow limit was achieved by using different speed-density functions.

The graphs in figures 70 and 71 compare the ramp metering scenario with the normal scenario. The graphs show *speed* as height of the bars and density as shade of the bar for the ramp metering scenario. The speed results for the normal rush hour scenario are displayed as a solid line in the same graph. The x-axis represents time.

A location upstream and a location further downstream on the I-66 were selected to display the results:

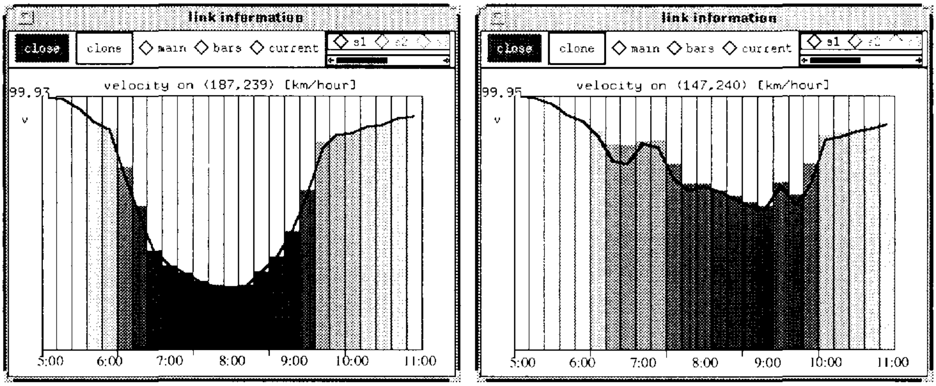


Figure 70: Velocity (km/hr) calculated using 3DAS for two locations on the I-66, the outcomes of the rampmetering scenario is represented by bars, the solid line represents the velocity calculated during the normal rush hour.

The graphs in figure 70 show that the impact of ramp metering is hardly noticeable on the I-66. The speed of traffic for the normal rush hour (solid line) is the same as during the metering scenario (bars). This is partly due to the low ramp flows.

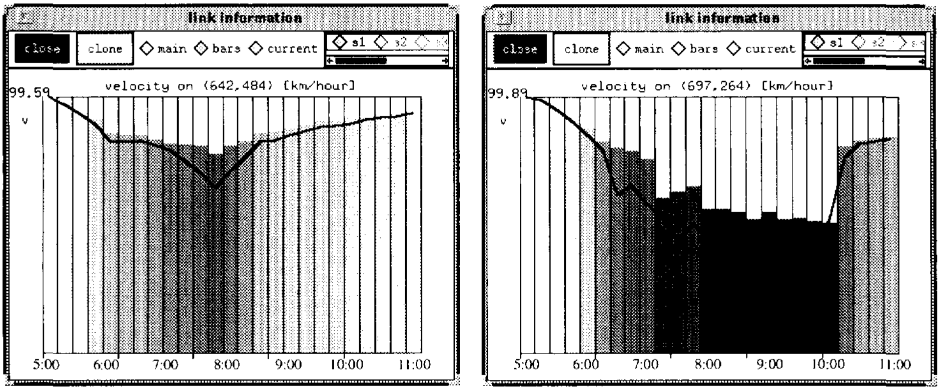


Figure 71: Velocity (km/hr) calculated using 3DAS at two locations on the I-395, the outcomes of the rampmetering scenario is represented by bars, the solid line represents the velocity calculated during the normal rush hour.

Figure 71 shows a location halfway along the I-395 eastbound, and one downstream the I-395 eastbound. The two graphs show that there is a noticeable impact on the I-395. Both the locations show a slight improvement in speed. In the left graph the temporal decrease in speed at 08:00h in the normal rush hour (solid line) is no longer there (bars). In the right graph, there is an improvement in speed over almost the total duration.

Figure 72 shows the impact ramp metering has on the arterial network, the figure shows a location at the end of the I-395.

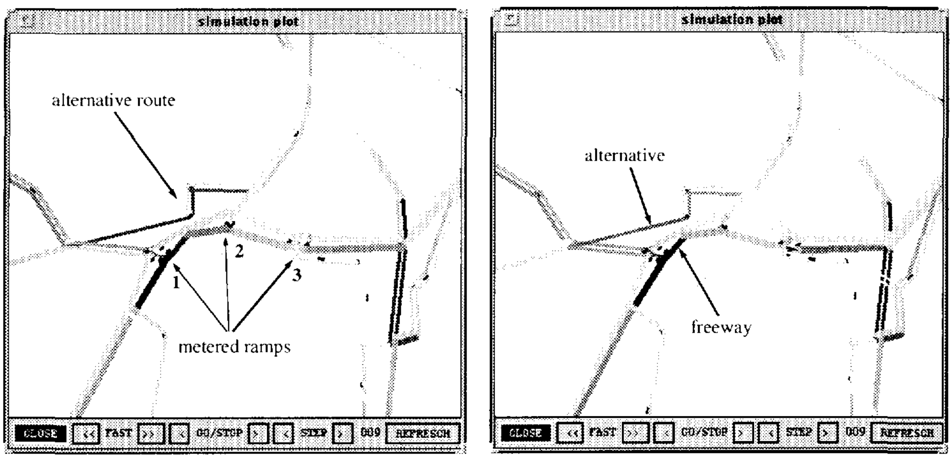


Figure 72: Rerouting behaviour. The density in the 13th period (09:00 - 09:15), the left figure displays the ramp metering scenario, the right figure displays the normal rush hour scenario.

The locations in figure 72 were chosen to illustrate that, due to ramp metering, alternative routes parallel to the freeway could be chosen. The left figure shows a slightly darker grey (higher density) at the alternative route then the right figure. A slightly lower flow can be detected at the freeway.

Table 9: Freeway versus Alternative route at 9:00am

Road	Scenario	Density (veh/km)	Flow (veh/hr)	Speed (km/hr)
Freeway	Normal	127	5800	45
	DTM	110	6000	50
Alternative	Normal	71	2000	30
	DTM	87	1800	25

Table 9 shows the values for several parameters of a link in the freeway and a link in an alternative route in the period between 09:00 and 09:15h. The two links are marked with arrows on the right map of figure 72. The normal morning rush hour scenario is compared with the DTM rampmetering scenario in table 9. The values show that traffic avoids ramp 1 and that a higher density is found at the alternative route. The flow at the alternative route has not increased because 2250 vehicles per hour is the capacity of that road; density has increased but throughput, due to slight congestion, has decreased.

7.6.4 Accident Scenario

To test whether the effects of incidents can be investigated using 3DAS an accident was simulated on the I-66. The accident was introduced by decreasing the capacity for the link with the accident by 60%. The OD-matrix and the departure functions were not changed.
Two different route choice strategies were used. One strategy using the same routes chosen during a normal morning rush hour, the other route choice strategy used an equilibrium assignment.
The first strategy represents a situation where travellers are unaware of the accident, while the second scenario gives a situation in which each traveller is diverted optimally.

In the first scenario (no diversion) there is a traffic jam on the I-66 which grows further upstream then normal for the morning rush hour, the average speed of the congested links is very low. The graphs in figures 73 and 74 show the situation for the middle section of the I-66. The density for each link is displayed in grey. The darker the grey the higher the density and the lower the speed. The situation for the 5th and 10th period is shown.

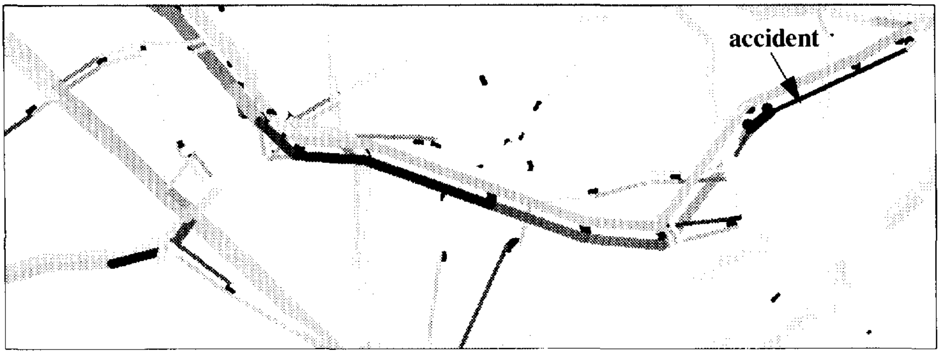


Figure 73: Accident on I-66; 5th period; 06:15, traffic jam starts

Figure 73 shows that at the link with the accident a queue starts to build.

Figure 74: Accident on I-66; 10th period; 07:30, traffic jam along entire section

Figure 74 shows the same section five periods (1¼ hours) later. In this figure the congestion has grown upstream and almost the entire section is jammed.

In this scenario the drivers did not divert to a different route because they were not aware of the accident. In the next scenario an equilibrium assignment was used, which means that all travellers were informed about the accident and chose their route accordingly.

The equilibrium assignment gave some remarkable results. The total length of the traffic jam that started due to the incident did not grow further upstream than in the *normal* morning rush hour. Comparison with the normal rush hour shows that the length of the queue was in fact shorter but the average speed was much lower. Arterials around the location of the accident were all more heavier loaded. Figures 75 and 76 show the I-66 again at the 5th and the 10th period for this scenario.

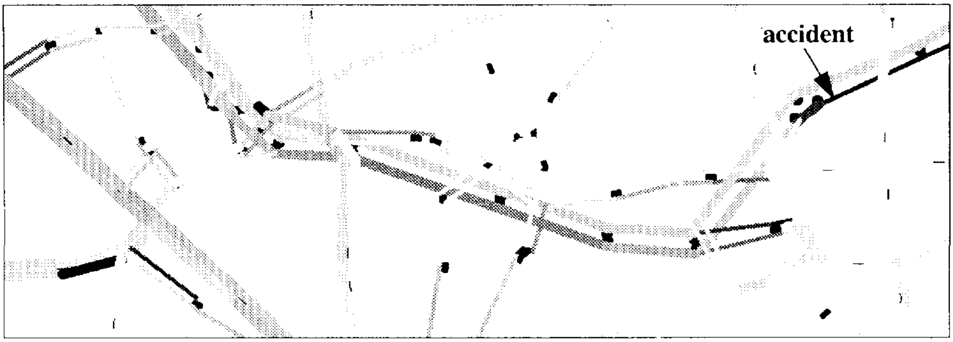


Figure 75: Accident on I-66; 5th period; 06:15, diversion scenario

Figure 75 shows that, again, in the 5th period the queue starts at the location of the accident. Further upstream density is lower because travellers start to divert. A higher density can be seen on the parallel route.

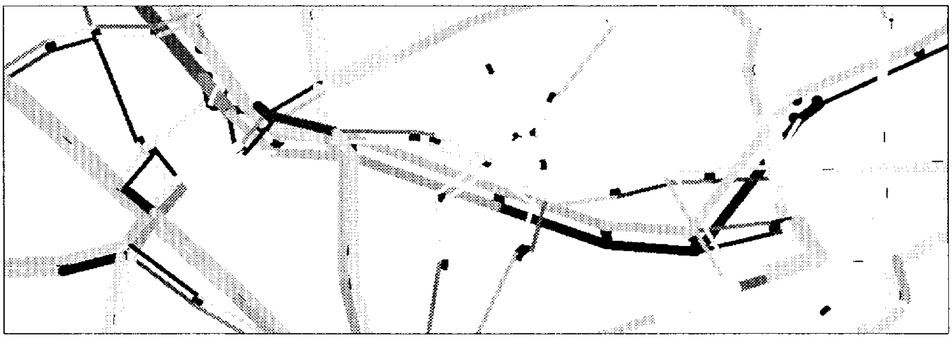


Figure 76: Accident on I-66; 10th period; 07:30, diversion scenario.

Figure 76 shows the same section five periods (1¼ hour) later. Here the congestion has grown upstream but not as far as in the other scenario. All the parallel routes show higher densities.

When the travel times to traverse the entire I-66 are compared, there is a significant difference between the two accident scenarios. The *normal* rush hour travel time is compared with the accident scenario and with the accident with diversion scenario in figure 77.

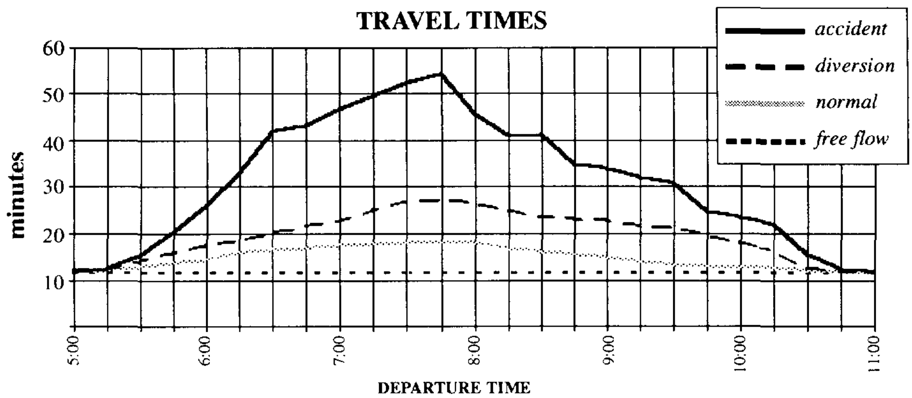


Figure 77: Travel Times to traverse the I-66 for four scenarios: the free flow situation, the normal rush hour, the accident scenario, and the accident with diversion scenario.

The free flow travel time on the I-66 is 11.5 minutes. For the normal morning rush hour it takes approximately 18 minutes to traverse the I-66 for traffic that departs at 07:30h. In the case of the accident, when the traffic is rerouted, travel time increases significantly, although the total length of the traffic jams is the same. When the traffic is not rerouted, travel time increases to almost an hour, to traverse the I-66, for traffic that departs at 07:45h.

Figure 77 shows that travel time during a normal rush hour is (obviously) the shortest. The scenario with the accident gives a travel time approximately 3 times as high. When diversion is allowed, travel times are approximately 1½ times as high. This case shows an improvement in travel time of approximately 50%. Of course this is an extreme case. The worst case is compared with the optimal, and there appears to be enough capacity on the arterial network.

7.7 Conclusions

This study showed that a dynamic assignment model can be very useful for planning applications. A number of clear advantages for using 3DAS instead of static assignment were demonstrated. The results give more detailed information about the occurrences of traffic jams, and the location or the cause of congestion can be identified more precisely. To alleviate congestion, DTM measures can be simulated, and all kinds of evaluations are possible, such as how travel times and traffic jam length might be influenced by various strategies.
Dynamic assignment has the advantage that all kind of *temporary* disturbances, such as accidents or road works, can be simulated, and the duration of delays can be derived. The study also showed that 3DAS can be used with larger networks

It must be stressed however, that data requirements are much more stringent, since by using 3DAS the level of detail is higher, the data must support this level of detail. The accuracy of time variance directly depends on the accuracy of the time variance of the OD-matrix. Due to the amount of data that 3DAS requires and produces, it is essential to find a good system of organizing and maintaining this large amount of data. In the beginning this may require a great effort, but with increased experience of using 3DAS this disadvantage will probably disappear.

The calculation time required for 3DAS is longer than that required for static assignment. For planning purposes however, the calculation time is not the main concern. Visualization of the results is much more important. Dynamic assignment gives flows *in time*. The best way to analyse the results is in a *movie like* manner. To do this a graphically powerful workstation must be used. This is one of the main reasons graphic workstations were used for this research. When the 3DAS model is used for ATMS a faster computation time is needed. This can be achieved by reducing the problem size (smaller network). When this is not possible a faster computer may be used. The computer design, described by H.J.M. van Grol in [29], about the development of special purpose hardware for assignment calculations gives a cost effective solution.

In this specific study the amount and the quality of the data was very poor, and there were limited possibilities to verify the data. Since we had limited insight into the local traffic patterns, we could not judge the quality of the OD-matrix. The time spent on this research was too short to use the results for any serious planning decisions. The study was therefore meant primarily to investigate the usefulness of dynamic assignment for planning purposes. For *real* planning decisions a more elaborate study would be required.

If Dynamic Traffic Management is to be successful there is a great need for more data and better (3D) OD-matrices. Hopefully new methods for OD-estimation and more data from induction loops and probe vehicles will give better results in the future.

Given more time for a study, more local knowledge of the study area, and more induction loop data, the 3DAS model has the potential to give reliable information for *real* planning strategies.

CHAPTER 8

APPLICATION OF 3DAS AS ATMS AMSTERDAM

8.1 Introduction

One of the requirements of dynamic traffic management is to provide a medium term prediction (one to three hours) of the traffic situation in a network. Based on this prediction a traffic operator can make decisions concerning rerouting measures, travel advice and ramp metering settings. In an on-line prediction system, the model must respond to accidents, weather changes and road works, and to react to these changing conditions, the model must be provided with data from various sources.
This chapter describes a research project in which the medium term prediction capabilities of the 3DAS model, with all the data preparing modules described in chapter 6, were linked together and tested.

The *primary* objective of the project was to find answers to the following two questions:
- Can the 3DAS model give a sufficiently accurate prediction of the future traffic conditions?
- Can the 3DAS model react to changing conditions, such as accidents and weather changes?

It is not easy to answer these questions. The reaction of the model to changing conditions depends heavily on available data, and the overall performance of the model depends on the traffic pattern of the specific day. The model might do very badly one day, and really well the another. This means that the entire system must be operational for several days, and that considerable effort must be put into making the required historical databases. Unfortunately the time for such an approach was not available.
The *secondary* objective of this study, was to test whether the model can give a prediction within a reasonable time. New predictions have to be made regularly for ATMS, and for every prediction on-line data is required to react to changing conditions. Predictions are clearly useless if they take more time than the predicted time span.

The Amsterdam beltway was selected as the pilot area. This network is well equipped with induction loops, has a manageable size, and several dynamic traffic management instruments in operation or planned for operation in the near future.
The research approach and how the data was obtained is discussed in this chapter. Several different scenario's are proposed in section 8.4, the performance of the hardware and the software is considered briefly in section 8.5, and the results are reported in section 8.6. The

chapter ends with conclusions.

8.2 Research Approach

To answer the questions given above, data was collected for three weeks at 141 locations in the Amsterdam network. The data was collected using induction loop detectors. The data was screened for errors and two days, April 14th and April 28th, were selected to be reproduced by the 3DAS model. A historical database was setup using data for several other days. The following research approach was used:

Can the 3DAS model give an accurate prediction of the traffic situation?

Dynamic assignment can be used for advanced traffic management if, given a network and several sources of actual data, it can predict a correct distribution of the traffic flow. The reproduction should not represent an average traffic flow, but the actual traffic flow at that moment. This includes all kinds of disturbances. To validate the 3DAS model, day specific demand and actual measured traffic flows are required.

Can the 3DAS model react to changing conditions, such as accidents and weather changes?

The second question partly overlaps the first question. To answer this question, a specific situation is required in which a particular condition disturbs the traffic pattern in the network. The model should be able to reproduce the effect of the disturbance. Reproduction should be made possible by changing the attributes of the network (e.g. available lanes) and adapting the speed-density functions.

During this research, various computers, including the LPA (see section 6.4.3), were used to make the calculations. Most modules of the system can work independently. The OD-estimator and the dynamic assignment model are the most computational intensive modules. An impression of the performance of the 3DAS model is also given in this chapter. The OD-estimation module is a separate research project, and the results of this research will be reported elsewhere.

8.3 Data

The data available for this research consisted of a network, represented by nodes and links, with several static attributes and induction loop data from 141 screenlines.

8.3.1 Network

The Amsterdam freeway network is a beltway with a diameter of approximately nine kilometres. The beltway has a total length of approximately 30 kilometres. There are two tunnels in the network and four major freeway intersections. In total there are 76 on and off ramps. The lay-out of the network is given in Appendix A. The network is constructed with a total of 286 nodes and 430 links. A small part of the arterial network is present near on- and off-ramps. This makes it possible to combine the origins and destination in the network as one node. In total there are 21 origins and 21 destinations. The network was extracted from a large national network. Considerable customizing was required to get a truthful reproduction of the reality. Detailed maps of the beltway were used during the customizing. An example of such a map, for the intersection at Watergraafsmeer, is given in appendix F. In total 17 maps were used.

8.3.2 Induction Loop Data

The Amsterdam network is equipped with hundreds of induction loops. One minute data was measured at all the screenlines for a duration of three weeks in April 1994. Each screenline has double induction loops on each lane. Every minute, flow and speed, is available at the traffic control centre. A schematic overview is given in appendix A with the location of all the screenlines. The exact locations of the induction loops are marked on the detailed maps of the beltway (see appendix F). The data sets were screened for errors and inconsistencies, which resulted in the deletion of several locations. In total 141 screenlines were used. The interval length of the induction loop data is one minute. This data was aggregated to a five minute interval, because an accuracy of one minute is not really feasible for network wide predictions.

Taking the completeness of the data into account it was decided to use five days for the historical database. OD-estimation, departure time patterns and speed-density functions were estimated based on this database. The procedures used are described briefly in the following three sections.

8.3.2.1 OD-matrix

The estimation of the OD-matrix was done externally. The only information used for the estimation was the induction loop data of the historical database. The method estimates the matrix by minimizing the difference between the observed time dependent link flows and the calculated time dependent link flows. The method is described in section 6.3.2 and is still experimental. The matrix estimation was carried out using a period length of 30 minutes. This was done to minimize the influence of travel times. This resulted in a historical matrix in which split ratios with a thirty minute interval were used.

The historical split ratios were multiplied by the historical on-ramp volumes (departure time patterns, see 8.3.2.2) to obtain a dynamic OD-matrix. As the on-ramp volumes were aggregated to five minute data, the resulting dynamic OD-matrix is also on a five minute basis, but with a fixed split ratio of 30 minutes. This OD-matrix is denoted as the *Future OD-matrix* in section 6.3.2. According to the general structure of the proposed ATMS in figure 56 on page 101, a second OD-matrix, the *Start OD-matrix* is required to reconstruct the current flows in the network. This OD-matrix was omitted from this case study. As a consequence it was not possible to start the calculations in the middle of the rush hour, but only at the beginning of the rush hour. An extra start hour has been added to the calculations to reconstruct the current flows.

8.3.2.2 Departure Time Functions

A departure time pattern was measured and stored in the database for each on-ramp in the network. For some on-ramps, no induction loop data was available, and the pattern was calculated by subtracting the volume measured upstream of the ramp from the volume measured downstream of the ramp. The resulting historical patterns were fitted with the measured patterns as described in section 6.3.2.3. A departure pattern could not be derived for three on-ramps in the network, so a departure pattern was assumed, based on the matrix estimation.

8.3.2.3 Speed-Density Functions

The speed-density function used by the model is the Smulders function, described in equation 74 in section 5.3. For each of the 141 measured locations the parameters of the speed-density functions were estimated, using a least squares minimization. Several links with similar characteristics were grouped together and shared the same speed-density relationship.
In total 19 speed-density functions were distinguished, these are given in table 10.

Table 10: Speed-density functions for Amsterdam

type	description	ρ^{max}	v^{max}
1	3 lane freeway	330	105
2	3 lane freeway	330	105
3	3 lane freeway	330	105
4	1 lane freeway intersections	138	105
5	2 lane freeway intersections	275	96
6	2 lane freeway	225	105
7	various on and off ramps	125	70
8		200	70
9		220	70
10	arterial network	var	50
11	4 lane freeway (3 lanes + merging lane	440	105
12	buffer links (near origins)	440	var
13	2 lane on and off ramps	220	70
14	merging section entering Coen tunnel	200	100
15	merging section Nieuwe Meer (east)	440	100
16	freeway intersections	var	90
17	merging section Nieuwe Meer (north)	440	100
18	1 lane section at Coen plein	137	100
19	2 lane Coen tunnel	225	100

The names of the intersections are given in appendix A. The parameters of the speed-density functions were estimated with the data of the historical databases for each link type. To reproduce a specific day, the parameters were adjusted using the actual data from that day.

8.4 Modelling & Calibration

It was decided to consider three different scenario's to answer the questions stated in the research approach:

I. The first scenario tried to reproduce the morning rush hour of Thursday, April 28th. The data on this day is error free and no major accidents, or other disturbances occurred. It was decided to reproduce the rush hour from 06:00h in the morning until 09:00h in the morning. A longer time was not feasible, because most of the traffic has left the network after two hours. The period length used in the calculation is five minutes. This period length was chosen, because a total of 36 periods could be managed by the software, and resulted in a feasible representation of the dynamics in the network. A shorter period length would suggest an accuracy that cannot be reached, and for a longer period length, the interaction between time periods diminishes. An impression of the capabilities of 3DAS can be achieved by comparing the results of the 3DAS model with the actual measured flows

II. The second scenario tried to reproduce the morning rush hour of Thursday, April 14th. On this day an accident occurred on one of the freeways which blocked the freeway for 20 minutes. The same time span and period length as in the first scenario was used. The results are again compared with the actual measured flows for this scenario.

III. To get an impression of the value of the 3DAS model for making predictions, a third scenario was considered that does not use the 3DAS model to make a prediction, but used the collected induction loop data to make a prediction. The measurements were summed and averaged for several measured days, with similar conditions. This scenario represents the historical average traffic pattern.

The first question of the research approach was answered by the first scenario, and the second question was answered by the second scenario. The first scenario is described in section 8.6.1, the second in 8.6.2, and the results of a comparison between the first two scenario's and the historical average is given in section 8.6.3.

8.5 Hardware & Software

The software described in section 6.4.2 was used for the computation. The use of extensive graphics again showed its value. Errors, present in the network, such as the wrong number of lanes, or an induction loop at the wrong location were easily located and fixed. The actual data was read by the model and displayed as a line in all the graphs used to represent the calculations. This made it possible to compare the actual data with the calculated data directly. Examples of the graphs are given in figures 78 through 82 in the next sections.

During calibration, the computations were done on a PC, Intel 486/66 using the Linux operating system, and on a SUN Sparc10. One iteration on the PC, using the trajectory assignment, took 11 seconds. With the surface assignment, one iteration took 46 seconds. On the SUN workstation these times were respectively 9 and 40 seconds. After calibration, the LPA described in section 6.4.3 with 14 processors was used to measure calculation time. One iteration on the LPA using the trajectory assignment took less than a 1 second, and using the surface assignment ±4 seconds. On average sixty iterations were required for the model to converge. To obtain a smooth result, the last 10 iterations were done using the surface method. So, for real-time applications, networks of this size are reasonably manageable on PC's or workstations, while for larger networks the LPA can give a substantial improvement in calculation time.

8.6 Results

The three proposed scenario's are described in the next three sections.

8.6.1 Morning Rush Hour at Thursday, April 28th.

A dynamic assignment was carried out based on the dynamic OD-matrix, and the speed-density functions. Calculation started at 06:00h. At this time the network was almost empty. During calibration numerous modifications were made to the network. The wrong number of lanes at one location can disturb traffic flow over a large part of the network. Several on-ramps were not well represented, and at some locations individual speed-density relationships were required. Detailed drawings were required to figure out the exact lay-out of the network. Finally a fairly good reproduction of the actual traffic flow was achieved. Four different locations in the network are showed to give an impression of the results. Figure 78 gives flow and speed at two locations near the Coen tunnel, and figure 79 gives flow and speed for two locations near the Zeeburger tunnel. .

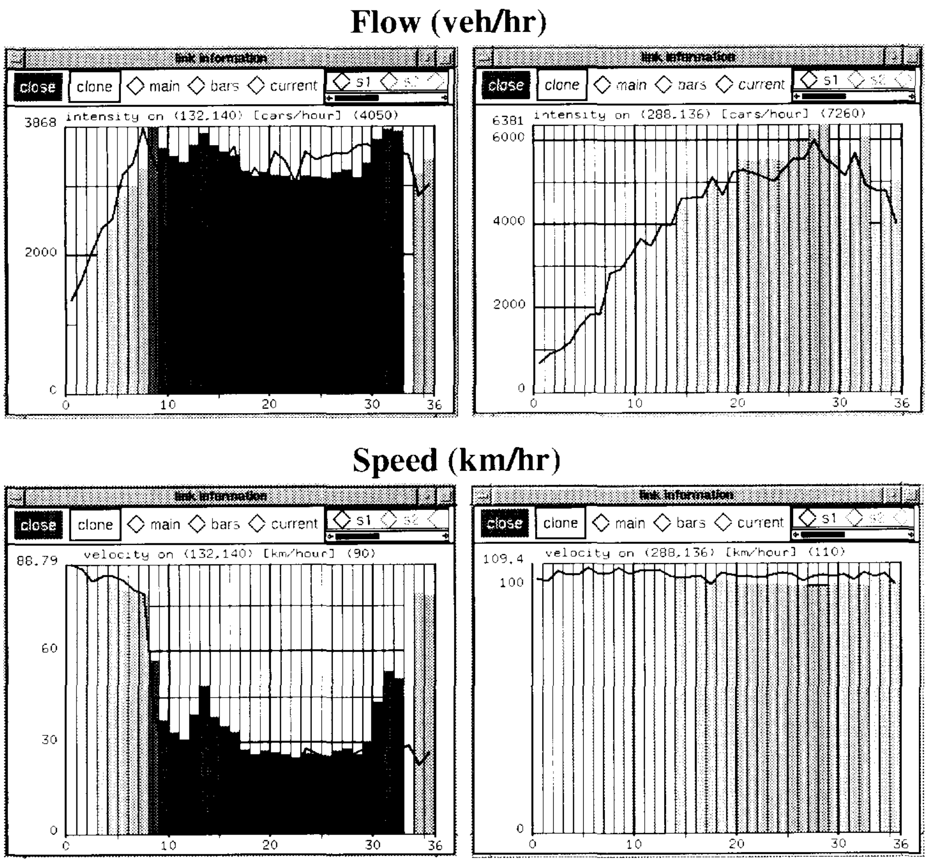


Figure 78: 3DAS results (bars), compared with the actual measured traffic flow (solid line) at two locations near the Coen tunnel.

The left side of figure 78 shows a location denoted by the number 108 in appendix A. The traffic is queuing before entering the Coen tunnel. The flow pattern is reproduced fairly well, except in the last 6 periods, speed increases, while in reality speed remains low. The error in the results could be caused by an inaccurate OD-matrix, or an error of the model. The right side of figure 78 shows a location in the other direction of location 108. This traffic is leaving the Coen tunnel and travelling north. This location shows a free flow situation and is a good reproduction.

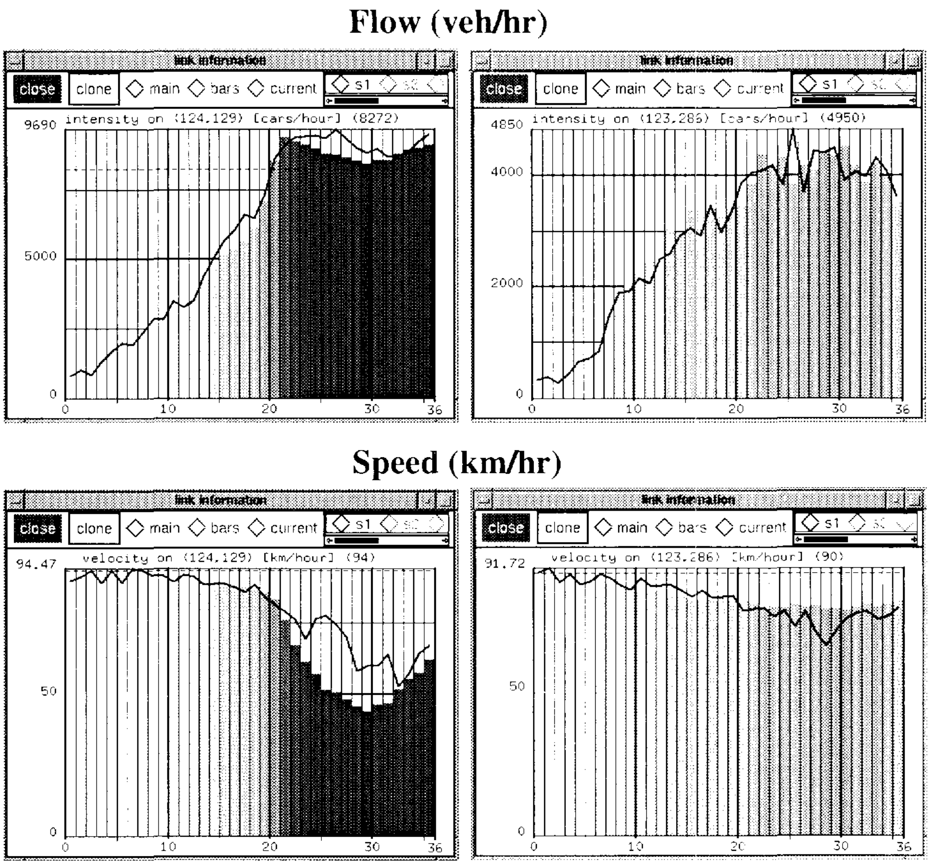


Figure 79: The 3DAS results (bars), compared with the actual measured traffic flow (solid line) at two locations near the Zeeburger tunnel.

The left side of figure 79 shows a location denoted by the number 157 in appendix A. On this section two freeways merge. The reproduction of speed by 3DAS is a little pessimistic. The right side of figure 79 shows a location in the other direction of location 157 and represents the traffic coming from the east and going to the north. At this link the traffic changes freeway, and this link is reproduced fairly accurately.

8.6.2 Morning Rush Hour at Thursday, April 14th

A dynamic assignment was made for the morning rush-hour of April 14th based on the dynamic OD-matrix of April 28th. Calculation started at 06:00h. At 06:50h an accident occurred north of the Nieuwe Meer intersection. This accident is clearly present in the induction loop data. Downstream flow is almost zero and a large traffic jam has built up in several directions. At 07:35h, a larger downstream flow was again measured and the traffic jam builds down. The accident was introduced into the model by setting the maximum density to 1 veh/km during the accident for the appropriate section.
Figure 80 shows the impact of the traffic jam during two different periods.

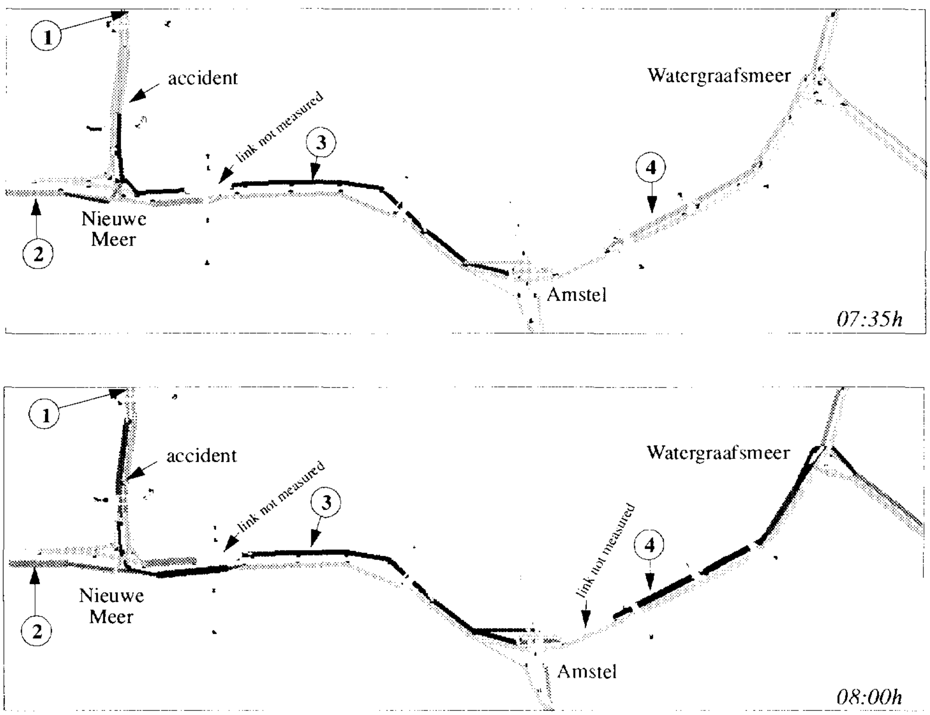


Figure 80: Traffic jam at the south part of the beltway on Thursday, April 14th at 07:35h, and 08.00h, the figure displays actual measured density at each section in a grey scale, the darker the colour the higher the density, some links are light grey, because they are not measured by induction loops.

Between 06.50h and 07:35h all the lanes were blocked at the location of the accident, after 07:35h all lanes became available again. The traffic jam started to build down from the start point, but it still builds up from the tail of the traffic jam. At 08:00h the tail of the traffic jam reached the Watergraafsmeer intersection.
Four different locations near the accident are shown in figures 81 and 82 to give an impression of the reproduction of the traffic jam. The exact locations are given in figure 80.

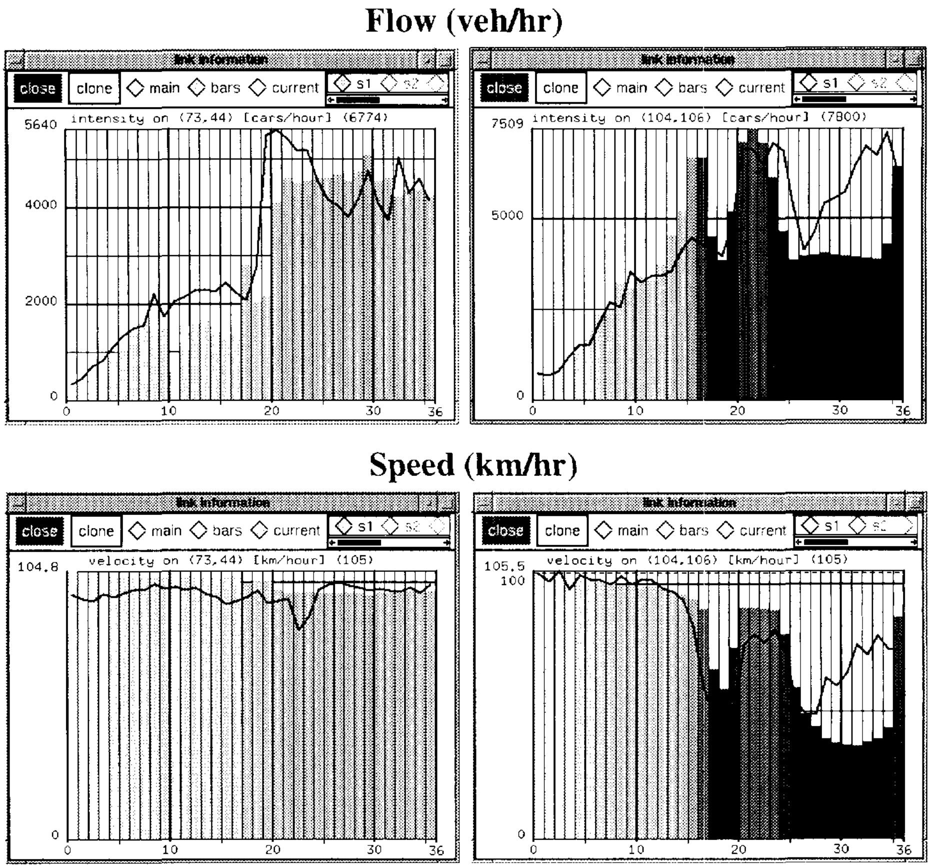


Figure 81: 3DAS results (bars), compared with actual measured traffic flow (solid line) at two locations near the accident; left figures is downstream of the accident, travelling north and denoted as location 1 in figure 80, the right figures is upstream of the accident travelling east, denoted location 2 in figure 80.

The left side of figure 81 shows a location denoted location 1 in figure 80. This location is downstream of the accident, and shows a restored flow and speed. Flow is low for the first twenty periods. Around the twentieth period (07:35h), the freeway is clear again and flow increases. The right side of figure 81 shows a location denoted as location 2 in figure 80. This location is upstream of the accident. The flow pattern is quite complicated here. The first decrease in speed around the 18th period is the result of the accident. The second decrease in speed, around the 27th period is the result of a traffic jam that appears east of Nieuwe Meer, for trips going east and has nothing to do with the accident. This traffic jam was not reproduced exactly by 3DAS, and for this reason speed is somewhat underestimated around the 30th period.

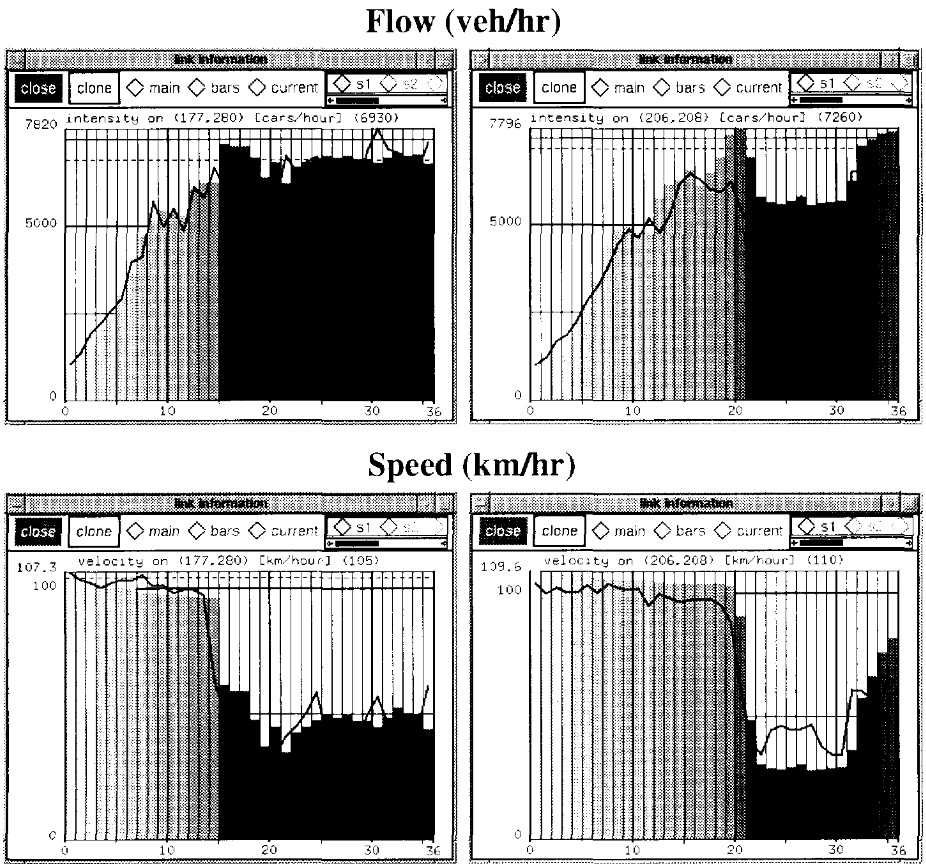


Figure 82: 3DAS results (bars), compared with the actual measured traffic flow (solid line) at two locations further downstream of the accident, travelling west, the locations are denoted 3 and 4 in figure 80.

The left side of figure 82 shows a location upstream of the accident, denoted as location 3 in figure 80. The graphs show that the traffic jam reached this link in the 25th period. Speed drops rapidly and remained low until the last period. The fluctuations in speed and flow were not reproduced by 3DAS, which gives a somewhat smoother pattern. The right side of figure 82 shows a location far downstream of the accident, denoted as location 4 in figure 80. The period in which the traffic jam reached this link is fairly well reproduced.

8.6.3 A Prediction Based on the Historical Average

The two scenario's described were validated with data measured for three weeks by induction loop detectors at 141 locations. Five days were selected from this data collection which contained no major disturbances. The traffic patterns for these days were summed and averaged. The resulting data set represents the average traffic pattern for the morning rush hour from 06:00h until 09:00h for a "normal" weekday.

This results in five different data sets for the morning rush hour, if the data of the two scenario's described earlier are included:

- **I.** The first data set represents the actual measured situation from 06:00h until 09:00h with 36 period of 5 minutes at Thursday, April 28th.
- **II.** The second data set represents the results of the 3DAS model of Thursday, April 28th.
- **III.** The third data set represents the actual measured situation from 06:00h until 09:00h with 36 period of 5 minutes at Thursday, April 14th.
- **IV.** The fourth data set represents the results of the 3DAS model of Thursday, April 14th.
- **V.** The fifth data set represents the average traffic pattern for a morning rush hour from 06:00h until 09:00h in 36 period of 5 minutes.

A comparison can be made between the results of the 3DAS model and the historical average for a normal day and for a day with an accident. The difference in travel time between I and II, and I and V for a normal day, and between III and IV, and III and V for a day with an accident.

The difference in travel time between two data sets was calculated for each location/link separately. Travel times were summed over all the periods for both data sets. The absolute difference between the travel times was chosen as the measure of difference. The difference for a specific link between scenario I and scenario II can be calculated according to equation 86.

$$\sum_{p} \left| t_{I}^{p} - t_{II}^{p} \right| \qquad \text{(eq 86)}$$

Where p is the period number and t_i^p the travel time in period p for scenario i. Figure 83 shows these absolute differences for each location and each period. The absolute difference between data set I and II is represented by a bar, and in the same figure, the difference between data set I and the historical average V is shown as a solid line.

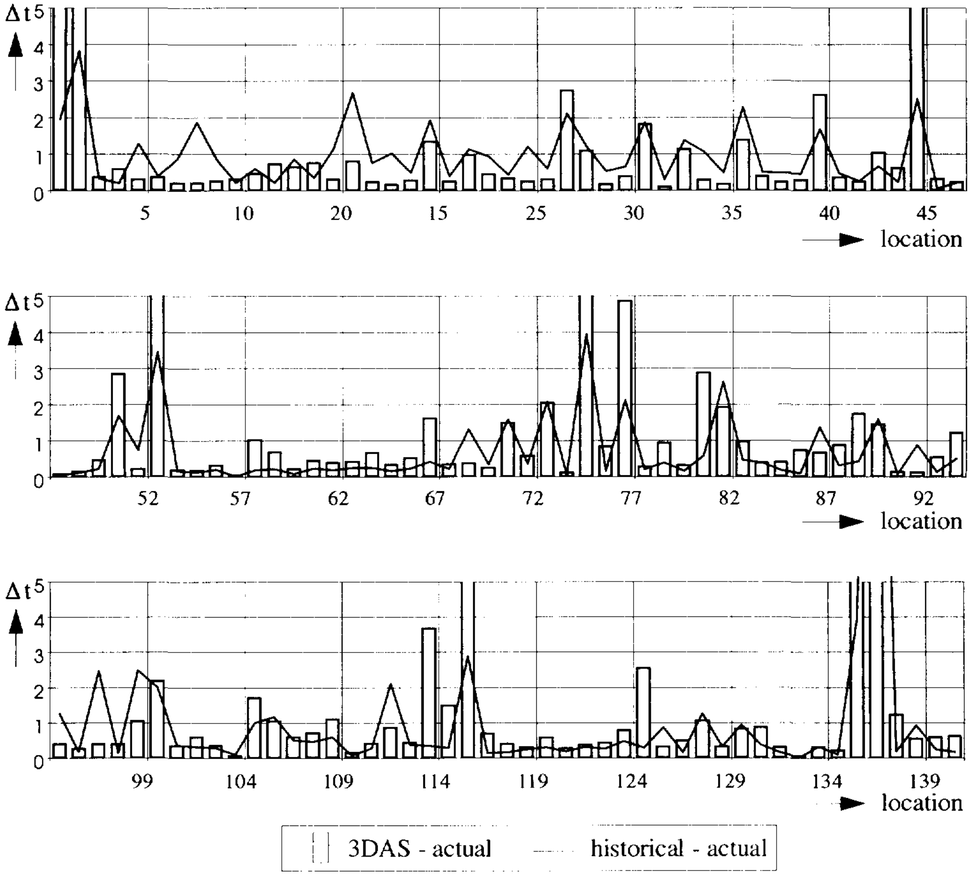


Figure 83: The absolute difference in minutes between 3DAS (II) and the actual data of April 28th (I) in bars, and the difference between the historical average (V) and the actual data (I) as a solid line.

Figure 83 shows that the 3DAS model is capable of reproducing the actual situation of April 28th, because for most links the difference in travel time summed over the entire time span is only a few minutes. The same figure, however, shows that the historical average is at least as good. To make a prediction for the morning rush hour of April 28th, it might be better advised to take the historical average of several similar days. So, for this specific case, it could be argued that a better prediction can be made by collecting historical data.

Figure 84 shows the same data as figure 83, but now for April 14th. The absolute difference between data set III and IV is represented by a bar for each location, and in the same figure, the difference between the data set III and the historical average V is shown as a solid line.

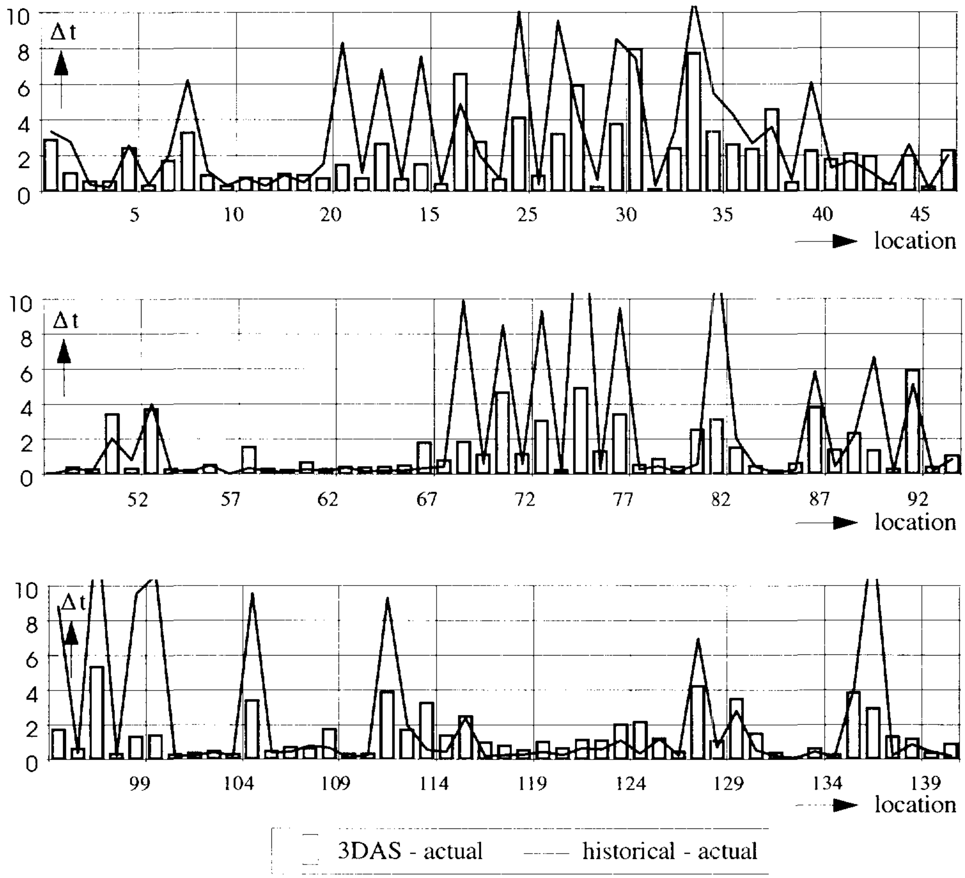


Figure 84: The absolute difference in minutes between 3DAS (IV) and the actual data of April 14th (III) in bars, and between the historical average (V) and the actual data (I) by a solid line.

Figure 84 shows that the 3DAS model is also capable of reproducing the actual situation on April 14th, it also shows that for several locations, the historical average gives a bad estimate.

8.7 Conclusions

This chapter illustrates that a dynamic assignment model can be used for ATMS predictions of the traffic flow in a network. OD-estimation and the 3DAS model are capable of reproducing a traffic pattern on a specific day, including accidents.

Correct representation of the network and speed-density functions is very important. Specifying the wrong number of lanes at a road section can have consequences for the entire network. These errors can be located easily and corrected using the available graphic software.

The model was tested on two different days. One day represented a "normal" morning rush

hour, the other day represented a morning rush hour with an accident which blocked all lanes for 45 minutes.
Comparing the traffic flow pattern of a specific day with the historical average pattern, it can be concluded that for "normal" days, the 3DAS model can reproduce the situation, but that the historical pattern is at least as good. It can be concluded from this observation that a dynamic assignment model is not required to make short term predictions for normal days, for days with abnormal conditions however, use of a dynamic assignment model is justifiable. The day with the accident was better reproduced by 3DAS than by the historical data.

Networks of this size are reasonably manageable on PC's or workstations for real-time applications. Insight is provided into the development of traffic for the next one to two hours within several minutes. The LPA can give a substantial improvement in calculation time for larger networks.

The data used for this case proved to be very useful, using this data set it was possible to improve and calibrate the 3DAS model further. The queuing algorithm, for example, needs further improvement, and some ideas for improvement are proposed in section 9.3. The OD-estimation was at an early stage of development during this research, and further improvements in OD-estimation may improve the results further.

CHAPTER 9

CONCLUSIONS

9.1 Introduction

The *main research issues* tackled in this dissertation were:

- The development of a traffic model to simulate traffic on a network and to test DTM instruments for their ability to solve the congestion problem.
- The development of a traffic management system which can predict traffic flow on a freeway network and utilize DTM instruments

This chapter summarizes and discusses the conclusions made in this dissertation. Some points for further research are proposed in section 9.3, some of these points have already been studied, but have never been tested using real data.

9.2 Conclusions

One solution to alleviate congestion in transportation systems is Dynamic Traffic Management (DTM). Dynamic Traffic Management employs technologies for real-time traffic monitoring, network wide management of traffic flows, and traffic-adaptive control to respond to changing traffic conditions while improving the efficiency, safety and travel conditions of a highway network.

At the present time there are very few models suitable for DTM. A new model, called 3-Dimensional Assignment (3DAS), is described and tested in this dissertation. Like most other dynamic assignment models, the simulation period is divided into intervals of equal duration, referred to as periods. The link parameters may be defined separately for each period. Network and travel demand are presumed given. An iterative process is used to solve the model. In every individual iteration, the optimal route for every OD-relation in each departure period is calculated. The OD-matrix is assigned to the network according to these routes. Before each iteration, new link travel times are calculated for every period, based on the conditions resulting from the previous iteration. The iteration process stops when the stop criterion is reached, i.e., when the loading of the network remains essentially unchanged from one iteration to the next. The model uses density as an explanatory parameter instead of flow, because the relationship between flow and travel time is dual, while the relation between density and travel time is not. This makes it possible to model a decreasing flow when a critical density has been reached, thus representing congestion.

The 3DAS model was tested and compared to the microscopic simulation model FOSIM, which describes traffic realistically. A bottleneck and an on ramp were tested. Four different methods for pathfinding and two different methods for assignment, and several parameters were calibrated. The 3DAS model was shown to be capable of reproducing these basic traffic situations. Although the results were not reproduced in detail, the effect a bottleneck and an on ramp would have *in a network* were reproduced.

The 3DAS model was developed and implemented to operate in planning and management systems. Little extra effort is required to use the 3DAS model for planning applications. Traditional *static* OD-matrices can be used, and made dynamic using departure time functions. For management applications, however, several extra modules are required to provide the 3DAS model with input and to make the model adaptive to prevailing conditions. One of the modules, the OD-split matrix module, is a research project by itself and is described elsewhere. Two other modules, the departure time functions module and the speed-density functions module were studied more elaborately. The departure time functions module uses historical data to predict future departure time functions. Empirical research showed that the departure time functions are fairly consistent each day. This gives the opportunity to use the historical function as a basis for the prediction. Good results were achieved with the on ramps in Amsterdam by fitting the historical function to the current measured function. The speed-density functions module makes the model adaptive to prevailing conditions. The functions are easily adapted using data from induction loops. It is not necessary to be aware of all kinds of weather circumstances or incidents.
A fast computing time is essential to make predictions. A prediction is clearly useless when the computing time exceeds the required prediction time. Parallel research towards special purpose hardware solves has solved computational problems [29].

Using 3DAS for planning applications was tested in Washington DC, USA. The planning case showed a number of clear advantages for using 3DAS instead of traditional static assignment. The results give more detailed information about the occurrences of traffic jams, and the location or the cause of congestion can be identified more precisely. DTM measures can be simulated to alleviate congestion, and all kinds of evaluations are possible. Influences on travel time and traffic jam length, effects of ramp metering and rerouting can be investigated. Dynamic assignment also has the advantage that it is possible to simulate all kind of *temporary* disturbances, such as accidents or road works, and that the duration of delays can be derived. The study also showed that 3DAS can be successfully used for larger networks.
Data requirements, however, are much more stringent, since by using 3DAS the level of detail is higher, the data must also support this level of detail. The accuracy of time variance depends directly on the accuracy of the time variance of the OD-matrix. No discretization problems were encountered because demand was represented with a rather smooth pattern.
It is essential to find a good system of organizing and maintaining the large amount of data 3DAS requires and produces. In the beginning this costs considerable effort, but with increasing experience this disadvantage will most probably disappear. Visualization of the results is very important. With dynamic assignment the results are flows *in time*. The best way to analyse the results is as a *movie*. To do this a graphically powerful workstation must be used, this is one of the main reasons graphic workstations were used for this research.

Using 3DAS for management applications was tested in Amsterdam in the Netherlands. Correct representation of the network and speed-density functions is very important. The wrong number of lanes at one section can have network wide consequences. Using the available graphic software these errors are easily located and corrected.
The model was tested on two different days. One day represented a "normal" morning rush hour, the other day represented a morning rush hour with abnormal conditions. In this case an accident occurred which blocked all lanes. The case showed that the 3DAS model is capable of reproducing both days.
Comparing the traffic flow pattern of the normal day with the historical average pattern, it can be concluded that the historical pattern is at least as good as the 3DAS prediction. From this observation it can be concluded that a dynamic assignment model is not required to make predictions. For days with awkward conditions, however, such as accidents and road works, the results of 3DAS were significantly better than the historical average.
The size of the Amsterdam network caused no problem for the 3DAS software and conventional hardware was able to make a prediction within several minutes. Special purpose hardware can give a substantial improvement in calculation time for larger networks.

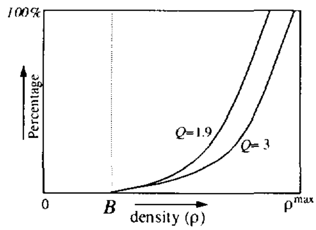
9.3 Further Research

One of the main problems during this research was the lack of suitable data to test the model. Suitable induction loop data became available only during the end of the research period. Further, with increasing research towards the estimation of dynamic OD-matrices, more validation and calibration is now possible. Several aspects of the model need some more fine tuning. Further research is required for three of these aspects, the queuing algorithm, the merging algorithm and the speed-density functions. Some steps in this direction have already been made during the Amsterdam case. Further extensions of the model towards dynamic departure times are desirable for planning purposes. The mathematical properties of the model and the possibility to optimize different DTM strategies is another aspect that requires further research. This section ends with some advice concerning the software.

The queuing algorithm

The queuing function, as described in section 4.7 was improved during the Amsterdam case study. The improved function is given in equation 87.

$$F(\rho) = A \cdot \left(\frac{\rho}{\rho^{max}} - B\right)^{Q} \qquad B \le \rho \le \rho^{max} \qquad \text{(eq 87)}$$



This function gives better queuing behaviour in the model. The function prevents reaching maximum density on a link, which improves the bottleneck behaviour of the 3DAS model. The function was tested and calibrated using the bottleneck network (see section 5.4.1) and by the Amsterdam case study. Recommended values for the parameters are: A=2.5, B=0.24 and Q=1.9. The parameter B is equal to the critical density, because at free-flow conditions no queuing is observed. The bottleneck network shows only a very small queue and the differences with the old function are small. The behaviour of this queuing function with different values for the parameters is similar to the old function. Further research is recommended, however, to gain more insight into the effect and sensitivity of the parameters. Explicit bottleneck situations are required to calibrate the function.

Several different theories exist about how a traffic jam starts due to high densities. Another theory worth testing is described by van Toorenburg [91], who describes the process in a bottleneck and claims that the flow in a bottleneck does not decrease below the capacity of the bottleneck (or only by a few percent). Upstream of the bottleneck, however, flow can decrease below the capacity as an effect of queuing. This theory has been implemented in the 3DAS model, by allowing density not to exceed the critical density (or by only a few percent) so that the flow cannot become lower than capacity. In the case of queuing, however, a higher density than critical density is allowed. Finding an appropriate data set to test this theory is difficult because very often several on or off ramps will disturb the pattern.

The merging algorithm

The FOSIM microscopic simulation model was especially calibrated to model the merging of traffic at weaving sections [82]. The results of this research gave several possibilities to study and to implement more advanced merging algorithms in the 3DAS model, however, more elaborate studies are required, especially to study the influence of ramp metering on the merging behaviour.

Speed-density functions

A database with speed-density function for various road types is desirable for ATPS applications. An attempt to setup such a database was made in section 5.3.4. More data and research is required to complete this database.
The speed-density function needs to be adapted to prevailing conditions for management applications. A disadvantage of the Smulders function is, however, the discontinuity and the many parameters. Therefore in the case of ATMS, with continuous updating, a simple speed-density function is desirable, which is easily adapted by on-line data to different situations.
Some research has been done to provide a function which is computationally cheap. Since every link in the network can have its own function, the adaptation of these functions to different conditions will take some computational effort.
The function is given in equation 88, and referred to as the 3DAS function:

$$v = \frac{\beta \cdot v^{max}}{\rho^{\alpha} + \beta} \qquad \text{(eq 88)}$$

Where v is the speed on a link, v^{max} is the free flow speed, ρ is the density. α and β are two parameters. The 3DAS function and its related functions are plotted in figure 85.

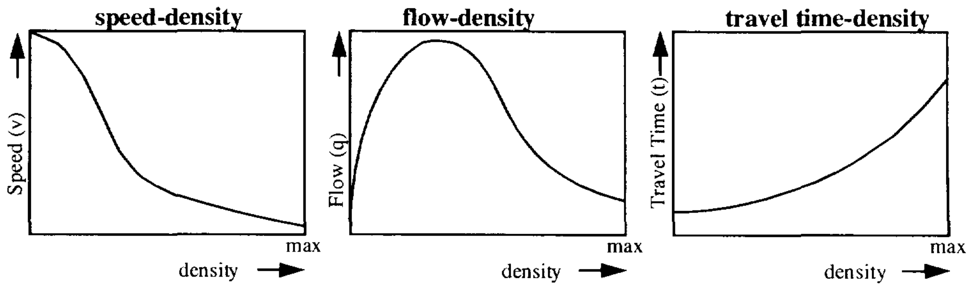


Figure 85: The 3DAS function and its related functions

The 3DAS function has three parameters: α, β and v_{max}.

The 3DAS function was calibrated using a least squares method for the ten chosen locations given in table 2 in section 5.3.2, and the four chosen days given in table 1 in section 5.3.2. The results are compared with the Smulders function. The mean square error is shown for both functions for each day at the ten locations in figure 86. The locations are given along the x-axis and the mean square error along the y-axis.

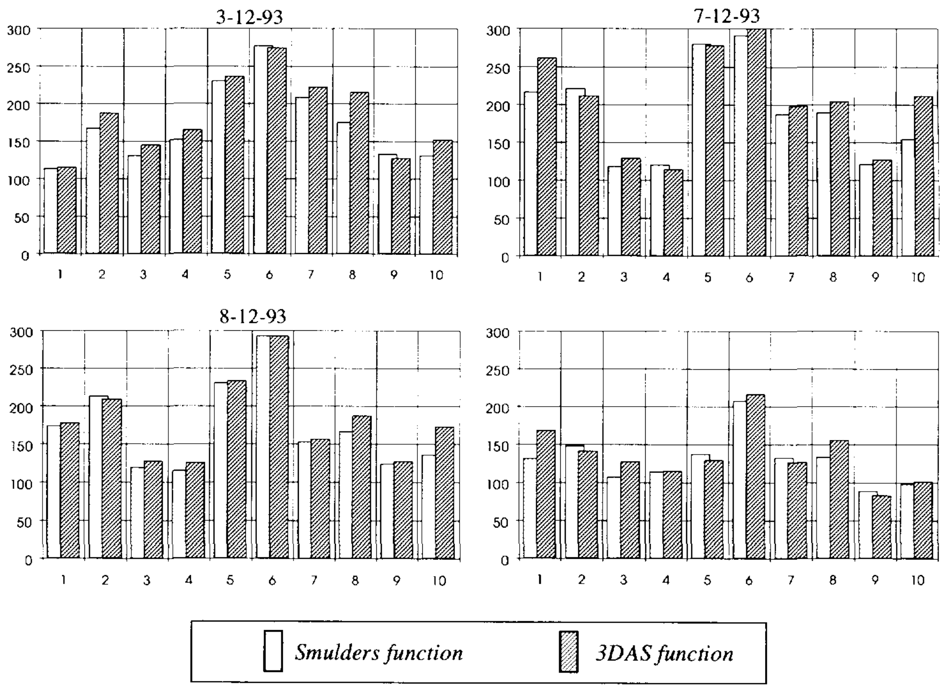


Figure 86: Mean Square Error between measured data and the Smulders speed-density function and the 3DAS speed-density function for ten locations at four different days.

The Smulders function results, in most cases, in a slightly better match than the 3DAS function. The 3DAS function is, however, easier to calibrate requiring less computational effort. The discontinuity of the Smulders function can give problems with the least squares method, because the minimal least square error is not uniquely defined.

Mathematical properties of the model

Several aspects of the model need further mathematical attention. The convergence properties of the model and the existence of a unique solution are still unsolved problems.
The model is not capable of generating optimal solutions for DTM scenarios automatically. There is, for example, a need for models which are able to determine optimal ramp metering rates, such that the total system travel time is minimized.

Improvement of the software

A lot of software was developed during the research. This software was made to test the theory of the model and has been found to be very helpful. Several unexpected effects were discovered and corrected using the software. The extensive use of visualization especially showed its value. Some of this software could be used for professional use and the model should be adopted and integrated into some kind of commercial traffic package.
Currently the 3DAS model runs on almost any UNIX based X-window system. The program is written in the C programming language.

REFERENCES

[1] Updates to COG/TPB Travel Forecasting Procedures, Highway Applications, 1990 Validation and 2010 Simulation, DRAFT, March 1993.

[2] Abkowitz M., "Understanding the Effect of Transit Service Reliability on Work Travel Behavior", *Transportation Research record, 794, pp 33-41, 1981.*

[3] Bakkenist Management Consultants, "Projectdefinition Monitoring System", *december 1989.*

[4] M.J. Beckman, C.B. McGuire & C.B. Winsten, "Studies in Economics of Transportation", *Yale University press, New Haven, 1956.*

[5] Ben-Akiva M., "Dynamic Network Equilibrium Research", *Transportation Research A, Vol 19A, pp 429-431, 1985.*

[6] Ben-Akiva M., A. de Palma & p. Kanaroglu, "Dynamic Model of Peak Period Traffic Congestion with Elastic Arrival Rates", *Transportation Science, 20, pp 164-181, 1986.*

[7] E. van Berkum & P. van der Mede, "The Impact of Traffic Information", *Ph.D. thesis, Delft University of Technology, ISBN 90-9006318-8,* 1993.

[8] D. Branston, "Link Capacity Functions: A Review", *Transportation Research, volume 10, pp 223-236, 1976.*

[9] J. Boesefeldt & M. Neuherz, "RTI/IVHS on European Highways", ITE 1993, The Hague, The Netherlands.

[10] Box, G.E. & G.M. Jenkins, "Time Series Analysis: Forecasting and Control", *San Francisco: Holden-Day, 1976.*

[11] D.E. Boyce, Bin Ran, L.J. LeBlanc, "Dynamic User-Optimal Traffic Assignment Model: A New Model and Solution Technique", submitted to *Transportation Science,* 1991.

[12] D.E. Boyce, Bin Ran, L.J. LeBlanc, "A Combined Dynamic Mode / Departure Time / Route Choice Model with Multiple Groups of Travelers", *Presented at the 72nd Annual Meeting of the Transportation Research Board, Washington DC., USA January 1993.*

[13] P.H.L. Bovy, "Toedelingen in congestievrije netwerken", *Ph.D. thesis (in Dutch), Delft University of Technology, 1991.*

[14] Bureau Goudappel Coffeng B.V., "Orientatie toeritdosering Californie" (in Dutch), *Rotterdam, june 1989, Rijkswaterstaat, Dienst Verkeerskunde.*

[15] Bureau Goudappel Coffeng B.V., "RIA nader bekeken, Eindrapportage deel I" (in Dutch), *Rotterdam, june 1989, Rijkswaterstaat, Dienst Verkeerskunde, Directie Noord-Holand.*

[16] Bureau of Public Roads, "Traffic Assignment Manual", U.S. Department of Commerce, Urban Planning Division, Washington D.C.

REFERENCES

[17] M. Carey, "Optimal Time-Varying Flows on Congested Networks", *Operations Research 35 (1), pp. 58-69, 1987.*

[18] E. Cascetta, G.E.Cantarella, A doubly dynamic assignment model, *PTRC, Brighton U.K. 1989.*

[19] E. Cascetta, G.E.Cantarella, M. Di Gangi, "Evaluation of Control Strategies trough a Doubly Dynamic Assignment Model", *Presented at the 70th Annual Meeting of the Transportation Research Board, Washington DC., USA 13-17 January 1991.*

[20] I. Catling, "Advanced Technology for Road Transport, IVHS and ATT", *Artech House, Boston-London, ISBN-0-89006-613-2, 1994.*

[21] K.B. Davidson, "A Flow Traveltime Relationship for Use in Transportation Planning", *Proceedings 3rd Australian Road Research Board Conference, Part 1, 1966, pp 183-194.*

[22] G.W. Euler, "IVHS: Definitions and Applications", *ITE Journal, November 1990.*

[23] Foley, van Dam, Feiner & Hughes, "Computer Graphics: Principles and Practice", *Addison Wesley System Programming Series, 1990.*

[24] T.L. Friesz, J. Luque, R.L. Tobin & B. Wie, "Dynamic Network Traffic Assignment Considered as a Continuous Time Optimal Time Control Problem", *Operations Research 37 (6), pp. 893-901, 1989.*

[25] D.C. Gazis, "Car Following Theory of Steady State Traffic Flow", *Operations Research, No 7. pp 499-505, 1961.*

[26] M.O. Ghali & M.J. Smith, "Approximation to Optimal Dynamic Traffic Assignment of peak Period Traffic to a Congested City Network", *Second International Capri Seminar on Urban Traffic Networks, 5-8 July 1992, Italy.*

[27] H.J.M. van Grol, A.F. Bakker, "Shortest path finding: from algorithm to special purpose hardware", *Proceedings of Seminar C, "Transportation Planning Methods" at the 17th PTRC Transport & Planning Summer Annual Meeting, University of Sussex, England, 11-15 September 1989.*

[28] H.J.M. van Grol, A.F. Bakker, "A Special-Purpose Parallel Computer for Traffic Simulation", *Transportation Research Record 1306, pp 40-48, 1991.*

[29] H.J.M. van Grol, "Traffic Assignment Problems solved by Special Purpose Hardware with Emphasis on Real Time Applications", *Ph.D. thesis, ISBN:90-9005441-3, 1992.*

[30] M.D. Hall, D. van Vliet & L. Willumsen, "SATURN: a simulation-assignment model for the evaluation of traffic management schemes", *Traffic Engineering and Control, pp 168-176, 1980.*

[31] R. Hamerslag, P.C.H. Opstal, "A Three-dimensional assignment method in Timespace", *Rapport 87-94 Faculty of Mathematics and Informatics, Delft University of Technology, The Netherlands, 1987.*

[32] R. Hamerslag, "Dynamic Assignment in the Three Dimensional Timespace, Mathematical Specification and Algorithm", *USA-Italy seminar on transportation networks, Napoli & Capri, 1989.*

[33] R.Hamerslag, "Dynamic Assignment in the Three Dimensional Timespace", *Transportation Research Record 1220, 1989.*

[34] Heidemij advies, "Evaluation homogenizing A2" (in Dutch), Rotterdam, June 1993. Rijkswaterstaat, Directie Utrecht, Adviesdienst Verkeer en Vervoer.

[35] C. Hendrickson & E. Plank, "The Flexibility of Departure Times for Work Trips", *Transportation Research, 18A, pp 25-36, 1984.*

[36] Highway Research Board, "Highway Capacity Manual, *Special Report 87 (1965).*

[37] Intercai, BSO AT, "Functionele Specificatie Monitoring Casco" (in Dutch: Functional Specification Monitoring Casco), *DHV august 1990.*

[38] B.N. Janson, "A Convergent Algorithm for Dynamic Traffic Assignment", *Presented at the 70th Annual Meeting of the Transportation Research Board, Washington DC., USA 13-17 January 1991.*

[39] B.N. Janson, "Dynamic Traffic Assignment with arrival time costs", *Transportation and Traffic Theory, pp 127-146, 1993.*

[40] D. Judycki & G Euler, "The IVHS program in the United States", ITE 1993 The Hague, The Netherlands.

[41] I.A. Kaysi, "Framework and Models for the Provision of Real-Time Driver Information", *Ph.D. dissertation, MIT, February 1992.*

[42] J.J. Klijnhout, "IVHS/RTI/ATT in Europe", ITE 1993, The Hague, The Netherlands.

[43] E. Kroes & R. Hamerslag, "A Practical Dynamic Network Equilibrium Model with Elastic Departure Times", Transportation Research Board, 69th Anual Meeting, Washington D.C., 1990. (paper 890788)

[44] J. Lambrechtsen & M. Westerman, "Dynamic Traffic Management in the Netherlands" (in Dutch), Delft University of Technology, The Netherlands, October 1993. ISN: LVV report 0920-0592.

[45] D.R. Leonard, P. Gower, N.B. Taylor, "CONTRAM; Structure of the model", *Research Report 178, TRRL Crowthorne 1989.*

[46] M.J. Lighthill, G.B. Witham, "On Kinematic Waves II. A theory of traffic flow on long crowded roads", *Proc. Roy. Soc., Series A, No 1178, Vol 229, pp 317-345, 1955, London.*

[47] J.F. Luque & T.L. Friesz, "Dynamic Traffic Assignment Considered as a Continuous Time Optimal Control Problem", *Paper presented at the TIMS/ORSA Joint National Meeting, Washington , D.C.,May 5-7, 1980.*

[48] A.S. Lyrintzis, G. Liu & P.G. Michalopoulos, "Development and Comparative Evaluation of High-Order Traffic Flow Models", *Transportation Research Board, 73rd Annual Meeting, January, 1994, Wahington, USA.*

[49] P.C.F. Maas, "Design and Implementation of Visualization Methods for Traffic Assignment Methods", *Masters Thesis, Delft University of Technology, Faculty of Applied Physics, Section Physics Informatics, Subsection Computational Physics, December 1993.*

[50] H.S. Mahmassani & R. Herman, "Dynamic User Equilibrium Departure Time and Route Choice on Idealized Traffic Arterials", *Transportation Science, 18, pp 362-384, 1984.*

[51] H.S. Mahmassani & S. Peeta, "Network Performance Under System Optimal and User Equilibrium Dynamic Assignments: Implications for Advanced Traveler Information Systems", *Transportation Research Record 1408, pp 83-93, Washington, D.C, 1993.*

[52] R.E. Maki & D.M. DeVaughn, "Guidelines For Implementing ATMS".

[53] A.D. May, "Traffic Flow Fundamentals", *Prentice Hall, Inc. ISBN: 0-13-926072-2, 1990.*

[54] D.K. Merchant & G.L. Nemhauser, "A Model and an Algorithm for the Dynamic Traffic Assignment Problem", *Transportation Science 12, pp. 183-199, 1978.*

[55] D.K. Merchant & G.L. Nemhauser, "Optimality Conditions for a Dynamic Traffic Assignment Model", *Transportation Science 12, pp. 200-207, 1978.*

[56] P.G. Michalopoulos & D.E. Beskos, "Improved continuum models of freeway flow", *Ninth International Symposium on Transportation and Traffic theory, VNU Science Press, pp 89-111, 1984.*

[57] P.G. Michalopoulos, P. Yi, P. Beskos & A.S. Lyrintzis, "Continuum Modeling of Traffic Dynamics", *Proceedings of the 2nd International Conference on Applications of Advanced Technologies in Transportation Engineering", Minneapolis, Minnesota, pp 36-40, August 18-21, 1991.*

[58] P.G. Michalopoulos, P. Yi & A.S. Lyrintzis, "Development of an Improved High Order Continuum Traffic Flow Model.", *Transportation Research Record 1365, pp.125-132, 1993.*

[59] F. Middelham, Ramp metering, applications elsewhere and in The Netherlands" (in Dutch), *Syllabus PAO cursus "Dynamisch Verkeersbeheer", Rotterdam 1991, Rijkswaterstaat, Dienst Verkeerskunde.*

[60] N.L. Nihan & K.O. Holmesland, "Use of the Box and Jenkins Time Series Technique in Traffic Forecasting', *Transportation nr.9, pp 125-142, Elsevier Scientific Publishing Company, Amsterdam, 1980.*

[61] J.de D. Ortúzar & L.G. Willumsen, "Modelling Transport" *Second Edition, JohnWiley & Sons, Chichester, England, ISBN 0 471 94193 X, 1990*

[62] D. Owens & M.J. Schofield, "Access control on the M6 motorway: evaluation of Britains' first ramp-metering scheme", *Traffic Engineering and Control, December 1988.*

[63] A. de Palma, M. Ben-Akiva, C. Lefevre & N. Litinas, "Stochastic Equilibrium Model of Peak Period Traffic Congestion", *Transportation Science, 17, pp 430-453, 1983.*

[64] M. Papageorgiou, "A Hierarchical Control System for Freeway Traffic", *Transportation Research, Vol. 17B, pp.251-261, 1983.*

[65] M. Papageorgiou, J-M Blosseville, H. Hadj-Salem, "Macroscopic Modelling of Traffic Flow on the Southern Part of Boulevard Peripherique in Paris", *Transportation Research B, Vol 23B, No 1, pp 29-47, 1989.*

[66] M. Papageorgiou, J-M Blosseville, H. Hadj-Salem, "Modelling and Real-time Control of Traffic Flow on the Southern Part of Boulevard Peripherique in Paris. Part I: Modelling", *Transportation Research A, Vol 23A, No 5, pp 345-359, 1990.*

[67] H.J. Payne, "Models of Freeway Traffic and Control", *Mathematical Models of Public Systems, Simulation Counc. proc. Ser., 1. pp 51-61, 1971.*

[68] H.J. Payne, "FREFLO: A Macroscopic Simulation Model of Freeway Traffic", *Transportation Research Records 722, pp 86-77, 1977.*

[69] S. Peeta, "System Optimal Dynamic Traffic Assignment in Congested Networks with Advanced Information Systems", *Ph.D. Dissertation, The University of Texas at Austin, May 1994.*

[70] D.H. Polak and T. Heijer, "Control Strategies for a Highway Network, The development and implementation of a theoretical model for local traffic control", *Control Strategies for a highway Network, 2nd Concept, Leidschendam, SWOV Institute for Road Safety Research, The Netherlands, December 1993.*

[71] B. Przebinda, "Comparison of a few Speed-Density functions in Application to Dynamic Traffic Management.", *A study report from Training in Placement at Delft University of Technology under TEMPUS "TREP" Joint European Project. Delft September 1991.*

[72] B. Ran, D.E. Boyce, L.J. LeBlanc,"An Instantaneous Dynamic User-Optimal Traffic Assignment Model and Solution Algorithm", *submitted to Operations research,* 1990.

[73] B. Ran, N. Rauphail, A. Tarko & D.E. Boyce, "Towards a set of Dynamic Link Traveltime Functions for Dynamic Traffic Assignment", *39th North American Meeting of the Regional Science Association International, 13-15 November 1992, Chicago Illinois USA*

[74] E. de Romph, H.J.M. van Grol & R. Hamerslag, "A Dynamic Assignment Model for Short Term Prediction", *Second International Capri Seminar on Urban Traffic Networks, 5-8 July 1992, Italy.*

[75] E. de Romph, H.J.M. van Grol & R. Hamerslag, "3DAS -3Dimensional ASsignment- A dynamic assignment model for short term predictions", *39th North American Meeting of the Regional Science Association International, 13-15 November 1992, Chicago Illinois USA.*

[76] E. de Romph, H.J.M. van Grol & R. Hamerslag, "Application of 3DAS (3-Dimensional Assignment) in the Washington Metropolitan Area", *73rd Transportation Research Board Annual Meeting, Washington, January 1994.*

[77] G. Rose, M.A.P. Taylor & P. Tisato, "Estimating Travel Time Functions for Urban Roads: Options and Issues", *Transportation Planning and Technology, 1989, Vol 14, pp 63-82.*

[78] C.A. Rothrock & L.A.Keefer, "Measurement of Urban Traffic Congestion", *Highway Research Board, Bulletin 156, pp 1-13, 1957.*

[79] H.Salem, "ALINEA: un outil de regulation d'acces isole sur autoroute", *Institut National de Recherche sur les Transports et leur Sécurité (INRETS), Paris, October 1988.*

[80] A.J. Santiago, "ATMS Technology What We Know and What We Don't Know", *IVHS Research Division, Federal Highway Administration, MCLean, Virginia, 1992.*

[81] H. Schuurman & R.G.G.M.Vermijs, "Verkeersafwikkeling bij discontinuiteiten" (in Dutch: Level of service at freeways), *Report TU Delft, VK 2205.305, ISSN: LVV report 0920-0592. Delft, June 1993.*

[82] H. Schuurman & R.G.G.M.Vermijs, Simulatie benuttingsmaatregelen, "Effectenstudie m.b.v. het mikrosimulatiemodel FOSIM" (in Dutch: Simulation of utilisation measures, a study of effects using the microsimuation model FOSIM), *Report TU Delft, VK 2205.309, ISSN: LVV report 0920-0592, Delft, March 1994.*

[83] S. Schwerdtfeger, "DYNEMO: A model for the simulation of traffic flow in motorway networks", *Ninth International Symposium on Transportation and Traffic theory, VNU Science Press, pp 89-111, 1984.*

[84] Y. Sheffy, "Urban Transportation Networks", *Englewood Cliffs, NJ,: Prentice Hall, 1985.*

[85] M.J. Smith, "A New Dynamic Traffic Model and the Existence and Calculation of Dynamic User-Equilibria on Congested Capacity-Constrained Road Networks", *Transportation Research Board, 71st Annual Meeting, January 12-16, 1992, Washington, USA.*

[86] S.Smulders, "Modelling and Filtering of Freeway Traffic Flow", *Report OS-R8706, Centre for Mathematics and Computer Science, November 1988.*

[87] S. Suh, C. Park & T.J. Kim, "A Highway Capacity Function in Korea: Measurement and Calibration", *Transportation Research 24A, No 3, pp 177-186, 1990.*

[88] J.M. Sussman, "IVHS, A Challenge Awaits the Transportation and OR/MS community", *OR/MS today, December 1992.*

[89] M.L.G. Thoone, "CARIN, een navigatie en informatie systeem voor auto's" (in Dutch), *Philips Technisch Tijdschrift, Jaargang 43, nr.11/12, sept. 1987.*

[90] C. Thor, "IVHS: Boon or Boondoggle?", *TR News, May-June 1992.*

[91] TRANSPUTE, Consult over het Basisscenario Verkeersbeheersing, *(in Dutch) uitgracht aan hoofddirectie Rijkswaterstaat, 1992.*

[92] J. Upchurch, "Status of IVHS Activities in Japan", *ITE 1993, The Hague, The Netherlands*

[93] .M. VanAerde & S. Yager, "Modelling Dynamic Integrated Freeway/Traffic Signal Networks: A Proposed Routing-Based Approach", *Transportation Research A, 22A, nr6, 1988.*

[94] R.G.G.M. Vermijs, "Het mikrosimulatiemodel FOSIM; beschrijving van het komputerprogramma" (in Dutch: The micro simulation model FOSIM; description of the software), *Transportation Research Laboratory, Report TU Delft, VK 2205.308, Oktober 1993.*

[95] D. Van Vliet, "SATURN, A modern assignmnent model", *Traffic Engineering & Control 23, pp. 578-581, 1982.*

[96] P.C. Vythoulkas, "A Dynamic Stochastic Assignment Model for the Analysis of General networks", *USA-Italy seminar, "Urban Traffic Networks: Dynamic Control and Flow Equilibrium", Napoli & Capri, 1989.*

[97] J.G. Wardrop, "Some Theoretical Aspects of Road Traffic Research", *Proceedings, Institute of civil Engineering II(1), pp 325-378.*

[98] D. Watling, "Urban Traffic Network Models and Dynamic Driver Information Systems", *Institute for Transportation Studies, University of Leeds, forthcoming in Transportation Studies, 1994.*

[99] M. Westerman, R.Hamerslag, "Collecting the Dynamics of Metropolitan Freeway Systems: A Simulation Study", *Proceedings of 26th ISATA conference, Aachen, Germany, pp 807-814, 1993*

[100] D.A.Wicks & E.B.Lieberman, "Development and Testing of INTRAS, a microscopic freeway simulation model", *vol 1, Program design, parameter calibration and freeway dynamics component development, Final report, National technical informaiton service, Springfield, Va.22161, october 1980.*

[101] B.W. Wie, "Dynamic Analysis of User Optimized Network Flows with Elastic Travel Demand", *Presented at the 70th Annual Meeting of the Transportation Research Board, Washington DC., USA 13- 17 January 1991.*

[102] D. van Willigen, J.J. Reijmers, *Verkeersbegeleidingssystemen I & II (in Dutch for Traffic Control Systems I & II)*, Delft University of Technology, Delft 1992/1993

[103] Witteveen & Bos, RIA logging May, August and September 1992" (in Dutch), Deventer 1992.

[104] N.J. van der Zijpp & R. Hamerslag, "The Real Time Estimation of Origin-Destination Matrices for Freeway Corridors", *26th International Symposium on Automotive technology and Automation ISATA, Aachen, Germany 13th-17th September 1993.*

[105] N.J. van der Zijpp & R. Hamerslag, "An Improved Kalman Filtering Approach to Estimate Origin-Destination Matrices for Freeway Corridors", *73rd Transportation Research Board Annual Meeting, Washington, January 1994, forthcoming in Transportation Research Records.*

REFERENCES

APPENDIX **A**

AMSTERDAM MONITORING SYSTEM

The data for several tests described troughout this dissertation was derived from the data collection system at the beltway of Amsterdam.

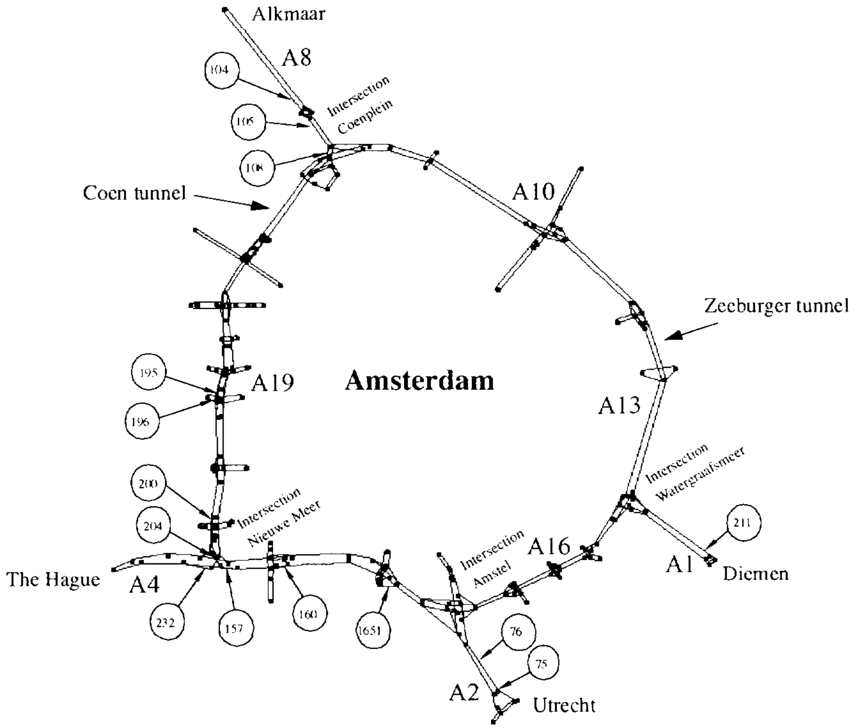


Figure 87: Amsterdam beltway with induction loops, diameter of this network is ±9km.

This network is equipped with several hundreds of induction loops used mainly for the Motorway Control and Signalling System (MCSS). (MCSS is described in chapter 2, paragraph 2.4.1.) The data from these induction loops are collected at screenline level and aggregated to one minute data by road side computers. Speeds (km/hr), volumes (veh/minute) and the status of signalling signs are measured. The one minute data is on-line available at the control centre and can be downloaded for research purposes.

Figure 87 shows the Amsterdam beltway. The location of the induction loops used throughout this dissertation are marked by their station numbers. The A numbers indicate the national road numbers. The location of all the induction loops of the control and signalling system at the Amsterdam beltway is given in figure 88.

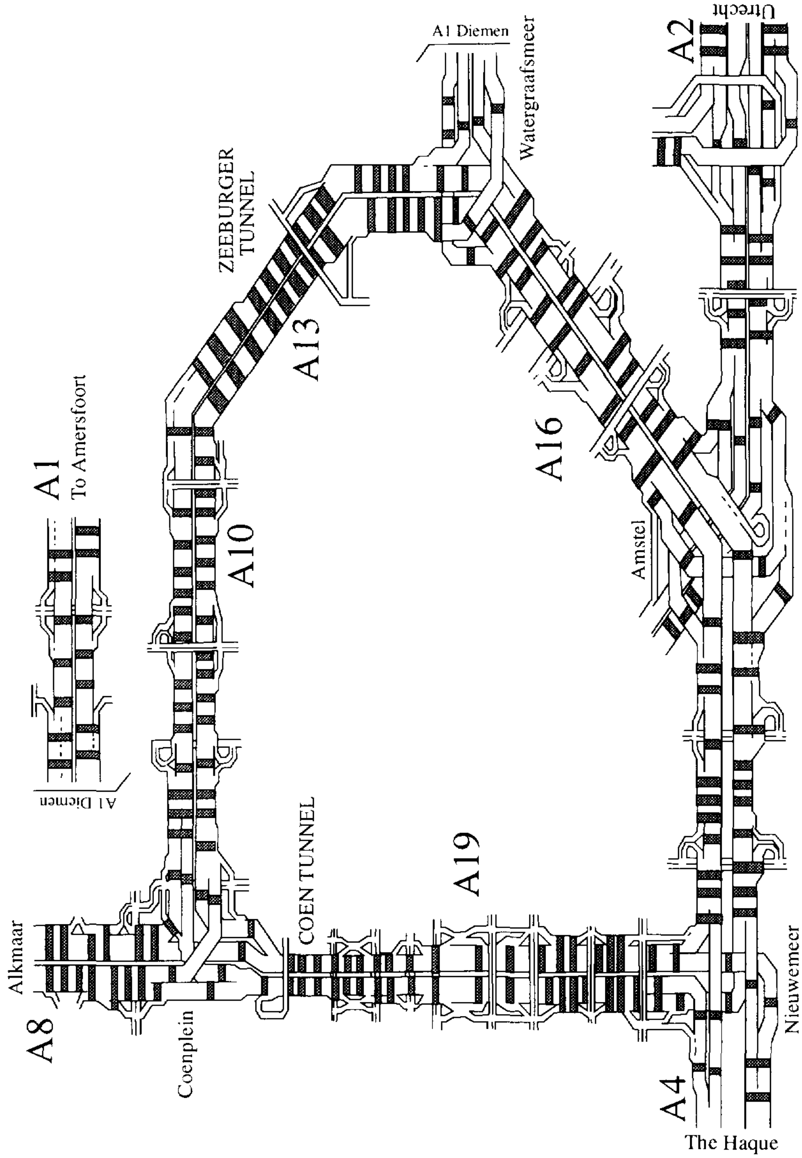


Figure 88: Location of all the induction loops on the Amsterdam beltway.

APPENDIX B

SPEED-DENSITY FUNCTION PARAMETERS

Table 11: Estimated parameters avaraged over 10 days

nr	loop	max dens	crit dens		max speed		alpha		beta		congested days
1	211	100.00	24.00	0.00	106.59	3.43	0.36	0.27	0.00	0.00	0
2	2111	200.00	45.31	5.38	97.67	2.52	0.32	0.22	1.53	0.05	2
3	215	200.00	47.20	1.81	104.49	1.47	0.25	0.15	1.36	0.04	2
4	2151	100.00	21.70	6.57	125.74	13.01	8.35	3.25	0.32	0.00	1
5	77	200.00	48.00	0.00	112.19	3.06	1.50	0.73	0.00	0.00	0
6	76	200.00	48.00	0.00	114.07	2.85	1.45	0.73	0.00	0.00	0
7	761	200.00	46.84	2.03	107.35	1.65	0.60	0.13	1.41	0.00	1
8	75	300.00	70.25	4.83	109.98	1.82	0.60	0.14	1.12	0.06	2
9	61	200.00	48.00	0.00	120.00	0.00	1.00	0.00	0.00	0.00	0
10	62	200.00	48.00	0.00	104.30	1.90	0.63	0.37	0.00	0.00	0
11	232	200.00	48.09	7.91	108.16	2.00	0.62	0.19	0.65	0.00	1
12	231	200.00	46.59	4.23	111.16	1.54	0.62	0.12	0.79	0.00	1
13	2311	200.00	48.00	0.00	104.95	1.60	0.74	0.18	0.00	0.00	0
14	111	200.00	48.00	0.00	103.80	1.79	0.32	0.27	0.00	0.00	0
15	110	200.00	48.00	0.00	106.28	1.68	0.63	0.27	0.00	0.00	0
16	109	200.00	48.00	0.00	108.94	1.83	0.94	0.18	0.00	0.00	0
17	1091	200.00	38.72	13.55	107.31	3.44	1.02	0.51	0.69	0.10	7
18	108	200.00	48.00	0.00	110.42	1.95	1.03	0.31	0.00	0.00	0
19	1081	200.00	35.23	9.10	108.46	2.21	0.99	0.29	0.77	0.03	7
20	107	300.00	63.60	10.39	112.06	5.43	0.82	0.24	0.69	0.03	4
21	106	300.00	60.23	9.84	107.74	1.88	0.53	0.20	0.72	0.05	5
22	1061	100.00	21.70	4.65	113.83	5.56	5.12	1.15	0.00	0.00	0

Table 11: Estimated parameters avaraged over 10 days

nr	loop	max dens	crit dens		max speed		alpha		beta		congested days
23	105	300.00	64.51	7.83	108.39	1.91	0.61	0.14	0.79	0.04	6
24	104	300.00	64.19	7.66	107.10	2.08	0.54	0.13	0.77	0.05	6
25	136	300.00	69.31	8.07	107.03	2.31	0.43	0.31	0.78	0.00	1
26	135	300.00	69.08	8.76	108.17	2.36	0.49	0.34	0.81	0.00	1
27	134	300.00	72.00	0.00	106.71	2.35	0.53	0.35	1.33	0.00	1
28	132	300.00	72.00	0.00	105.24	2.15	0.41	0.25	0.00	0.00	0
29	131	300.00	72.00	0.00	109.29	2.97	0.56	0.70	0.00	0.00	0
30	130	300.00	72.00	0.00	110.39	2.97	0.49	0.61	0.00	0.00	0
31	128	300.00	72.00	0.00	132.24	1.87	0.57	0.29	0.00	0.00	0
32	127	300.00	72.00	0.00	113.28	2.03	0.49	0.26	0.00	0.00	0
33	126	300.00	72.00	0.00	108.81	2.20	0.29	0.30	0.00	0.00	0
34	49	300.00	72.00	0.00	108.41	2.15	0.20	0.20	0.00	0.00	0
35	46	300.00	72.00	0.00	105.98	1.39	0.54	0.23	0.00	0.00	0
36	44	300.00	72.00	0.00	111.54	1.60	0.58	0.30	0.00	0.00	0
37	42	300.00	72.00	0.00	112.52	2.19	0.53	0.33	0.00	0.00	0
38	40	300.00	72.00	0.00	108.23	2.71	0.40	0.29	0.00	0.00	0
39	39	300.00	72.00	0.00	110.78	5.38	0.55	0.35	0.00	0.00	0
40	36	300.00	72.00	0.00	109.96	1.77	0.66	0.49	0.00	0.00	0
41	35	300.00	72.00	0.00	108.96	7.59	0.69	0.54	0.00	0.00	0
42	34	300.00	72.00	0.00	106.91	2.41	0.47	0.61	0.00	0.00	0
43	33	300.00	72.00	0.00	108.96	2.13	0.41	0.37	0.00	0.00	0
44	32	300.00	72.00	0.00	134.16	2.26	1.04	0.49	0.00	0.00	0
45	321	100.00	24.00	0.00	105.40	2.47	0.91	0.16	0.00	0.00	0
46	31	300.00	72.00	0.00	105.74	3.33	0.66	0.27	0.00	0.00	0
47	174	300.00	68.08	8.76	109.83	1.13	0.53	0.31	0.63	0.00	1
48	1741	100.00	22.00	2.24	105.27	5.31	2.31	1.22	0.00	0.00	2
49	173	300.00	60.50	14.77	104.84	1.44	0.35	0.11	0.55	0.12	2
50	1731	100.00	23.76	0.30	105.70	3.80	2.45	0.87	0.17	0.03	2
51	172	300.00	67.84	8.35	105.04	2.95	0.67	0.23	0.75	0.06	2
52	171	300.00	68.10	7.91	110.31	7.10	1.00	0.75	0.70	0.00	1
53	1711	100.00	23.51	1.47	117.61	20.53	3.29	4.54	0.31	0.00	1
54	170	300.00	72.00	0.00	126.29	4.44	1.48	0.95	0.00	0.00	0
55	169	300.00	64.14	12.05	107.16	3.55	0.66	0.22	0.82	0.15	4
56	1691	100.00	23.16	1.89	109.86	14.50	3.98	3.78	0.59	0.22	4
57	168	300.00	62.51	14.53	112.27	1.52	0.75	0.15	0.79	0.06	3
58	167	300.00	61.92	15.40	107.97	1.32	0.46	0.16	0.90	0.02	2

Table 11: Estimated parameters avaraged over 10 days

nr	loop	max dens	crit dens		max speed		alpha		beta		congested days
59	166	300.00	68.42	10.74	110.24	2.01	0.59	0.19	1.05	0.00	1
60	165	300.00	72.00	0.00	102.21	2.66	0.61	0.27	0.00	0.00	0
61	1651	100.00	26.06	2.88	103.39	2.67	1.04	0.19	0.00	0.00	0
62	164	300.00	72.00	0.00	108.73	1.75	0.60	0.13	1.11	0.00	1
63	163	300.00	82.67	9.21	103.77	1.66	0.49	0.10	1.01	0.00	1
64	161	300.00	76.60	3.53	107.96	1.41	0.54	0.08	0.88	0.11	4
65	160	300.00	70.48	3.51	108.29	1.42	0.58	0.09	0.88	0.15	5
66	159	300.00	71.45	1.97	106.49	1.74	0.62	0.13	0.80	0.03	2
67	158	300.00	73.22	8.82	102.74	2.17	0.59	0.11	0.76	0.11	4
68	1581	100.00	32.19	13.38	101.70	2.38	1.08	0.19	0.53	0.13	3
69	157	400.00	99.61	16.11	95.95	2.01	0.76	0.10	0.56	0.11	5
70	204	200.00	55.92	5.28	89.76	2.04	0.37	0.09	0.63	0.09	8
71	203	200.00	54.20	5.13	99.84	1.83	0.63	0.07	0.65	0.07	8
72	202	200.00	45.07	4.98	105.51	1.51	0.84	0.25	0.00	0.00	0
73	2021	200.00	44.77	2.77	100.14	1.83	0.65	0.09	0.70	0.03	8
74	201	200.00	41.25	10.58	107.53	2.44	0.74	0.24	0.58	0.13	3
75	2011	200.00	46.98	3.98	102.58	1.95	0.72	0.10	0.81	0.06	6
76	200	200.00	42.66	8.44	105.19	2.30	0.57	0.19	0.88	0.36	4
77	2001	200.00	48.26	2.36	98.91	1.47	0.76	0.08	0.92	0.09	5
78	199	400.00	91.09	6.68	99.65	1.89	0.70	0.07	0.85	0.08	4
79	198	300.00	70.50	4.50	106.04	1.64	0.61	0.13	0.94	0.00	1
80	197	300.00	69.16	5.82	104.09	1.84	0.52	0.11	0.87	0.02	2
81	1971	100.00	23.68	0.96	101.90	1.85	1.51	0.28	0.26	0.03	2
82	196	300.00	69.04	5.98	102.13	1.68	0.49	0.09	0.83	0.02	2
83	1961	100.00	23.25	1.29	100.80	1.80	1.74	0.27	0.28	0.03	2
84	195	300.00	67.07	9.87	104.08	1.65	0.52	0.11	1.00	0.26	3
85	194	300.00	69.51	7.47	101.17	1.84	0.50	0.16	0.00	0.00	0
86	193	300.00	72.00	0.00	102.59	1.45	0.61	0.10	0.00	0.00	0
87	192	300.00	72.00	0.00	101.50	1.94	0.55	0.08	0.00	0.00	0
88	191	100.00	24.00	0.80	111.63	5.14	1.77	1.01	1.15	1.48	3
89	190	100.00	22.66	4.02	109.50	2.97	0.97	0.51	1.13	0.00	1
90	1901	200.00	47.97	0.09	105.50	2.72	0.59	0.31	0.00	0.00	0
91	81	200.00	46.71	2.86	96.50	1.67	0.29	0.16	1.19	0.49	2
92	91	100.00	22.01	2.80	92.41	2.44	0.50	0.38	0.66	0.04	2
93	233	200.00	46.75	3.75	92.73	1.12	0.33	0.12	0.73	0.00	1
94	235	200.00	48.00	0.00	94.33	1.49	0.36	0.18	0.00	0.00	0

Table 11: Estimated parameters avaraged over 10 days

nr	loop	max dens	crit dens		max speed		alpha		beta		congested days
95	113	200.00	36.37	10.12	105.86	2.75	0.90	0.44	0.78	0.06	7
96	112	100.00	24.00	0.00	107.84	4.15	4.00	0.86	0.00	0.00	0
97	222	100.00	24.00	0.00	100.44	2.21	0.37	0.22	0.00	0.00	0
98	59	200.00	48.00	0.00	87.94	1.09	0.11	0.16	0.00	0.00	0
99	58	200.00	49.78	3.73	109.24	3.94	0.82	1.65	0.00	0.00	0
100	224	200.00	60.20	7.80	99.71	1.60	0.56	0.07	0.00	0.00	0

The loopnumbers in the second column determine the location on the map in appendix A. The columns with an empty header contains the standard deviation for the preceeding parameter. The locations with nr: 4, 9, 39, 41, 48, 50, 53, 56, 88, 96 and 99 were not used to generate table 4 in paragraph 5.3.4, because these locations showed detection errors or other anomalities.

APPENDIX C

DEFINITION OF THE TEST NETWORKS

Test network: bottleneck

This network has eleven nodes and ten links. The first four links are 3 lane and the last six links are 2 lane. Each link has a length of one kilometre. Traffic travels from left to right. There is one OD-relation.:

Problem specification

nr of periods	40
period length	0.5 minute

:

Network specification

link	length	max dens	max speed	Function	type
1	1 km	375 veh/km	120 km/hr	Smulders	1
2	1 km	375 veh/km	120 km/hr	Smulders	1
3	1 km	375 veh/km	120 km/hr	Smulders	1
4	1 km	375 veh/km	120 km/hr	Smulders	1
5	1 km	375 veh/km	120 km/hr	Smulders	1
6	1 km	250 veh/km	120 km/hr	Smulders	2
7	1 km	250 veh/km	120 km/hr	Smulders	3
8	1 km	250 veh/km	120 km/hr	Smulders	3
9	1 km	250 veh/km	120 km/hr	Smulders	3
10	1 km	250 veh/km	120 km/hr	Smulders	3

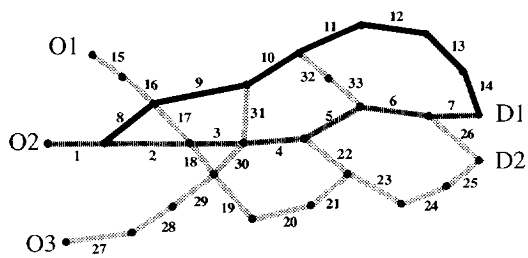
The Smulders function has a critical density of 24% of the maximum density.
During the first 15 minutes traffic departs from the origin. Each minute represents two periods, and the number of cars depart in two equal halfs during the two periods in each minute.

.

Departure values

minute	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
#cars	40	56	72	88	104	104	104	104	104	88	72	56	40	20	8	0	0	0	0	0

Test network: alternative routes



In this network the traffic travels from left to right. There are 34 links, numbered in the figure. There are three origins and two destinations. Every OD-relation uses the same departure table.

Problem specification

nr of periods	30
periodlength	0.5 minute

All the links in the network have a maximum density of 200 vehicles per kilometre, and a maximum speed of 90 km/hr. The speed density relationship is the default Smulders function.

The length of each link is given in the following table.

Network specification

link	length
1	1.20
2	2.00
3	1.20
4	1.30
5	1.47
6	1.51
7	1.20
8	1.50
9	2.14
10	1.30
11	1.52

link	length
12	1.51
13	1.13
14	1.08
15	0.86
16	0.92
17	1.20
18	0.86
19	1.35
20	1.33
21	1.06
22	1.28

link	length
23	1.39
24	1.08
25	1.00
26	1.56
27	1.51
28	1.08
29	1.14
30	0.99
31	1.30
32	1.00
33	0.85

The OD-table is determined by the departure value in the following table
:

Departure values

period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
O1-D1	3	3	4.2	4.2	5.4	5.4	6.6	6.6	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8
O1-D2	3	3	4.2	4.2	5.4	5.4	6.6	6.6	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8
O2-D1	8	8	11.2	11.2	14.4	14.4	17.6	17.6	20.8	20.8	20.8	20.8	20.8	20.8	20.8	20.8	20.8	20.8
O2-D2	3	3	4.2	4.2	5.4	5.4	6.6	6.6	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8
O3-D1	3	3	4.2	4.2	5.4	5.4	6.6	6.6	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8
O3-D2	3	3	4.2	4.2	5.4	5.4	6.6	6.6	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8

APPENDIX C: DEFINITION OF TEST NETWORKS

APPENDIX D

CALCULATION OF TRAVEL TIME

The experiment in paragraph 5.2.1 on page 68 uses a triangular density distribution. This density distribution is symmetrical around the point $x=½L$. At the beginning of the road the density is zero and increases towards $\alpha\rho^{max}$ in the middle of the road, than it decreases linearly back to zero, α is between zero and one. As the distribution is symmetrical, travel time is two times the travel time over the first half of the road. The density distribution is shown in figure 89.

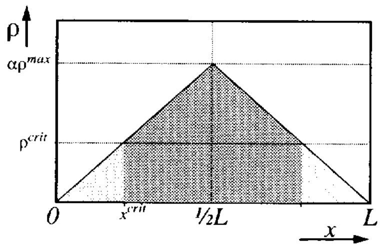


Figure 89: Density distribution on a link of length L

The distance along the link is denoted by x. The density distribution is given in equation 89.

$$\frac{\rho}{\rho^{max}} = \frac{2\alpha}{L}x \qquad 0 \le x \le ½L \tag{eq 89}$$

The travel time to traverse the link is given in equation 66 on page 68. This equation is repeated in equation 90:

$$t_{se}^{\infty} = \lim_{n \to \infty} t_{se}^{n} = \lim_{n \to \infty} \sum_{a \in P_{se}} \frac{L/n}{v(\rho_a)} = \int_0^L \frac{dx}{v(\rho(x))} \tag{eq 90}$$

1. This appendix is taken from P.C.F. Maas [49].

The Smulders speed-density function is used to calculate the travel time to traverse the link. This function is given in equation 21 on page 38, it has a linear and a non-linear part. The linear part changes into the non linear part at the critical density, ρ^{crit}, with $\rho^{crit} = ¼\rho^{max}$.
At the moment the density on the link exceeds the critical density the non linear equation to calculate the travel time is used. The density exceeds the critical density when $\alpha > ¼$. So for $0 \le \alpha \le ¼$ the linear part $v_1(\rho)$ of equation 21 must be used and for $¼ \le \alpha \le 1$ the non linear part $v_2(\rho)$ of equation 21 must be used.

$$v_1(\rho) = \left(1 - \frac{\rho}{\rho^{max}}\right)v_0 \qquad v_2(\rho) = ¼\left(\frac{\rho^{max}}{\rho} - 1\right)v_0 \qquad \text{(eq 91)}$$

v_0 denotes the free flow speed.
The following two equations are obtained by substituting equation 89 and 91 in equation 90:

for $0 \le \alpha \le ¼$:

$$t_\infty = 2\int_0^{½L} \frac{1}{v_1(\rho)}dx = \frac{2}{v_0}\int_0^{½L} \frac{1}{1 - \frac{\rho}{\rho^{max}}}dx = \frac{2}{v_0}\int_0^{½L} \frac{1}{1 - \frac{2\alpha}{L}x}dx$$
$$= -\frac{L}{\alpha v_0}\frac{2}{v_0}\ln\left(1 - \frac{2\alpha}{L}x\right)\Bigg|_0^{½L} = -\frac{L}{\alpha v_0}\ln(1-\alpha) \qquad \text{(eq 92)}$$

for $¼ \le \alpha \le 1$:

$$t_\infty = 2\int_0^{x^{crit}} \frac{1}{v_1(\rho)}dx + 2\int_{x^{crit}}^{½L} \frac{1}{v_2(\rho)}dx = 2t_{1,\infty} + 2t_{2,\infty} \qquad \text{(eq 93)}$$

The point where the density exceeds the critical density ($\rho > \rho^{crit}$) is called x^{crit}.

$$x^{crit} = \frac{\rho^{crit}}{\rho^{max}} \cdot \frac{L}{2\alpha} = \frac{L}{8\alpha} \qquad \text{(eq 94)}$$

substituting equation 94 into equation 93 and solving the integrals, results in:

$$2t_{1,\infty} = -\frac{L}{\alpha v_0}\ln(¾) \qquad 2t_{2,\infty} = -\frac{4L}{v_0}\left(1 - \frac{1}{4\alpha} - \frac{1}{\alpha}\ln\frac{¾}{1-\alpha}\right) \qquad \text{(eq 95)}$$

Using equation 95, 93 and 92, results in the following expression for the actual travel time:

$$t_\infty = \begin{cases} \dfrac{L}{\alpha v_0}\ln\left(\dfrac{1}{1-\alpha}\right) & 0<\alpha<\tfrac{1}{4} \\[2ex] \dfrac{4L}{v_0}\left(\dfrac{1}{4\alpha}-1+\dfrac{1}{\alpha}\ln\dfrac{3/4}{1-\alpha}-\dfrac{1}{4\alpha}\ln\left(3/4\right)\right) & \tfrac{1}{4}<\alpha<1 \end{cases} \qquad \text{(eq 96)}$$

APPENDIX E

MATHEMATICAL SYMBOLS

Various sets

N	= set of all nodes
A	= set of all links
D	= set of all departure periods
Z	= set of all zones (origins and destinations)

List of sub and superscripts

a	= link between node i and node j
d	= departure period
ij	= link between node i and node j
k	= path
n	= node
p	= period
r	= origin
s	= destination
t	= time (moment)

List of variables

Capitals

I_n^p	= trips generated at node n in period p.
$I_n(t)$	= trips generated at node n at time t.
O_n^p	= trips attracted by node n in period p.
$O_n(t)$	= trips attracted by node n at time t.
T	= period length.

Lower case

d_i = period in which a trip departs/arrives from/in node i
d_i^p = arrival/departure time in/from node i in period p
f_k^{rs} = flow on path k originating from trip rs.
f(x) = cost function
m_{ij}^p = trips, exiting node i to node j, in period p
$m_{ij}(t)$ = trips, exiting node i to node j, at time t.
p = period number
p_i = period in which the traffic departs from node i
q = flow (veh/hr) (also intensity or volume)'
q_{rs} = trip rate from origin r to destination s
q_{rs}^d = trip rate between origin r and destination s departing in period d.
q_{rn}^d = trip rate between origin r and node n departing in period d.
$q_{rs,k}^d$ = trip rate between origin r and destination s departing in period d, and taking path k.
x_{ij} = flow on link ij
x_{ij}^p = flow on link ij in period p
$x_a(t)$ = flow on link a at time t
t_{ij} = travel time on link ij
$t_{ij}(x)$ = travel time function for link ij
$t_{ij}^p(x)$ = travel time function for link ij in period p
$u_a(t)$ = the inflow of link a at time t
$v_a(t)$ = the outflow of link a at time t
v = speed (km/hr)
v^{max} = maximum (free flow) speed
v_{rij}^{dp} = flow on arc ij in period p for the trips departing from r in departure period d.

Greek

α = parameter to adjust the angle of the linear part of the modified Smulders function
α^p = fraction of the link length covered in period p
α_{ri}^{dp} = zero-one variable indicating whether trips departing from origin zone r in departure period d reach node i in period p.
β = parameter to adjust the curve of the non-linear part of the modified Smulders function
γ_{ars}^{pd} = the fraction in which the trip r-s departing in period d contributes to the density on link a in period p
$\delta_{ij,k}^{rs}$ = indicator variable, equal to one if link ij is on path k from trip rs, zero otherwise.
ζ_{rs}^d = travel time for the trip r-s departing in period d.
ϕ_r^d = fraction of the trip rate departing from origin r in period d.
μ_{ij}^p = trips entering node j from node i in period p.
$\mu_{ij}(t)$ = trips entering node j from nide i at time t

ρ = density (veh/km)
ρ_a^p = density at link a in period p
$\rho_a^{p,\,old}$ = density at link a in period p, as a result from all the previous iterations
$\rho_a^{p,\,new}$ = density at link a in period p, as a result of the last iteration only
ρ_{ars}^{pd} = contribution to the density at link a in period p from trip rsd
ρ^{crit} = critical density (density at which the maximum flow is reached)
ρ^{max} = maximum density
σ_i = fraction of the link length covered of a link starting in node i
σ_j = fraction of the link length covered of a link ending in node j
τ_a^p. = travel time on a link expressed in number of periods (not necessarily an integer value)

APPENDIX F

DETAILED MAP INTERSECTION AMSTEL

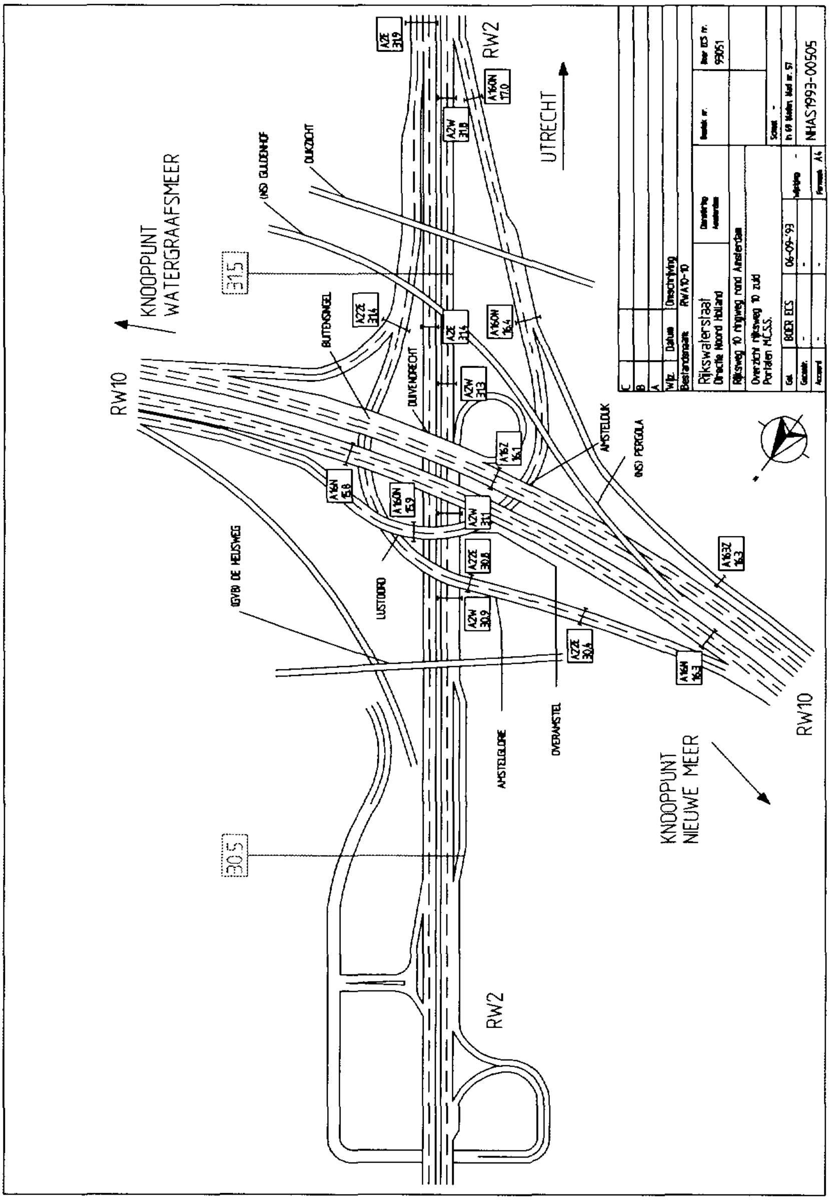


Figure 90: One of the in total 17 maps of the Amsterdam beltway, induction loops are indicated with road number, direction and kilometre position (e.g. A22E, 30.4).

APPENDIX F: DETAILED MAP INTERSECTION AMSTEL

INDEX

SUMMARY

A DYNAMIC TRAFFIC ASSIGNMENT MODEL
Theory & Applications

A recent solution to alleviate congestion in traffic systems is Dynamic Traffic Management (DTM).

> *"Dynamic Traffic Management employs technologies for real-time traffic monitoring, network wide management of traffic flows, and traffic-adaptive control to respond to changing traffic conditions while improving the efficiency, safety and travel conditions of a highway network"* .

Within traffic management systems, various instruments are available, and they can be divided in two basic types. *Road side based systems*, designed to manage and control traffic flow, e.g. ramp meters, and give road side information, e.g. changeable message signs, and *In-Vehicle systems* which use electronic equipment inside the car to provide information to the driver (navigation) or to control the behaviour of the car (speed and distance control). All these systems can work in a complementary way.

With the introduction of DTM, three different levels of decision making are involved:
- Firstly, decisions have to be made concerning long term planning of DTM (years). Questions about the feasibility of DTM have to be answered, and the location and impact of its instruments have to be studied, sometimes in combination with traditional traffic engineering, such as the construction of new roads.
- Secondly, decisions have to be made concerning medium term management of DTM (a few hours). When several DTM instruments are in use, questions have to be answered about the settings of certain ramp meters and changeable message signs. In the case of an accident advice about the impact of the accident and the measures to be taken is needed quickly. Rather than analyse and react to information, decisions have to be based on forecasts, making action prior to degradation of the traffic flows possible.
- Thirdly, decisions must be made concerning short term control of DTM (a few minutes). These decisions are usually made by computers without human interference. The control of a local ramp meter, or the alarm signal raised by a truck entering a tunnel that is too low for it, are two examples.

The three different levels of decision making are named respectively: Advanced Traffic Planning Systems (ATPS), Advanced Traffic Management Systems[1] (ATMS) and Advanced Traffic Control Systems (ATCS).

The kind of decisions required to introduce these systems are difficult to make. Most of the proposed systems are very expensive to run and maintain, introduce new technologies, and have an uncertain impact.
Using *traffic models* various strategies can be tested and fine tuned without having to disturb

1. Caution: ATMS is redefined here, and narrowed down to the actual management of traffic flow. Automatic traffic control systems have been excluded.

traffic, they avoid risk of liability when problems in a strategy are detected only after implementation, and they save the capital required to acquire and install traffic control hardware so that strategies, which may or may not work, can be field tested [80]. These traffic models should be able to simulate traffic realistically and be suitable for calculating the effects of DTM in networks, and for ATMS applications in particular, they must be able to do this within a reasonable time.

In the recent literature several models can be found that have the potential to be applied for DTM. These models can be divided into two categories: simulation models and assignment models. Simulation models are, in general, not well suited for network problems. They do not consider route choice and are designed for small networks, but they are well suited to study the local impact of dynamic traffic management measures, such as the local impact of ramp metering. For ATPS and ATMS models are required that can manage entire networks and consider route choice, and provide insight into rerouting and the network wide impact of ramp metering, i.e. coordinated ramp metering.
Models that consider route choice, are assignment models. These models are divided into *static* assignment models and *dynamic* assignment models. Static assignment models lack the ability to model traffic realistically. The absence of dynamics in these models makes them unsuitable to model congestion. Dynamic assignment models, however, show several promising features. These models simulate the evolution of traffic along the links of the network and can therefore model traffic more realistically. Several models have incorporated mechanisms to model DTM measures. Dynamic assignment models are divided into *heuristic* models and *mathematical* models:
The mathematical approach contributes several theoretically sound models. Proof of convergence and the existence of a unique solution are provided. A consequence of the strict mathematics is, however, that the amount of realism in these models is rather limited, and until now the literature has contained very few implementation or computational works, and testing has only been done on very small networks.
Several of the heuristic dynamic assignment models have the potential to meet the requirements for ATMS and ATPS. Some of the models described, however, lack the ability to calculate routes that are based on travel times while traversing the links, i.e. experienced route choice. Experienced route choice is very important for traffic predictions, because instantaneous route choice does not divert traffic optimally.

The *main research issues* addressed in this dissertation are:

- The development of a traffic model to simulate traffic flows in a network and to test DTM instruments on their ability to solve the congestion problem.
- The development of a traffic management system which can predict traffic flow in a freeway network and utilize DTM instruments

The newly developed model, called 3-Dimensional Assignment (3DAS), determines the time varying traffic conditions, flow and travel time, for every link in a network during a certain simulation period. Like most other dynamic assignment models is the simulation period divided into intervals of equal duration, that are referred to as periods. The link parameters may therefore be defined separately for each period. The network and the travel demand (OD-matrix) are presumed to be given.

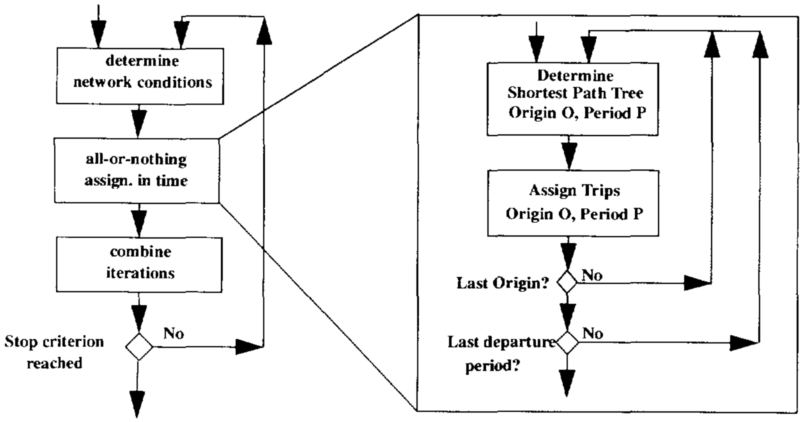
Two commonly used assumptions are made concerning route choice:
- all travellers are completely informed and take future congestion into account.
- all travellers choose their minimum cost path.

The theory of the model, as it is described in this dissertation, is primarily constructed as a user optimal model. Travellers choose their minimum cost path, and take *future* congestion into account. A second strategy is to choose the *instantaneous* shortest path. This path is based on travel times at the moment of departure or at the moment of a choice (at a junction). Instantaneous route choice does *not* take future congestion into account. A third strategy, to model a more realistic route choice, is to use *historic* information for route choice.
The user optimal approach is described as the main approach, instantaneous path finding is described as a simplification, and path finding based on historical information as an extra feature.

As discrete time units are used and for simplicity, the two following assumptions concerning the links, are made:
- traffic conditions are homogeneous along a link and constant for the duration of each time period.
- the variables of a link, flow (q), density (ρ) and speed (v), in a certain period p are subject to a fixed relationship described by:

$$q^p = \rho^p \cdot v^p$$

The model is based on the iteration scheme in the next figure, and is essentially the same as is used for *static* assignment models. The difference lies in the "all-or-nothing-assignment-in-time" module. In this module an extra iteration over the departure period is needed, and the shortest pathfinding and the assignment have to be performed *in time*.



The paths are defined using the travel time on a link for the period in which trips actually traverse the link, i.e. the trajectory a trip follows in time is calculated. The trips are assigned to the network, based on these trajectories. During assignment the contribution of a trip to the traffic load on a link in a certain period is determined by calculating the duration of presence on the link for that period.

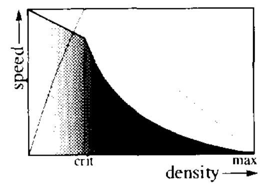
If one traveller is focused on than two situations can occur:

1. *Several links are traversed within one period.* In this case the traveller is only present on a link for a part of the period, and therefore should only be assigned for this part.
2. *One link is traversed in several periods.* The traveller is present on a link during multiple periods and should be assigned entirely for each individual period.

At the start of each iteration the travel times on the links are derived from the load of the previous iteration. The travel time is calculated for each link using a speed-density function. Speed-density functions are used instead of traditional traveltime-flow functions to model a decreasing flow in cases of congestion. Conservation of traffic and the continuity of flow is maintained. In cases of overflow, the overflow is assigned to the preceding link of the path in the same period. The process stops if the results of the last iteration are sufficiently close to the link flows of the previous iteration.

A speed-density function of the following form is used (Smulders [86]):

$$v(\rho) = \begin{cases} v^{max} \cdot \left(1 - \dfrac{\rho}{\rho^{max}}\right) & 0 < \rho < \rho^{crit} \\ \phi \cdot \left(\dfrac{1}{\rho} - \dfrac{1}{\rho^{max}}\right) & \rho^{crit} < \rho < \rho^{max} \end{cases}$$



Where v^{max} is the free flow speed, ρ^{crit} is the critical density and ρ^{max} is the maximum density. The maximum density represents a no motion traffic jam, and at the critical density the maximum flow (capacity) is reached. The parameter ϕ is chosen to make the function continuous at ρ^{crit} ($\phi = v^{max} \rho^{max}$).

Four different pathfinding methods and two different assignment methods were developed and tested. A method for traffic jam building and a method for merging traffic were added to the model to give a realistic representation of traffic. Several extra features are appended to the model for DTM applications, such as ramp metering and variable message signs.

The consequences of discretization of time and space were investigated. The most obvious conclusion concerning link length is that longer links are less accurate. Areas with an expected high level of congestion should be modelled with short links. It is important that the definition of the network and the length of the periods should be adapted to fluctuations in demand. When demand fluctuates highly, a more detailed network and shorter periods are advisable. An important advantage of the 3DAS model, however, is that detail does not get lost for downstream links, though underestimation of travel time will occur on long links. This underestimation increases with higher densities.

It is argued, using induction loop data from Amsterdam, that there is no generally applicable speed-density function for a network. The relationship should at least be tailored to the location. A merging section shows, for example, a different function than an exit ramp. For ATMS applications, the function also needs to be adapted to current conditions, such as the influences of weather, time of day and accidents. Using the Amsterdam data it is demonstrated, that a modified Smulders speed-density function gives a good fit with the data, and is easily adapted to different situations.

To test and calibrate several parameters of the 3DAS model, it was compared to the microscopic simulation model FOSIM. A bottleneck and an on-ramp were tested. The 3DAS model was shown to be capable of reproducing these basic traffic situations. Although the results were not reproduced in detail, the effect a bottleneck and an on-ramp would have *in a network* were reproduced.

The 3DAS model was developed and implemented to operate in ATPS and ATMS. Little extra effort is required to use the 3DAS model for ATPS applications. Traditional *static* OD-matrices can be used, and made dynamic using departure time patterns. Several extra modules are required for ATMS applications, to provide the 3DAS model with input and to make the model adaptive to prevailing conditions. One of the modules, the OD-split matrix module, is a research project by itself and is described elsewhere. The departure time patterns module uses historical data to predict future departure time patterns. Empirical research has shown that departure patterns are fairly consistent each day. This allows historical patterns to be used as a basis for prediction. Good results were achieved with the on-ramps in Amsterdam, by fitting the historical pattern to the current measured pattern. Using the speed-density functions module, the model can be made adaptive to prevailing conditions. The speed-density functions are easily adapted using data from induction loops, and it is not necessary to be aware of all kind of weather circumstances or incidents.
A fast computing time is essential for traffic predictions. A prediction is clearly useless when the computing time exceeds the required prediction time. When conventional computers are not fast enough, dedicated computers can solve computational problems [29].

Using 3DAS for ATPS was tested in Washington and for ATMS in Amsterdam. The ATPS case showed a number of clear advantages for using 3DAS instead of the traditional static assignment model. The results give more detailed information about the occurrences of traffic jams, and the location or the cause of congestion can be identified more precisely. To alleviate congestion, DTM measures can be simulated, and all kinds of evaluations are possible. Influences on travel time and traffic jam length, effects of ramp metering and rerouting can be investigated. Dynamic assignment also has the advantage that all kind of *temporary* disturbances, such as accidents or road works, can be simulated, and the duration of delays can be derived. The study also showed that 3DAS can be used successfully with larger networks. Data requirements, however, are much more stringent, because the level of detail required for 3DAS is higher and the data must support this level of detail. The accuracy of time variance depends directly on the accuracy of the time variance of the OD-matrix. As the demand was represented by a rather smooth pattern no discretization problems were encountered.
It is essential to find a good system of organizing and maintaining the large amount of data that 3DAS requires and produces. In the beginning this costs a great deal of effort, but with increasing experience this disadvantage will most probably disappear. Visualization of the results is very important, because with dynamic assignment the results are flows *in time*, and the best way to analyse the flows is as a *movie*. To do this a graphically powerful workstation must be used. This is one of the main reasons graphic workstations were used for this research.

Using 3DAS for ATMS was tested in Amsterdam. One minute data was measured at 141

screenlines for a duration of three weeks in April 1994. The model was tested on two different days. One day represented a "normal" morning rush hour, the other day represented a morning rush hour with an accident which blocked all lanes. A third data set representing the average of the measurements of five days was used to compare the results. This data set was named: 'the historical pattern'.
It can be concluded that for "normal" days, the 3DAS model can reproduce the traffic situation, but that the historical pattern is at least as good. The day with the accident, however, showed that the 3DAS model gave a better reproduction than the historical pattern. It can be concluded from this observation that a dynamic assignment model is not required to make traffic predictions for normal days, however, the use of a dynamic assignment model is useful for days with awkward conditions.
The case also illustrated that correct representation of the network and the speed-density functions is very important. The wrong number of lanes at one section can have network wide consequences. Using the available graphic software these errors can be located easily and corrected. The size of the Amsterdam network caused no problem for the software developed and conventional hardware which could make a prediction within several minutes. The dedicated computer gave a substantial improvement in calculation time.

A new traffic model to simulate traffic on a network and to test DTM instruments on their ability to solve the congestion problem is now available, the dynamic traffic assignment model **3DAS** (3-Dimensional ASsignment). The 3DAS model is capable of providing predictions for Advanced Traffic Planning Systems (ATPS) and Advanced Traffic Management Systems (ATMS) and can simulate large networks within a sufficiently short time.

SAMENVATTING (SUMMARY IN DUTCH)

EEN DYNAMISCH TOEDELINGSMODEL
Theorie & Toepassingen

Een recente benaderingswijze voor het aanpakken van het congestie probleem in verkeersnetwerken is Dynamic Traffic Management (DTM), of dynamische verkeersbeheersing.

> *"Dynamic Traffic Management omvat het management en de controle van verkeersstromen aan de hand van actuele verkeersgegevens, teneinde de efficiëntie, de veiligheid en de reis omstandigheden op een verkeersnetwerk te verbeteren".*

Verscheidene instrumenten zijn beschikbaar voor het ten uitvoer brengen van dynamisch verkeersmanagement. Men kan onderscheid maken tussen *'road side based systems'* en *'in-vehicle systems'*. Road-side based systems zijn ontworpen voor management en beheersing van de verkeersstromen, zoals toeritdoseerinstallaties, en om informatie aan weggebruikers te geven, zoals routeinformatieborden. In-vehicle systemen geven actuele informatie aan de automobilist met behulp van elektronische apparatuur in de auto, zoals dynamische routegeleiding. Er zijn ook systemen die het rijgedrag van de auto beïnvloeden , zoals snelheids- en afstandsregeling. Al deze systemen moeten complementair kunnen werken.

Voor een effectieve toepassing van DTM zijn beslissingen nodig op drie verschillende niveaus, die betrekking hebben op abstractie en tijdschaal. Op basis van deze kenmerken wordt DTM onderverdeeld in de volgende drie systeem categorieën:

- Advanced Traffic Planning Systems (ATPS), voor lange termijn planning (jaren). Vragen over de haalbaarheid van DTM moeten beantwoord worden, en de lokatie en de invloed van de toegepaste instrumenten moet bestudeerd worden. Bovendien zal het nodig zijn te combineren met traditionele beslissingen, zoals infrastructurele uitbreidingen en aanpassingen.
- Advanced Traffic Management Systems (ATMS), voor middellange termijn vraagstukken (enkele uren). Als verscheidene DTM instrumenten operationeel zijn, moeten vragen beantwoord worden over de locatie en de instellingen van deze instrumenten. In geval van een incident is er snel advies nodig over het effect van het incident op de verkeersafwikkeling en de te nemen maatregelen. Beschikbare informatie moet geanalyseerd worden en de nodige acties moeten uitgevoerd worden. Bovendien zijn voorspellingen essentieel, zodat anticiperende acties genomen kunnen worden *voordat* de verkeersprestatie van het wegennet ernstig verslechtert.
- Advanced Traffic Control Systems (ATCS) voor de korte termijn gevolgen van DTM (enkele minuten). Dit soort beslissingen worden in de regel genomen door computers, zonder menselijke tussenkomst. De besturing van een toeritdoseerlicht, of het alarm dat afgaat als een vrachtwagen te hoog is voor een tunnel, zijn voorbeelden van dit soort systemen.

De beslissingen die benodigd zijn voor elk van deze drie DTM categorieën zijn niet triviaal. Bovendien zijn de meeste systemen duur om operationeel te maken en te houden, gebaseerd op nieuwe technologieën en is de effectiviteit nog onzeker.
Om inzicht te krijgen in de invloed van allerlei maatregelen wordt in de verkeerskunde een

verscheidenheid aan modellen voor het berekenen van verkeersstromen gebruikt. Met behulp van deze modellen is het mogelijk verschillende DTM strategieën te testen zonder het verkeer te verstoren. Deze verkeersmodellen moeten uiteraard wel in staat zijn het verkeer op een realistische manier te modelleren, en geschikt zijn om het effect van DTM instrumenten op netwerk niveau te simuleren. Voor de middellange en de korte termijn moeten deze modellen bovendien voldoende snel zijn om in de realiteit toepasbaar te zijn.

In de recente literatuur kan men verscheidene verkeersmodellen vinden die de potentie hebben om voor DTM te worden toegepast. Deze modellen kunnen worden ingedeeld in twee categorieën: Simulatie modellen en Toedelingsmodellen. Simulatie modellen beschouwen meestal geen routekeuze en zijn in het algemeen minder geschikt voor netwerk problemen. Ze zijn het meest geschikt voor kleine netwerken, en om de lokale invloed van DTM te bestuderen, zoals het effect van een toeritdoseerinstallatie op de lokale verkeersstromen. Voor ATPS en ATMS is er een behoefte aan modellen die wel routekeuze beschouwen en de invloed van DTM instrumenten op netwerkniveau kunnen doorrekenen.
Verkeersmodellen die wel routekeuze beschouwen zijn toedelingsmodellen. Deze modellen kunnen worden ingedeeld in *statische* toedelingsmodellen en *dynamische* toedelingsmodellen. Statische toedelingsmodellen zijn niet goed in staat congestie te modelleren omdat geen rekening wordt gehouden met veranderingen in de tijd (dynamiek). Dynamische toedelingsmodellen zijn beter geschikt, want deze modellen modelleren het verloop van het verkeer wel in de tijd en beschrijven verkeerstromen derhalve op een meer realistische wijze. Verscheidene van deze modellen beschikken over mechanismen om DTM instrumenten te modelleren. Dynamische toedelingsmodellen kunnen ingedeeld worden in *heuristische* modellen en *mathematische* modellen. Mathematische modellen zijn strikt geformuleerde modellen waarvan de convergentie naar een unieke oplossing is bewezen. De realiteit van het verkeersproces is in deze modellen echter vrij beperkt, en tot nu toe bevat de literatuur slecht weinig implementatie voorbeelden, en zijn tests slechts uitgevoerd op kleine netwerken. Voor de heuristische aanpak komt het mathematische bewijs van convergentie niet op de eerste plaats, maar de realiteit van het verkeersproces, dat deze modellen beter geschikt maakt voor DTM toepassingen. Slecht enkele van deze modellen zijn echter in staat om routes te bepalen die gebaseerd zijn op de reistijden van het moment waarop de links gepasseerd worden (*experienced* routekeuze).

De belangrijkste research doelstellingen van deze dissertatie zijn:
- Het ontwikkelen van een verkeersmodel waarmee verkeersstromen gemodelleerd kunnen worden, en DTM instrumenten getoetst kunnen worden.
- Het ontwikkelen van een verkeersbeheersing systeem waarmee verkeersstromen voorspeld kunnen worden en DTM instrumenten toegepast kunnen worden.

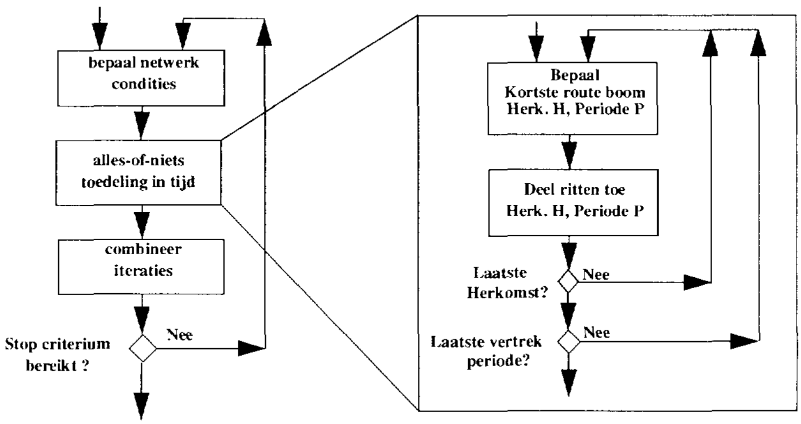
Om aan de bovenstaande doelstellingen te voldoen is het dynamisch toedelingsmodel 3DAS (3-Dimensional Assignment) ontwikkeld. Hierbij is, zoals in de meeste dynamische toedelingsmodellen, de simulatie periode verdeeld in intervallen van gelijke duur die perioden genoemd worden. Het netwerk en het aanbod aan verkeer (HB-matrix) worden als bekend verondersteld, en zijn voor elke periode apart gedefinieerd. Het model berekent voor elke link en elke periode de in de tijd variërende grootheden, zoals dichtheid, intensiteit en reistijd.

De volgende twee veronderstellingen over routekeuze zijn aangenomen:
- alle reizigers zijn volledig geïnformeerd en op de hoogte van toekomstige congestie,
- alle reizigers kiezen de kortste route.

De theorie van het model is dus in eerste instantie gebaseerd op het evenwichtsmodel. Alle reizigers kiezen de kortste route en houden rekening met toekomstige congestie (*experienced* routekeuze). Een tweede strategie baseert de kortste route op basis van de omstandigheden die gelden op het moment van vertrek. Deze strategie wordt *instantaneous* routekeuze genoemd, en houdt geen rekening met toekomstige congestie. Een derde strategie is een meer realistische routekeuze die gebaseerd op historische reistijden.
Het model is beschreven als een evenwichtsmodel, het instantaneous routekeuze model is beschreven als een simplificatie en het historische routekeuze model als een extra toevoeging.

Doordat discrete tijdseenheden worden gebruikt, en voor de eenvoud, zijn de volgende twee veronderstellingen aangenomen:
- verkeers condities op een link zijn homogeen over de lengte ervan en constant voor de duur van de periode,
- de variabelen van een link, intensiteit (q), dichtheid (ρ), en snelheid (v), zijn binnen een bepaalde periode p gerelateerd door middel van $q^p = \rho^p \cdot v^p$

Het model is gebaseerd op het onderstaand iteratie schema, dat in principe gelijk is aan het iteratie schema van een statisch toedelingsmodel. De module "alles-of-niets-toedeling-in-tijd"is echter uitgebreid met een extra iteratie over de vertrekperiode en voor het berekenen van de kortste route en het toedelen van de ritten moet rekening gehouden worden met het tijdstip.



Figuur 91: Iteratie schema

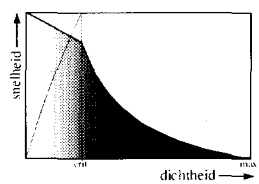
De routes worden berekend op basis van de reistijden van de links in de periodes waarin de link ook daadwerkelijk gepasseerd wordt. Om de juiste periode te bepalen, wordt een trajectorie berekend. De ritten worden aan het netwerk toegedeeld aan de hand van deze trajectorieën. De bijdrage van een rit aan de belasting op een link wordt bepaald door de *aanwezigheidsduur* te berekenen op de link in die periode. Als we ons richten op één reiziger dan kunnen er in principe twee situaties ontstaan:

1. *Meerdere links worden gepasseerd binnen één periode.* In dit geval is de reiziger maar voor een gedeelte van de periode aanwezig op een link, en wordt dus ook maar voor een gedeelte toegedeeld aan een link.
2. *Meerdere periodes zijn nodig om één link te passeren.* De reiziger is aanwezig op dezelfde link gedurende meerdere perioden en wordt derhalve voor elke periode in zijn geheel toegedeeld.

Aan het begin van iedere iteratie worden de reistijden voor elke link berekend op basis van de belasting van een link, gebaseerd op de situatie in de vorige iteratie. De relatie tussen reistijd en belasting wordt bepaald door een snelheid-dichtheid functie. Deze functie wordt gebruikt in plaats van de traditionele reistijd-intensiteit functie, omdat de relatie tussen reistijd en intensiteit in werkelijkheid een dualiteit vertoont. Bij één intensiteit kan zowel een *free-flow* reistijd als een *file* reistijd optreden.
In het geval dat de maximale dichtheid wordt overschreden op een link, wordt het teveel toegedeeld aan de voorgaande link op de route in dezelfde periode. Het stop criterium van het iteratieproces wordt bereikt als er vrijwel geen verschil meer is tussen de resultaten van twee opeenvolgende iteraties.

De snelheid-dichtheid functie is gebaseerd op de volgende functie (Smulders [86]):

$$v(\rho) = \begin{cases} v^{max} \cdot \left(1 - \dfrac{\rho}{\rho^{max}}\right) & 0 < \rho < \rho^{crit} \\ \phi \cdot \left(\dfrac{1}{\rho} - \dfrac{1}{\rho^{max}}\right) & \rho^{crit} < \rho < \rho^{max} \end{cases}$$



waarin v^{max} de maximale snelheid weergeeft, ρ^{crit} de kritieke dichtheid en ρ^{max} de maximale dichtheid. Bij maximale dichtheid staat het verkeer stil, bij de kritieke dichtheid wordt de maximale intensiteit (capaciteit) bereikt. De parameter ϕ wordt zodanig gekozen zodat de functie continu is in ρ^{crit} ($\phi = v^{max}\rho^{max}$).

Voor het routezoeken zijn vier verschillende methoden en voor het toedelen zijn twee verschillende methoden ontwikkeld en onderzocht. Het 3DAS model is uitgebreid met een methode voor het op- en afbouwen van files en voor weefgedrag bij opritten, en voor het simuleren van ATMS instrumenten, zoals voor toeritdosering en voor dynamische route informatiepanelen.

De consequenties van de discretisatie in tijd en ruimte zijn onderzocht. Als meest voor de hand liggende conclusie kan gesteld worden dat kortere links nauwkeuriger zijn dan lange links. Belangrijk daarbij is echter dat de lengte van de links en de duur van de periodes vooral aangepast moeten worden aan de mate van fluctuatie van het aanbod. Een belangrijk voordeel van het model is echter dat bij een lange link de fluctuatie niet verloren gaat voor achterliggende links. Alleen een onderschatting van de reistijd kan optreden op de te lange link.

Met behulp van gemeten snelheden en intensiteiten op de ringweg van Amsterdam is aangetoond dat het niet mogelijk is een algemeen toepasbare snelheid-dichtheid functie toe te passen. Voor ATPS toepassingen moet de functie aangepast worden aan de aard van de link. Een weefsectie vertoont een heel andere relatie dan bijvoorbeeld een afrit. Voor ATMS toepassingen moet de functie bovendien aangepast worden aan de op dat moment geldende condities, zoals weersomstandigheden, tijdstip van de dag en aan ongevallen. Aan de hand van de verkeersmetingen is aangetoond dat de gebruikte snelheid-dichtheid functie goed is te fitten op de data en eenvoudig aangepast kan worden aan verschillende situaties.

Om het 3DAS model te testen en de parameters te kalibreren, is het model voor twee situaties vergeleken met het microscopisch simulatie model FOSIM, een bottleneck (een wegversmalling van drie naar twee stroken) en een weg met een oprit zijn getest. De resultaten werden niet tot in het detail gereproduceerd maar het effect van een wegversmalling of een oprit in een netwerk werd goed gereproduceerd.

Toepassingen van het 3DAS model moeten met name gezocht worden in ATPS en ATMS. Voor ATPS toepassingen is het model vrijwel op gelijke wijze toe te passen als statische modellen. Traditionele, statische HB-matrices kunnen dynamisch gemaakt worden met behulp van vertrektijd functies. Voor ATMS toepassingen daarentegen zijn extra modules nodig om de benodigde invoer te verwerken en het model gevoelig te maken voor de op dat moment heersende omstandigheden. Eén van deze extra modules, het schatten van de HB-matrix, is een apart onderzoek en niet beschreven in dit proefschrift. De extra module voor het bepalen van de vertrektijd functies maakt gebruik van historische data voor het schatten van toekomstige vertrektijd functies. De extra module voor de snelheid-dichtheid functies zorgt ervoor dat het model reageert op veranderende omstandigheden, zoals weersomstandigheden of ongevallen. De functies zijn eenvoudig aan te passen met behulp van inductielusgegevens. Voor het maken van korte termijn voorspellingen is een snelle rekentijd essentieel. In het geval dat PC's of werkstations niet snel genoeg zijn, is er de beschikking over een special purpose computer om de rekentijd te verkorten [29].

Het 3DAS model is getest in Washington voor ATPS toepassingen en in Amsterdam voor ATMS toepassingen. De ATPS case toonde een aantal duidelijke voordelen aan van 3DAS boven de traditionele, statische toedelingsmodellen. Men beschikt over meer gedetailleerde informatie over het tijdstip van optreden en de lokatie van files, waarmee meer inzicht ontstaat in de oorzaak van een file. Met het model kunnen verschillende ATMS scenario's geëvalueerd worden. Tijdelijke verstoringen, zoals ongevallen en wegwerkzaamheden, kunnen gesimuleerd worden waarmee inzicht wordt verkregen in de lengte van file's en de duur van een verplaatsing. De afmetingen van het Washington netwerk (2098 links) gaven geen problemen voor het model. De eisen die gesteld moeten worden aan de data zijn echter hoger dan bij traditionele statische toedelingsmodellen. Omdat de uitkomsten een hoger detailniveau hebben moet de invoer ook een hoger detail niveau hebben. De nauwkeurigheid waarmee de periodelengte wordt gespecificeerd is direct afhankelijk van de definitie van de periodelengte in de HB-matrix. Doordat de fluctuatie in de HB-matrix geen extreme pieken vertoonde waren er geen problemen met discretisatie verschijnselen. De hoeveelheid data die het 3DAS model nodig heeft en produceert, stelt hoge eisen aan de organisatie daarvan. Het is erg belangrijk een goed management systeem voor deze hoeveelheden data te vinden. Naarmate men meer ervaring krijgt met 3DAS wordt dit probleem minder.
De visualisatie van de resultaten bleek erg belangrijk te zijn. Omdat de resultaten van een dynamisch model tijdseries zijn, moeten de resultaten als een soort film gepresenteerd worden. Dit was een van de belangrijkste redenen dat grafische werkstations zijn gekozen, zowel voor de ontwikkeling als voor de toepassing van 3DAS.

In Amsterdam is het 3DAS model getest voor ATMS toepassingen. Gedurende drie weken is op 141 lokaties in het netwerk elke minuut data verzameld. Het 3DAS model is vervolgens getest op twee verschillende dagen. Eén dag vertegenwoordigde een normale ochtendspits, de andere dag een situatie met een ongeval, dat de gehele rijbaan blokkeerde. Een derde dataset werd samengesteld door het gemiddelde van de metingen van vijf verschillende dagen te nemen. Deze dataset wordt aangeduid als het 'historische patroon'.
Voor de normale dag werden de verkeersstromen goed gereproduceerd door 3DAS, het historische patroon bleek echter nog beter te zijn. Voor de dag met het ongeval bleek het 3DAS model een betere reproduktie te geven dan het historische patroon. Hieruit kan geconcludeerd worden dat een dynamisch toedelingsmodel niet nodig is om voorspellingen te maken voor normale dagen, maar wel voor dagen met afwijkende omstandigheden.

Een correcte representatie van het netwerk en de snelheid-dichtheid functies bleek erg belangrijk. Het verkeerde aantal rijstroken kan grote gevolgen hebben en het patroon in het hele netwerk beïnvloeden. Met behulp van de uitgebreide visualisatie mogelijkheden van de software werden deze fouten meestal snel gelokaliseerd en verholpen. De afmetingen van het netwerk geven geen problemen voor de rekentijd. Op een werkstation kan het netwerk doorgerekend worden binnen enkele minuten, met de special purpose computer nog aanzienlijk sneller.

Met het dynamisch toedelingsmodel 3DAS is een nieuw verkeersmodel beschikbaar gekomen, waarmee de verkeersstromen in netwerken gemodelleerd kunnen worden en het effect van DTM instrumenten onderzocht kan worden. Het model is in staat om voorspellingen te genereren voor Dynamic Traffic Management en kan grote netwerken doorrekenen in een voldoende korte tijd.

ABOUT THE AUTHOR

Erik de Romph was born at the 13th of February 1965 in Velsen. In 1984 he received a VWO diploma from the "scholengemeenschap Hendrik van der Vlist" in Utrecht. Erik graduated in Computer Science from the Delft University of Technology in the Netherlands in 1989. During this study he designed and implemented a database system at the "Antoni van Leeuwenhoek" Laboratory in Amsterdam. He spent the last eighth months of this study period at the traffic consultancy BGC in Deventer, working for his masters thesis. The MSc. thesis describes the design, development and implementation of a Decision Support System for the estimation of multi-modal Origin-Destination matrices for traffic forecasting studies.
After travelling around the world, Erik returned to Delft and joined the Traffic Modelling section of the Department of Infrastructure at the faculty of Civil Engineering as a research fellow in 1990, where he started the research on Advanced Traffic Management Systems and dynamic traffic assignment models. During this research he has presented various papers in the Netherlands, and at several international conferences, in Germany, Italy and the USA, with papers published in conference proceedings and in the Transportation Research Records. During his period at the TU Delft he was responsible for teaching a class in Computer Science for Civil Engineers, and has assisted students with practical work. In 1993 Erik spent five months at The Virginia Polytechnic and State University (Virginia Tech) in Blacksburg, Virginia, USA; here he took some courses on optimization and visualisation and carries out a case study on dynamic traffic assignment in the Washington Metropolitan Area.