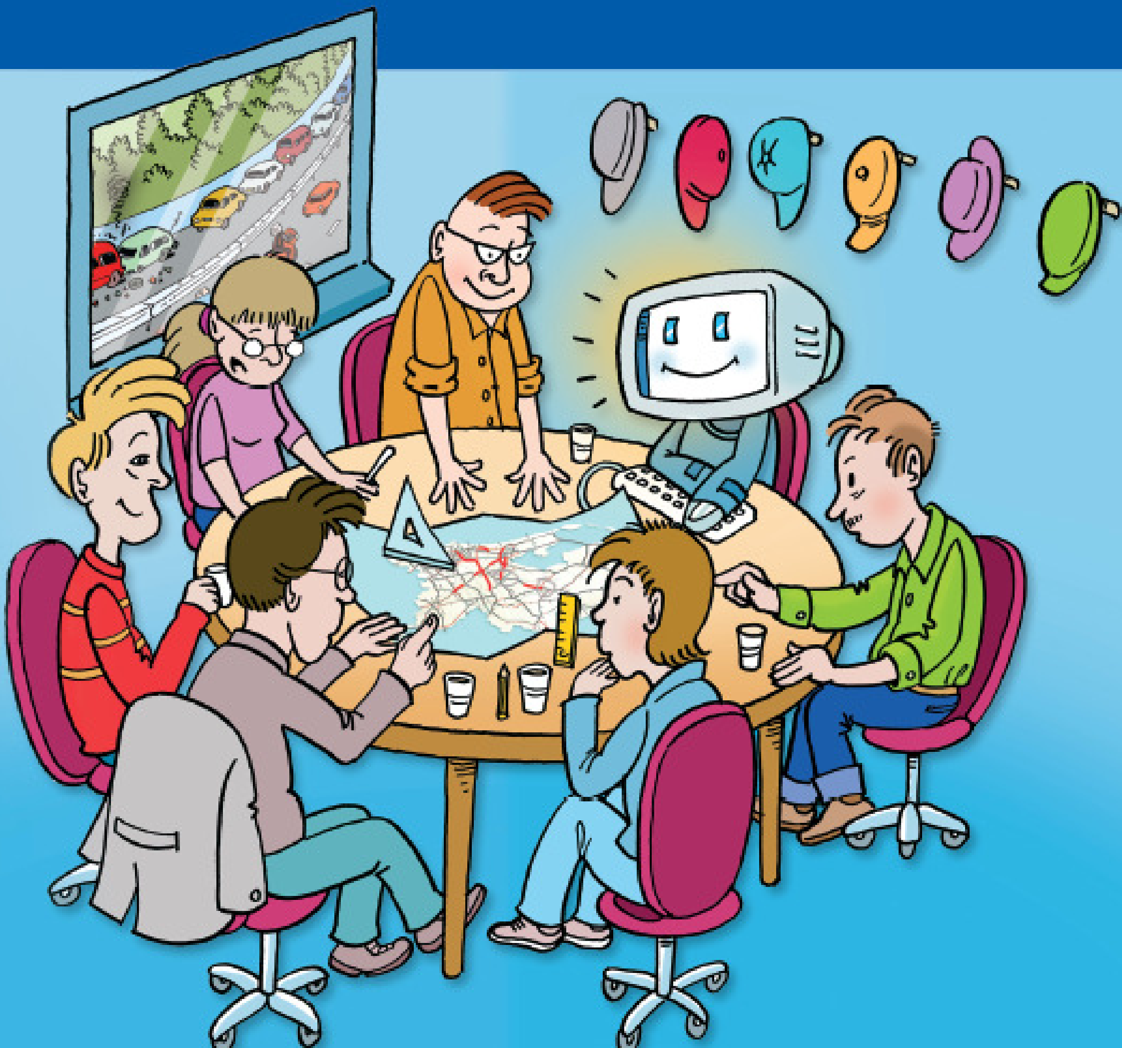
Designing Robust Road Networks

A general design method applied to the Netherlands



Maaike Snelder

Designing Robust Road Networks A general design method applied to the Netherlands | Maaike Snelder

Designing Robust Road Networks
A general design method applied to the Netherlands

Maaike Snelder

This thesis is supported by the research programmes Transumo and Next Generation Infrastructures, the Netherlands Orgnaisation of Applied Scientific Research TNO, the Delft University of Technology and the Netherlands Research School for Transport, Infrastructure and Logistics TRAIL

Cover illustration: petervandorst.nl

Designing Robust Road Networks
A general design method applied to the Netherlands

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties
in het openbaar te verdedigen op woensdag 1 december 2010 om 10.00 uur
door Maaike SNELDER
doctorandus in de econometrie
geboren te Rotterdam.

Dit proefschrift is goedgekeurd door de promotoren:
Prof. dr. H.J. van Zuylen
Prof. ir. L.H. Immers

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. H.J. van Zuylen,	Technische Universiteit Delft, promotor
Prof. ir. L.H. Immers,	Technische Universiteit Delft, promotor
Prof. dr. M.G.H. Bell,	Imperial College London
Prof. dr. M. Florian,	Université de Montréal
Prof. dr. ir. L.A. Tavasszy,	Technische Universiteit Delft
Prof. dr. A.P.M. Wagelmans,	Erasmus Universiteit
Dr. K. Vuorenmaa Berdica,	Faveo Projektledning AB

TNO
Schoemakerstraat 97
NL-2628 VK Delft
T +31 15 269 69 69
F +31 15 261 24 03
E infodesk@tno.nl

tno.nl

ISBN nummer: 978-90-5986-370-5

Keywords: road network design, robustness, spare capacity, optimization method.

Printed in the Netherlands

Een rit door Nederland

Alleen in de auto,
Files op straat,
Eén grote ellende,
Gestrest, ik kom te laat.

Voor me auto’s,
Achter me auto’s,
Naast me auto’s,
Of kijk ik niet goed?

Groene weilanden,
Grazende koeien,
Door zon verlichte wolken,
En een bruisende stad.

Voor me is thuis,
Achter me ervaring,
Naast me natuur,
De tijd verstrijkt langzaam uur voor uur,

Ik geniet.

A trip through the Netherlands

Alone in my car,
Congestion everywhere,
One big disaster,
Stressed, I’ll never make it there.

Cars in front of me,
Cars behind me,
Cars next to me,
Or do I see it wrong?

Green meadows,
Grazing cows,
Sunshine through the clouds,
And a sparkling town.

Home is in front of me,
Experience behind me,
Nature is next to me,
Time passes slowly,

I am free.

- Maaike Snelder

Preface

Here it is: my thesis! I have been working on this thesis for about five years. In this period I have learned a lot about robustness, road networks and doing research and consultancy work in general. Moreover, it has been a very nice period, which I will never forget. In this period I was supported, encouraged and challenged by many people and organizations. Although I prefer to thank you in person, this preface is an excellent opportunity to express my thanks.

First of all, I would like to thank my promotors Henk van Zuylen and Ben Immers. You were the ones who believed in me from the start. You asked me if I wanted to do a PhD on robust road network design after having seen my master thesis and after having seen me at work at TNO for one year. I was honored that you asked me to do this and I didn't have to think very long before I said yes. It took some time to arrange everything, but we ended up with a construction in which I had a 36 hour contract at TNO and an 8 hour contract at the Delft University. It was not easy to spend the appropriate amount of time on my thesis, because of the many projects that I did at TNO which all had a deadline way before the deadline of my thesis. You watched me and this process from a distance but always gave me the confidence that I/we would succeed. I think you knew more than I did. I think you knew from the beginning that my work at TNO would be an imported input for my thesis. So, it is safe to say that you guided me in the right direction by sharing your knowledge on robustness with me, by making sure that I was surrounded by the right people and the right projects and by making me believe in myself. Especially, at the end you gave me a lot of comments that lifted my thesis from practical applications to a scientific level. I know that this was the point that worried you the most. Hopefully, you are pleased with the result.

In your footsteps there is a third professor and my direct manager for many years at TNO, Lori Tavasszy, who has been of great importance and who enabled me to combine my work at TNO and the TU. Lori, I think that we form a great team. You know how to organize everything in such a way that it works. You were the one with the fantastic idea to give me a

44 hour contract. You were the one who brought in the economic perspective of robustness. You were the one who challenged me to make examples that showed why robustness is an important issue and you were the one who started walking with those examples. Thanks to all your efforts, we were able to do a lot of projects about the importance of robustness and the evaluation of robustness for different organizations. You always asked me if it was possible to go one step further, to do something that hadn't been done before. I think, or at least hope, that I managed to do that and that is what makes us a great team.

Besides Henk, Ben and Lori, I would like to thank my other committee members for their willingness to participate in my committee, for their willingness to travel to Delft and their useful remarks. I really appreciated all your efforts.

At TNO I worked with a lot of colleagues who made, that I could go to work with a smile on my face every day. I will not mention everybody by name, because that brings along a great risk that I will forget somebody. When you read this, I hope that you know that this is about you. We had a lot of nice conversations in the open spaces of our building about the content, about our personal life's and quite often simply about nothing. That was just what I needed, so thank you very much for that. I would like to thank the 'Indy team' for helping me out with some practical issues and for their efforts in calibrating the Amsterdam network. There are three master students that I supervised whom I would like to thank: Michiel Muller, Silvie Ketelaars and Aleksander Jacimovic. You did a great job in validating Indy for incident situations, in analyzing the effects of unbundling of traffic flows and in making a link between complex network theory and robustness of road networks. I made reference to your work in this thesis. Furthermore, I would like to thank Rob Kooij for his help with complex network theory and robustness of telecommunication networks.

I would also like to express my thanks to two ex-colleagues of TNO: Jeroen Schrijver and Ramon Landman. You were both great persons to work with and we had a lot of fun during work time as well, but there is more to it.

Jeroen, you were the one who taught me almost everything I know about traffic and transport and about carrying out and leading projects as well. When I came to TNO as a 'Duale Student' you were my supervisor. I was an econometrician who only knew things about the road network from my own experience as a car driver. You taught me what the difference was between 'banen' and 'stroken' and you taught me about generation, distribution, substitution and route choice effects. Furthermore, together with Hans Meeuwissen and Albert Wagelmans, you supervised my master thesis project that became quite a success years later. The redesign of the Dutch road network, which was the starting point of my PhD research, made it to almost every news paper and even the museum that you visited weeks or months before I did myself. I hope you know that this was your success as well. Last but not least, I would like to thank you, Ben, Bart and Lori for the teamwork in many projects that we did about robustness.

Ramon, you were a very special colleague first at TNO and later on at the TU and we became friends. We had many interesting, sometimes professional, but in most cases personal conversations of which I of course hope that we will continue having them. You helped me quite a lot in putting things in perspective which is of course needed to write a dissertation. Or, in other words, you taught me something that the Buddhist would express as follows: "Meester, hoe kan ik een volmaakt schilderij maken?", "Wees volmaakt en schilder dan gewoon!" Ramon, thank you for being you and thank you for accepting my invitation to be my paranymph.

This automatically brings me to my colleagues at the Delft University. First of all, I have to apologize for not being near you as often as you deserved. You organized a lot of great being-togethers ranging from nice trips, to table tennis competitions and PhD meetings. I hope you know that the fact that I hardly ever joined these fantastic initiatives is not personal, but just a result of scarcity in time and the fact that we as human beings can just not be at two places at the same. The times that I was at the university, I mostly spent with my roommates. You made my time at the university nice and you also helped with all kinds of practical daily problems. Victor, it was very nice to travel with you to conferences and to work closely together. There wasn't a question that you could not answer. You helped me enormously. An obvious example of this, are the numerous times that you had to explain the shock wave theory to me. You did a lot research into the effects of incidents that was complementary to my work. People who are familiar with your thesis might recognize a subtle reference to your work on my cover. Finally, I would like to thank Minwei who was my supervisor at the TU in the beginning and later on a colleague at TNO. You helped me in many ways, but I was especially honored that you asked me to be your paranymph. This was a nice way of experiencing what it is to be on that side of the room during a PhD defence and was in that sense a nice preparation for my own defence. Furthermore, the dinner afterwards brought me in contact with Professor Michael Florian, which appeared to be important for my own thesis.

In June and July 2009, I visited Michael Florian and his colleague Michael Mahut in Canada. I had an excellent time and you were fantastic hosts. I learned quite a lot about dynamic traffic assignments and I learned quite a lot from being abroad. In Canada I had a lot of time to work on my PhD research. So, a relatively large part of the work was done over there. Furthermore, I had the privilege to extend my visit to Canada with a ten days stay in a very nice resort in Los Angeles where I wrote a large part of my thesis with an ocean view. It turned out that I needed a view on big waters. A year later, I wrote the last part of my thesis in Lisbon/Estoril. Life isn't that bad at all! The people from Montreal were not the only ones from abroad that were important for my research. In Leuven there are some excellent researchers (Chris, Ruben, Roderic and Isaak) who did work that was important for my research. I was happy to use the Link Transmission Model in Indy and the Marginal Incident Computation model that were developed by Isaak Yperman and Ruben Corthout. Finally, I would like to thank Warren Walker for correcting the English in my dissertation.

In this word of thanks, a lot of organizations and initiatives absolutely deserve a place. First of all I would like to thank TNO, the Delft University of Technology and the TRAIL Research School for enabling and allowing me to do a PhD research. I am proud to be a member of these organizations. Furthermore, my work was supported by Next Generation Infrastructures, TRANSUMO-ATMA and TRANSUMO-NiVeS. These kinds of research programs make it possible that PhD-projects are being carried out, which is of course fantastic. Furthermore, in random order, I am very grateful to the ANWB, the Ministry of Transport, Public Works and Water Management, the Advisory Council for Transport, Public Works and Water Management, the Directorate-General for Public Works and Water Management, the stadsregio Amsterdam and Bart Egeter Advies for assigning projects about robustness to TNO and, more important, for taking this topic serious. From my point of view, research most definitely gets the most value, when it has practical implications. The interaction with these organizations forced me to think about practical applicability which most definitely improved my work. In fact, I am very proud that my thesis became a combination of science and applied research. The fact that robustness has made it to the policy document the "MobiliteitsAanpak" and the fact that organizations like the ones mentioned are willing to do robustness analyses and are willing to make visions on robust road networks, makes me very happy. I hope that my work contributed to that somehow and I hope that we can keep working together in order to really create that robust road network.

Writing a thesis is not only about doing research. It is also a learning process that requires a lot of persistence. Without any doubt a drive is needed to keep on going. Therefore, you need to know very well what you are doing and why you are doing it. You need to know what your dream is and how you can chase after that. I have been in the lucky circumstances that a lot of people are willing to coach me in that. So thank you all very much. You made me a better and more complete researcher, consultant and, moreover a better and more complete person. Without you, the process of writing this thesis would not have been as successful and as much fun as it has been. Gerard, as always, you were very right when you quoted “het loon is met u”.

Finally, I would like to thank my parents, sisters and brothers. You enabled me to do this. You always stood behind me and you challenged me to do it right. The knowledge that you are proud of me has given me the confidence and the drive to do this. Out of curiosity and simply just because you cared, you asked me at least a thousand times when I would be finished and what my research was exactly about. Well, here is the answer to all your questions. My thesis is finished once you read this, and in this thesis you can read in detail how you can design a robust road network. Almost all of you work in the medical sector. I chose to go in another direction, but of course I do have an interest in medicine and the human body as well. Therefore, it was great fun to make a comparison between the robustness of a road network and the robustness of the cardiovascular system and the nervous system. Sanne and dad, thanks for your help with that. Sanne, I hope you like being a paranymph for the first time in your life. Hopefully, your boyfriend Leon will take nice pictures of us. Leon, thank you in advance for this. Martin, thank you for the finishing touch. Removing the word “dus” from the summary was really an enormous improvement ;-)

To you and to everybody else I would like to say: I am sorry that you are often delayed by traffic jams and that you have to wait quite a long time at the intersection on specific locations in Rotterdam, Bergschenhoek, Dordrecht or anywhere else. However, this is out of my circle of influence. You might not want to hear this, but the primary cause of congestion is you when you decide to take your car to go on the road. Unfortunately, this thesis is not a complete solution for traffic jams, especially not for regular congestion. I do hope that it contributes somehow to preventing traffic jams as a result of irregular conditions like incidents. Moreover, I hope that it contributes to offering more reliable travel times.

I wish you a lot of fun with reading this thesis. I challenge you to read it from the beginning to the end, but of course you can also just start with the summary.

Sincerely yours,
- Maaike

Collega’s, teamgenoten en ook jij,
Van Groningen naar Amsterdam, Rotterdam en Brussel,
Ik rijd dus ik ben vrij.

Table of Contents

List of Figures

List of Tables

Notation

Table N.1: Sets

A	Set of directed links (arcs)
H	Set of trip purposes
$\mathbf{I}^{\mathbf{w}}$	Set of incident types (varying in duration, capacity reduction and road type *w*)
J	Set of all routes on the network, $\mathbf{J} = \bigcup_{rs} \mathbf{J}^{\mathbf{rs}}$
$\mathbf{J}^{\mathbf{rs}}$	Set of all routes on the network between origin *r* and destination *s*
K	Total departure time period, $\mathbf{K} \subseteq \mathbf{T}$
M	Set of modes
N	Set of nodes
R	Set of origin nodes, $\mathbf{R} \subseteq \mathbf{N}$
S	Set of destination nodes, $\mathbf{S} \subseteq \mathbf{N}$
T	Total simulation period
V	Set of driver types {1,2,3}
W	Set of road types
Ω	Set of feasible route flow rates

Table N.2: Indices

a	Link index, $a \in \mathbf{A}$
h	Index of trip purposes
i	Index of incident types, $i \in \mathbf{I}^{\mathrm{w}}$
j	Route (path) index, $j \in \mathbf{J}, \mathrm{J}_{\mathrm{v}}^{\mathrm{rs}}, \mathrm{J}^{\mathrm{rs}}$
k	Departure time index, $k \in \mathbf{K}$
m	Index of modes, $m \in \mathbf{M}$
n	Index of nodes, $n \in \mathrm{N}$
r	Origin node index, $r \in \mathrm{R}$
s	Destination node index, $s \in \mathrm{S}$
t	Time index, $t \in \mathrm{T}$
v	Index of driver types, $v \in \mathrm{V}$
w	Index of road types, $w \in \mathrm{W}$

Table N.3: Input parameters

α_w^{ext}	Average value of external costs on road type w [€/km] (emission factor * costs of emission) (also for safety and noise)
α_w^{roadp}	Average costs of road pricing on road type w [€/km] (emission factor * costs of emission) (also for safety and noise)
α_w^{vod}	Average value of distance on road type w over all trip purposes [€/km]
α^{vot}	Average value of travel time over all trip purposes [€/hour]
α^{voro}	value of robustness [€/hour]
β^{m}	Parameter of the distribution model for mode m
γ^{m}	Parameter of the distribution model for mode m
κ	Parameter for node distribution fuction
Cap_a	Capacity of link a [pcu/hour]
l_a	Number of existing lanes of link a
L_a	Length of link a [km]
P	General term used for probabilities
p_{iw}	Probability per vehicle kilometre of an incident of type i on road type w
h_v	Proportion of drivers of type v
η	Size of departure time interval [hours]
ω	Link aggregation time interval [hours]
μ	Route choice dispersion parameter
$\varsigma 1$, $\varsigma 2$	Parameters that represent the importance of the distance from alternative routes
ψ_h^z, ψ_h^a	Trips produced (z) or attracted (a) per resident with trip purpose h.
$\upsilon_h^z, \upsilon_h^a$	Trips produced (z) or attracted (a) per job with trip purpose h.
C_fixed_a	Fixed costs of infrastructure on link a [€/km per simulation time $\mathbf{T}$]
C_var_a	Variable costs of infrastructure on link a [€/lanekm per simulation time $\mathbf{T}$]

Table N.4: Route variables

c_{jk}^{rs}	Actual travel costs of vehicles using route *j* from origin *r* to destination *s* departing at time *k* [€]
$\tilde{c}_{jk}^{rs}$	Perceived travel costs of vehicles using route *j* from origin *r* to destination *s* departing at time *k* [€]
π_{k}^{rs}	Minimum travel costs from origin *r* to destination *s* departing at time *k* [€]
$\overline{dist}_{jk}^{rs}$	Actual equilibrium travel distance of vehicles using route *j* from origin *r* to destination *s* departing at time *k* [km]
f_{jkv}^{rs}	Flow rate of driving class *v* on route *j* from origin *r* to destination *s* departing at time *k* [pcu/hour]
$\bar{f}_{jkv}^{rs}$	Equilibrium flow rate of driving class *v* on route *j* from origin *r* to destination *s* departing at time *k* [pcu/hour]
y_{jkv}^{rs}	Proportion of driving class *v* taking route *j* when departing at time *k* from origin *r* to destination *s*
τ_{jk}^{rs}	Actual travel time of vehicles using route *j* from origin *r* to destination *s* departing at time *k* [hours]
$\bar{\tau}_{jk}^{rs}$	Actual equilibrium travel time of vehicles using route *j* from origin *r* to destination *s* departing at time *k* [hours]
E_{jk}^{rs}	Other costs of vehicles using route *j* from origin *r* to destination *s* departing at time *k* [€]
ε_{jk}^{rs}	Perception error component for vehicles taking route *j* from origin *r* to destination *s* departing at time *k* [€]

Table N.5: Link variables

δ_{a}^{new}	$\delta_{a}^{new} = 1$ if a new link *a* is constructed, $\delta_{a}^{new} = 0$ otherwise
l_{a}^{new}	Number of new lanes of link *a*
v_{at}	Outflow rate of link *a* at time *t* [pcu/hour]
$Speed_a$	Speed on link *a* [km/hour]
rc_a	Spare capacity on link *a*

Table N.6: Node variables

b	Degree of a node

Table N.7: Demand variables

D_{mk}^{rs}	Travel demand of mode *m* from origin *r* to destination *s* departing at time *k* [pcu/hour]

Table N.8: Zone variables

Z_r	Trip production per zone *r* [trips]
A_s	Trip attraction per zone *s* [trips]
ζ_r	Balancing factor production in zone *r*
φ_s	Balancing factor attraction in zone *s*

Table N.9: Functions

F_{rsm}	Trip distribution function for origin *r,* destination *s,* and mode *m*
G_1	Set of restriction functions for the top level network design problem: optimize capacities or number of lanes
G_2	Set of restriction functions for the lower level network design problem: assignment
G_3	Set of restriction functions for the lower level network design problem: demand modelling
P	Probability density function
TTC(B)	Function of the total travel time costs [€/simulation time **T**] (*B= benefits)*
TDC(B)	Function of total distance related costs [€/simulation time **T**] (*B= benefits)*
TCV(B)	Function of the total vulnerability/reliability related costs [€/simulation time **T**] (*B= benefits)*
TCI	Function of the total infrastructure related costs [€/simulation time **T**]
Z_1	Objective function of the top level network design problem: optimize capacities or number of lanes [€/simulation time **T**]
Z_2	Objective function of the lower level network design problem: assignment [€/simulation time **T**]
Z_3	Objective function of the lower level network design problem: demand modelling [€/simulation time **T**]

1 Introduction

1.1 Problem description

The Dutch road network is, like many other road networks in the world, congested in the morning and evening peaks. The locations of congestion are quite often the same; this makes it relatively easy to take the delay of this regular congestion into account when planning a trip. However, as a result of unforeseen disturbances, also unexpectedly large delays occur. These delays can easily add up to more than one hour per trip and cause a lot of inconvenience and welfare losses for travellers and companies. This implies that the road network is vulnerable.

If no measures are taken, the Dutch road network, especially in major urban areas, will become more and more vulnerable to unforeseen disturbances, like incidents, because of the fact that the network is being used more and more intensely. Furthermore it becomes more difficult to recover from unforeseen disturbances, since the spare capacity in the network has been reduced both in place and in time. The next chapter describes the developments that result in vulnerable road networks in more detail. Disturbances like incidents cause economic damage. The opportunity costs of vulnerability in the Netherlands might increase to more than four billion Euros per year in 2030 (Snelder et al, 2008). This raises the question of what measures can be taken to reduce the vulnerability or to increase the robustness of the road network, and how these measures could be applied. These questions refer to a so called ‘network design problem’ (Yang and Bell, 1998). To be more precise, these questions refer to a ‘robust road network design problem’ (see chapter 6).

Designing and developing road networks is a difficult task, because many different aspects have to be taken into account:

- *Different objectives:* different, sometimes conflicting, objectives have to be considered which could for instance relate to economic growth, lower travel times, more reliable travel times, improved social equity, improved environmental conditions, and an improvement of liveability.

- *Different stakeholders are involved:* The national, regional and local governments, companies and travellers, travellers associations, environmental movements, and associations of people living in the neighbourhood of roads all have objectives that are more or less different from each other. All these stakeholders have to a certain extent a vote in decisions with respect to the extension of road networks.
- *Interdependencies between choices and investments:* Travellers make a lot of choices: location choice, trip departure choice, destination choice, mode choice, departure time choice and route choice. All these choices depend to a certain extent on the travel times and travel costs. The capacity investment decision depends on the travel times as well. However, if we add capacity, the travel times change and, therewith, the other choices of the travellers change as well.
- *Short term versus long term:* In practice, it takes about 15 years (PBL, 2006), from plan to realization, before a complete new road can actually be used by travellers and truck drivers. From there on, the economic lifetime of a road is quite long, about 30-100 years. This explains why these kinds of major decisions have to be taken with great care. Opposed to this long term decision, measures have to be taken that relieve the short time problems. This could be any kind of measure, like small infrastructure investments. It is desirable that short term and long term investments are in line with each other, but it could also be that they contradict.
- *Long term uncertainties in demand:* The future demand is uncertain. This makes it very difficult to make investment decisions. If we make a large investment and the demand does not increase much, we overinvest in infrastructure. However, if the increase in demand is larger than expected, we under invest, which will result in high congestion levels and a high vulnerability cost.
- *Short term uncertainties:* Both the demand and supply vary in the short term as a result of accidents, roadwork, bad weather conditions, events, seasonality etc. That means that we can no longer make a deterministic design, but that we will have to make a probabilistic design that considers these short term risks.
- *Laws and procedures:* In the Netherlands, there are several laws with respect to the construction of new roads or the extension of existing roads like the 'Tracéwet'[1], 'Wet geluidhinder'[2], 'Wet luchtkwaliteitseisen'[3] and 'Onteigeningswet'[4]. These laws require that procedures like the 'tracé-procedure', 'm.e.r.-procedure'[5], expropriation procedure and advice from commissions and governments are followed. Most other countries have similar laws and procedures that ensure that decisions with respect to infrastructure are taken with great care. As a result of these laws and procedures (and other problems), it takes a long time before the construction of new roads or lanes can actually start.
- *The political environment*: Since governments in the Netherlands change at least every four years, also the attitude towards the construction of new roads and the extension of existing roads changes. A decision, taken by a certain government, to build a new road immediately leads to costs, but the people only experience the gains of new roads years and years later under a different government. This makes it difficult to take these kinds of decisions. What makes it even more complex is that, in total, a government can only

[1] The 'Tracéwet' is a law that requires that the tracé-procedure is followed. The tracé-procedure regulates all the steps that need to be followed before a new road can be consturcted or the capacity of an existing road can be extended.
[2] The 'Wet geluidhinder' describes the rules and norms with respect to noise polution.
[3] The 'Wet luchtkwaliteitseisen' describes the rules and norms with respect to air quality.
[4] The 'Onteigeningswet' require that expropriation procedure are followed.
[5] The goal of the m.e.r.-procedure is to make sure that environmental aspects are considered fully in the preparation and adaptation of plans and the decision making process for initiatives of public and private parties.

invest a certain amount of money every year, which implies that a balanced choice has to be made among investing in infrastructure, healthcare, safety, developing countries, etc.

- *Different measures:* Above only capacity expansion has been mentioned as a measure to improve the outcomes from the transport system. However, many other kinds of measures are possible, such as those that focus on mobility management (like encouraging working at home), encourage the use of other modes than a car (like the train), introduce road pricing, apply traffic management and incident management strategies, and introduce intelligent transport systems (like adaptive cruise control). A balanced choice has to be made among these measures.

The decision about where new roads are to be constructed or which roads get increased capacity is also difficult because of the fact that there are many interactions with other systems. For instance, there is a strong relationship between land use and infrastructure planning. On the one hand, the transport infrastructure is partly determined by the spatial structure, and, on the other hand, the (economic) spatial structure is partly determined by the transport infrastructure. The background between such interactions is that both persons and companies usually have a preference for settling at well accessible places, and that extra infrastructure is usually built at places where congestion is worst. Furthermore, there is a relationship between road transport and other modes. For instance, passenger cars compete with different public transport modes and with bicycles and walking, and trucks for freight transport compete with other modes like trains, inland waterway shipping, sea and air transport. The interaction among modes implies that in the development of road networks, the other infrastructure networks or service networks also have to be taken into account. Another level of interaction is the interaction between the demand for transport and the level of service for the different modes. It is likely that an improvement of the road network will attract new users to the roads due to latent demand and induced demand, i.e. changes in destination and mode choice.

The above mentioned aspects/dilemmas show that the network design problem is complex from a practical point of view. From a theoretical point of view, this problem is complex as well. In literature, the Network Design Problem (NDP) has been discussed extensively. In chapter 6, this literature is discussed in more detail. Many models have been proposed that can solve the NDP. However, they are hardly ever used in practice, because they require many simplifications with respect to demand for travel, the behaviour of travellers, and the objectives and/or the size of the network. Furthermore, computation time is still a problem. Nevertheless, much work is being done in this area in order to improve the models and to make them suitable for usage in practice. Including robustness is one step further. The robustness of transportation networks is a relatively new research area. There are several definitions, but none of these are commonly accepted. Specific indicators for robustness are scarce and robustness against short-term uncertainties is, as far as this author knows, not yet explicitly considered in the network design problem.

1.2 Research scope

In the previous section, the full problem of robust road network design was described. However, in this thesis it is impossible to include all elements of this problem. Therefore, some choices had to be made. In this section the scope of the research is presented.

Scale:
This research focuses on robust road network design in the Netherlands. Of course, congestion and unreliable travel times are not specific to the Netherlands. Many other countries are dealing with similar problems. Therefore, robust road networks are also important for other countries. The design method for designing robust road networks that is presented in this thesis is easily transferable to other countries. Therefore, this thesis is relevant for other countries as well.

Furthermore, the focus of this research is on the road network. This does not imply that robustness of other transportation networks, such as the rail network and the inland waterways network, is not important. On these networks disturbances also occur. Furthermore, these other networks offer the travellers other options for travelling and offer backup options for the road network. In this thesis a simple comparison between the road network and other networks is made to learn from the way in which robustness analysis is done in those networks. Within the road network, the focus is on motorways and regional roads and not on local roads.

Interactions with spatial developments:
The interactions between spatial developments and infrastructure are not explicitly considered in the research. The network design method can consider changes in spatial developments exogenously. For instance, land use models and spatial general equilibrium models can be used to forecast the spatial developments. However, as far as is known to this author, these models cannot yet deal with uncertainties in travel times. Another way to consider the interaction between spatial developments and infrastructure is by making expert judgments and translating these judgments into input that can be used in the design method.

Disturbances:
Road networks can be made robust against many long term and short term uncertainties in demand and supply. This thesis focuses on the short term uncertainties. As is shown in section 3.3, there are many short term uncertainties. We will focus on traffic incidents, because incidents often cause unexpected large travel time delays in the Netherlands. Therefore, robustness against incidents is important in the Netherlands. It is likely to be beneficial to improve the robustness of the road network in order to reduce the effects of these kinds of disturbances.

Measures:
Robustness is only one of the factors that influence the reliability of travel times. The behaviour of drivers and network managers is, for instance, important as well. The same is true for other possible changes in the future, like developments with intelligent transport systems, road pricing and climate change. All these factors can be considered in the network design method that is presented in this thesis by making exogenous assumptions about these developments and translating them into input that can be used in the design method.

Finally, this research focuses on the structure of the road network and not on the construction of roads. This implies that types of asphalt, road curves, etc. are not considered. All kinds of other measures, like traffic management measures including offering information, are important in order to really benefit from a robust structure. Therefore, these measures are briefly addressed. However, since the focus is on the structure of the network, these measures are not worked out in detail.

1.3 Research objective, research questions and thesis outline

Our objective is to develop a network design method for designing road networks that are robust against incidents that is applicable to real sized networks and can be used in practice.

In order to develop a network design method for designing robust road networks, we need to answer the following research questions:

Research question 1: What is the importance of robustness for the road network?
In section 1, we explained that drivers more and more often are facing long unexpected delays. Of course, this brings along costs, either because travellers have to take into account a buffer and therefore might arrive early if nothing happens during the trip, or because they arrive late, which causes economic damage (depending on the purpose of the trip) and discomfort. The same is true for freight transport. The question is how large these costs are. The answer to this question illustrates how large the vulnerability problem is and indicates how much money can be invested in order to make the network more robust. Chapter 2 addresses this question.

Research question 2: How is robustness defined?
In this chapter, the terms robustness and vulnerability have often been mentioned, but what is exactly meant by these terms? In Chapter 3, definitions from literature are reviewed and the definitions that are used in this thesis are presented. For reasons of clarity we already present the answer to this question here:

> Robustness is the extent to which, under pre-specified circumstances, a network is able to maintain the function for which it was originally designed.

Vulnerability is the opposite of robustness. A network that is vulnerable is not robust, and vice versa.

Research question 3: Against which disturbances should the network be made robust?
A lot of supply and demand related disturbances can occur. We mention for example incidents, road works, flooding, earthquakes, snow, rain, events and holiday traffic. However, do we really want to make our network robust against all these disturbances? In Chapter 3, the different disturbances are mentioned and a choice is made on which disturbances this thesis focuses.

Research question 4: Which elements determine the robustness of a road network?
The definition of robustness is still relatively broad. If we want to focus on making a network more robust, it has to be known what makes a network robust, or what makes a network vulnerable. Chapter 2 describes some elements that make different networks vulnerable or robust. These elements are combined in Chapter 3 into characteristics that make a road network robust.

Research question 5: What is the relationship between robustness, travel times and travel time reliability?
In this chapter the words robustness, vulnerability, travel times, and reliability have already been mentioned several times. The question is how these terms are related. In Chapter 3, this question is addressed.

Research question 6: Which indicators can be used to measure robustness?
For robustness, many indicators can be used. However, there is not yet one commonly accepted indicator. Chapter 3 addresses the question which indicator(s) can best be used for the purpose of designing robust road networks.

Research question 7: How can robustness be evaluated?
Evaluating the robustness of a road network is complex, because many different disturbances can occur on many different locations. Furthermore, the response of people to these incidents can differ and the way in which traffic jams build up or dissolve is very important. Knoop (2009) shows this in detail. Because of this complexity, it is quite difficult to evaluate the robustness of a network. In road network design the robustness of many different possible networks has to be evaluated. This raises the question how robustness can be evaluated in a short computation time by making as few as possible compromises on the quality of the indicator for robustness? This question is addressed in Chapter 4.

Furthermore, the question of how gains in robustness can be weighed against the investment costs that are needed for improving robustness is addressed in Chapter 4.

Research question 8: Which measures can be taken to improve the robustness of a road network, and what are their effects?
Once it is known how vulnerable the road network is, the question is which measures can be taken to improve the robustness of the network and what effects can be expected of these measures. These questions are addressed in Chapter 5.

Research question 9: How can robustness be integrated into a network design method?
The 9th question is how robustness can be included in a network design method. This question is difficult to answer because network design is complex by itself and including robustness in this problem is even more complex. Chapters 6 and 7 address this question.

Research question 10: How can the method be applied to large scale networks?
The final question is whether the method that is presented in this thesis can really be applied to large scale networks and, when doing so, what is the quality of the method and what do we learn from applying the method. These questions are addressed in Chapter 8.

The above presented explanation of the research questions already gives an indication of the outline of this thesis. To be more precise, in the next chapter (Chapter 2) vulnerability is described in more detail. It is shown how vulnerable the road network is compared to the vulnerability of other networks like the human cardiovascular system.

In order to develop a method for improving the robustness (or reduce the vulnerability), we first need to define robustness together with all related aspects. Furthermore, we need to specify indicators that can be used to quantify robustness. This is done in Chapter 3.

In Chapter 4, a method for evaluating the robustness (expressed by the indicators chosen in Chapter 3) is presented, as well as a method for including robustness in cost-benefit analysis.

The remaining chapters focus on the design method. First we present a long list of measures that can be used for improving the robustness of a road network (Chapter 5). Thereafter, in Chapter 6, we give a mathematical formulation of the robust road network design problem. In

Chapter 7 a solution algorithm is presented. Finally, in Chapter 8 the solution algorithm is applied to Amsterdam and its surroundings.

In the last chapter (Chapter 9) the main conclusions and implications of the research are discussed and recommendations for future work are presented. The outline of this thesis is summarized in Figure 1.1.

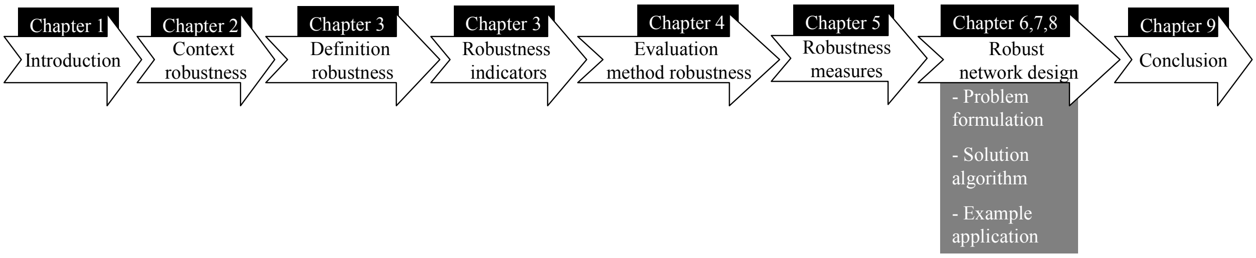


Figure 1.1: Thesis outline

1.4 Scientific, methodological and practical contribution

This thesis contributes to the problem of making a network more robust in several ways. In this section the scientific, methodological and practical contribution is explained.

Scientific and methodological contribution
This thesis contributes to the understanding of the concept of robustness. A framework is developed in which the relationship between robustness, vulnerability, disturbances, and reliability is clearly shown. Different indicators for robustness are compared and combined to a new indicator that can be computed in a short computation time and can therewith be used in designing robust road networks.

Related to this is the fact that we developed an evaluation framework that can be applied to the current network or to a planned future network in order to identify vulnerable road sections within a shorter computation time than the traditional evaluation methods for robustness and without compromising too much on the requirements for a detailed evaluation of robustness. For instance, the evaluation method considers flow dynamics, spillback effects and alternative route choice. The evaluation method combines existing evaluation models (a demand model, a dynamic traffic assignment model and a marginal incident computation model) in such a way that they can be applied to evaluate the robustness of real-size road networks. Forecasting the effects of incidents is complex and requires dynamic traffic assignment models that take spillback effects into account, as is explained for instance by Knoop (2009). In order to evaluate the robustness of a road network against incidents, many simulations have to be carried out with these dynamic traffic assignment models. This results in a computation time that is too long to be of practical use. In this thesis, it is explained which simplifications can be made in order to make the existing methods applicable to real-size networks and what the implications of these simplifications are.

In this thesis, we present a comprehensive overview of all kinds of measures for improving the robustness of a road network, and for a selected set of measures the individual effects of these measures on robustness are indicated.

A network design method is presented by which robust road networks can be designed by using the robustness measures that relate to the network structure. This method has several advantages compared to the methods that are usually used in practice and found in literature for designing networks:

- The method that is presented combines expert knowledge with advanced modelling techniques. In the literature, usually only models are used; in practice, usually only design methods are used that do not make use of optimization models. By involving experts (including stakeholders) in the design process, a common ground is created that makes it easier to actually realize the network in practice. Furthermore, models are by definition a simplification of reality. By including expert judgement in the design process, a lot of knowledge can be added and many different kinds of measures can be included in the network design. Usually, in optimization models, the capacity or number of lanes is the only design variable.
- By combining expert knowledge with advanced models, it becomes possible to apply the design method to large scale networks. Therewith, this method can be applied in practice. This sounds like a minimum requirement of a network design method. However, in the literature the network design problem is usually addressed from a theoretical point of view, which results in advanced network design models that can in most cases be applied only to very small networks consisting of several links that can be optimized. Of course, these theoretical developments have their value, because they give insight into the problem and because they can most likely be applied to larger networks in the future. However, the advantage of our method is that it can already be applied to larger networks and still uses many of the newest insights with respect to modelling techniques.
- We integrated short term variation in supply (incidents) into the robust network design problem. As far as is known to the author, the robust network design problem is only referred to in relation to long term variations in demand. Furthermore, although optimizing the reliability of travel times in the network design problem is often mentioned in the literature, the focus in this thesis differs from that, because it focuses on incidents that require a different set of measures to be taken.
- The optimization model that is used optimizes the spare capacity in a network by taking time dynamics, spillback effects, demand effects, and alternative routes into account. This is a combination that has not been considered yet in the literature. In fact, in network design, dynamic traffic assignment models are hardly ever used. Furthermore, demand effects are not often considered in the network design problem either.

Finally, in this thesis different applications of the evaluation and design method are presented that give insights into the robustness of the road network of the Netherlands and, in general, into which structures are more robust than other structures.

Practical contribution
A part of the work presented here has already found its way to practice. We have shown the importance of robust road networks to the Advisory Council for Transport, Public Works and Water Management (Snelder et al., 2008). This was one of the factors that cleared the way for actually including robustness in different policies.

The framework for robustness (definition, indicators, etc.) appeared useful in explaining what robustness is about and in operationalizing robustness. This framework was used as an input into discussions about robustness with the Advisory Council for Transport, Public Works and Water Management and the Directorate-General for Public Works and Water Management and worked out in more detail with the help of these organizations (Snelder et al., 2008 and 2010d).

Furthermore, we analysed the vulnerability of future networks for the National Market and Capacity Analysis (NMCA). The method that was used for doing that (Snelder et al., 2010a) differs from the method that is presented in this thesis. However, it is based on the insights about robustness that are presented in this thesis. The vulnerability analysis that was carried out enables policymakers to develop a strategy for improving the robustness of future road networks.

We analysed the robustness/reliability benefits of different measures of the policy document *MobiliteitsAanpak* (Dutch Ministry of Transport, Public Works and Water Management, 2008) by using the approach presented in this thesis. This analysis showed that, compared to the traditional travel time benefits, other (robustness/reliability) benefits need to be considered and can be quite large (Snelder et al., 2009c). This makes it easier to invest in robustness measures, because the benefit-cost ratio gets larger.

Elements of the design method that combines expert knowledge with models were used for making a vision on a robust road network for the area The Hague – Rotterdam in cooperation with the ANWB (Schrijver et al., 2008). This vision illustrates which measures can be taken to improve the robustness of a road network and how this can be done. The complete design method that is presented in this thesis can be used to do the same in a more advanced way. This enables policymakers to get a clear vision on how to improve the robustness of the road network.

Although elements of this thesis have already found their way into practice, this thesis and the policy documents to which it contributed still need to be followed up in practice. The true value of this thesis for society can be seen only in the future when, hopefully, road networks are actually made more robust against disturbances like incidents. This should result in fewer unexpected delays and more reliable travel times.

2 The vulnerability of networks

2.1 Introduction

As Boccaletti et al. (2006) explained, networks are all around us, and we are ourselves, as individuals, the units of a network of social relationships of different kinds and, as biological systems, the delicate result of a network of biochemical reactions. There are physical networks, such as electric power grids, the Internet, highways or subway systems, and neural networks, and there are abstract networks, such as networks of acquaintances or collaborations between individuals. During the last 15 years, a lot of research has been done into these complex networks, i.e. networks whose structure is irregular, complex, and dynamically evolving in time, with the main focus moving from the analysis of small networks to that of systems with thousands or millions of nodes, and with a renewed attention to the properties of networks of dynamic units. The road network is one of these complex networks. For an extensive description about the theory behind complex networks, we refer to (Boccaletti et al., 2006) in which an overview is given of complex network theory and reference is made to many other papers about complex network theory. Below, we give a short summary of the parts of complex network theory that are most relevant for robustness analysis of the road network.

Graph theory (e.g. Bollobás, 1985; Bollobás, 1998; West, 1995; Harary, 1995) is the natural framework for the exact mathematical treatment of complex networks and, formally, a complex network can be represented as a graph. A graph $G = (\mathbf{N}, \mathbf{A})$ is represented by a set of nodes **N** and a set of links **A** that connect the nodes. The links can be directed or undirected and they can have a weight. In road networks, the weights can for instance be the lengths of the links or the travel times over the links. Figure 2.1 shows schematically the network elements that are used throughout this thesis:

- *Region/zone*: a region is a collection of origins (starting point of trips) and destinations (end points of trips). In practice all trips start and end at a specific location, such as a home or a company. Since this thesis considers transport networks on a macroscopic level, trips are not considered on the microscopic level of homes and companies. Instead whole regions, also referred to as zones, are considered.

- *Centroïd*: a centroïd is a fictive point in a zone from which car or truck drivers enter the network or exit the network.
- *Connector/feeder link*: a connector, also referred to as a feeder link, connects a centroïd to a node in the 'real' road network. One centroïd can be connected to the road network by multiple connectors.
- *Node*: a node/vertex is a point in the network where multiple links come together.
- *Link*: a link/edge is a connection between two nodes. This implies that the roads in the road network are links.

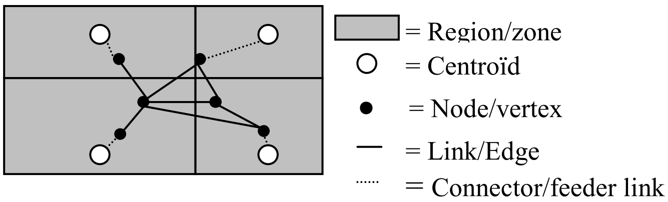


Figure 2.1: Network elements

The *Handbook of Graphs and Networks* (Bornholdt and Schuster, 2003) explains that graph theory initially focused on regular graphs. Since the 1950's, large networks with no apparent design principles were described as random graphs, proposed as the simplest and most straightforward realization of a complex network. Two Hungarian mathematicians, Paul Erdós and Alfréd Rényi, had the biggest impact on the development of a mathematical theory of random graphs. The so called Erdós-Rényi (ER) model connects every pair of nodes with probability p, creating a graph with approximately $pn(n-1)/2$ undirected randomly distributed edges. In this formula n is the number of nodes. Thereafter it became clear that most real-life networks are not random, since not all nodes have the same number of edges. The spread in the number of edges of the nodes, or the nodal degree, is characterized by the degree distribution $P(b)$, which gives the probability that a randomly selected node has exactly b edges. Since in a random graph the edges are placed at random, the majority of nodes have approximately the same degree, close to the average degree $(n\text{-}1)p$ of the network. Indeed, the degrees in a random graph follow a Poisson distribution with a peak at b. It turned out that in most real-life networks the degrees do not follow a Poisson distribution. Instead, the degree distribution has a power-law tail $P(b) \sim b^{-\kappa}$. Networks with a power-law degree distribution are called scale-free. Scale-free networks arise because new nodes are preferably attached to highly connected nodes. Hub and spoke networks are then likely to arise. Real-life networks significantly deviate from the Erdós-Rényi random graphs because they exhibit a power law (scale-free) tail with an exponent κ taking a value between 2 and 3. Moreover, real-life networks are characterized by correlations in the node degrees, by having relatively short paths between any two nodes (small world property[6]), and by the presence of a large number of short cycles or specific motifs[7].

According to Boccaletti et al. (2006), robustness refers to the ability of a network to avoid malfunctioning when a fraction of its constituents is damaged. Robustness was one of the first issues to be explored in the literature on complex networks. Boccaletti et al. (2006) review the

[6] The small world property is mathematically characterized by an average shortest path length that depends at most logarithmically on the network size (Watts, 1999).
[7] A motif is a pattern of interconnections.

main results concerning the resilience of networks to both random failures and intentional attacks. The problem can be encountered in two different variants. The first one, referred to as static robustness, is meant as the act of deleting nodes without the need for redistributing any quantity that is being transported by the network. This is the case, for instance, in a social network in which we cut relationships between individuals forming the system. On the other hand, dynamic robustness refers to the case in which the dynamics of redistribution of flows should be taken into account. Obviously, the road network is an example of a complex network in which dynamic robustness should be considered. For instance, an accident could block a road completely or partially, which causes delays for the road users, and some of the road users will change their route choice or maybe even their departure time choice, mode choice or destination choice. Boccaletti et al. (2006) explain that the two types of robustness are conceptually similar, but while the first can be analytically treated, e.g. by using the tools of statistical physics such as percolation theory, the analytical treatment of the second case is harder and in almost all cases one has to rely on numerical simulations.

Different researchers argue that the prevalence of small world networks in biological systems may reflect an evolutionary advantage of such an architecture that makes them more robust to perturbations than other network architectures. In a power law distributed small world network, deletion of a random node rarely causes a dramatic increase in mean shortest path length, because most shortest paths between nodes flow through hubs, and if a peripheral node is deleted it is unlikely to interfere with flows between other peripheral nodes. However, if a random deletion of a node hits a hub, the average path length can dramatically increase. By contrast, in a random network, in which all nodes have roughly the same number of connections, deleting a random node is likely to increase the mean shortest path length slightly but significantly for almost any node deleted. In this sense, random networks are vulnerable to random perturbations, whereas small world networks are robust. However, small world networks are vulnerable to a targeted attack of hubs, whereas random networks cannot be targeted for catastrophic failure (Albert et al., 2000).

Jamakovic et al. (2006) applied graph theory to the Dutch road network. They discussed how the underlying complex principles, captured in a wide range of topological characteristics, are related to the robustness of the road graph (static robustness). They conclude that the characteristics of the road network substantially differ from that of many other real-life networks. Despite the awareness that the Dutch road infrastructure substantially differs from many other real-world networks, its topological characteristics do resemble specific complex structures, for instance the power grid. In fact, it turned out the road network could best be described by a 'hopcount' distribution $\Pr[H = h]$, which is the probability that a random pair of nodes are at h hops (or links) from each other. The road network belongs to the class of D-lattice graphs, where $E[H] \sim (D/3)*N^{(1/D)}$ and D is the lattice dimension.

In (Boccaletti et al., 2006), (Nagurney and Qiang, 2009), and (Bornholdt and Schuster, 2003), different networks are described not only from a static but also from a dynamic point of view. The vulnerability of these networks is also described from the same perspective. Bornholdt and Schuster conclude for transportation networks that these networks are particularly interesting, since the one-dimensional dynamics on the links (the occurrence and solving of congestion) interacts with the network aspects. For example, kinematic waves can travel through an intersection, causing complicated dynamics there. In fact, very little seems to be known of these link-network interactions, especially for large systems with many links (roads) and vertices (intersections). In addition, the particles/agents in traffic systems are "intelligent". This means that they have strategic goals, with the consequence that no two particles are interchangeable, and that different particles, when confronted with the same situation, can make different decisions. In practical terms, for transportation simulations this "intelligence" involves aspects like route choice, mode choice, or activity generation. Moreover, agents adapt or learn, which means that they should be able to remember past behaviour and past performance, to construct new plans, and to try them out.

In addition to the work that has been done on vulnerability analysis according to the graph theoretical approach, this chapter describes in more detail why the road network is vulnerable and how vulnerable the road network is. This chapter starts with some general developments that explain why the Dutch road network is vulnerable and will become more vulnerable in the future if no measures are taken. The Dutch road network is taken as an example throughout this thesis. However, in many urbanized areas in other countries similar problems occur. In the remainder of this chapter, a comparison is made with a selection of other networks (rail network, inland waterways network, cardiovascular and nervous system and telecommunication networks) in order to learn what makes them vulnerable or robust and how that can be translated to the road network.

2.2 The vulnerability of road networks

This section explains how vulnerable the road network is and why it is vulnerable. The objective of this section is to show that it is important to improve the robustness of the road network. The terms and indicators and methods that are used in this section are explained in more detail in the following chapters.

Development of the Dutch motorway network (road supply)
In Figure 2.2**,** the development of the Dutch motorway network is presented. This network was developed according to the then current spatial and transportation planning philosophies. The figure shows that between 1960 and 1980 the largest part of the motorway network was constructed. Thereafter, only a few new motorways were constructed and, on some motorways, the number of lanes was extended.

The left hand side of Figure 2.3 shows the number of lanes on the road network of 2008. Most of the motorways have two lanes in each direction. Only between and around the large cities are there three or more lanes in each direction. The right hand side shows the locations on which extra lanes will be built up to 2020. The total lane length of the motorways in the Netherlands grew from 8800 km in 1980 to 11090 km in 2000 and will increase further by about 10% (Dutch Ministry of Transport, Public Works and Water Management and VROM, 2004) up to 2020. The most recent policy document "MobiliteitsAanpak" (Dutch Ministry of Transport, Public Works and Water Management, 2008) announces that the capacity of the most important motorways in the western part of the country (the Randstad) will be extended to at least four lanes in each direction.

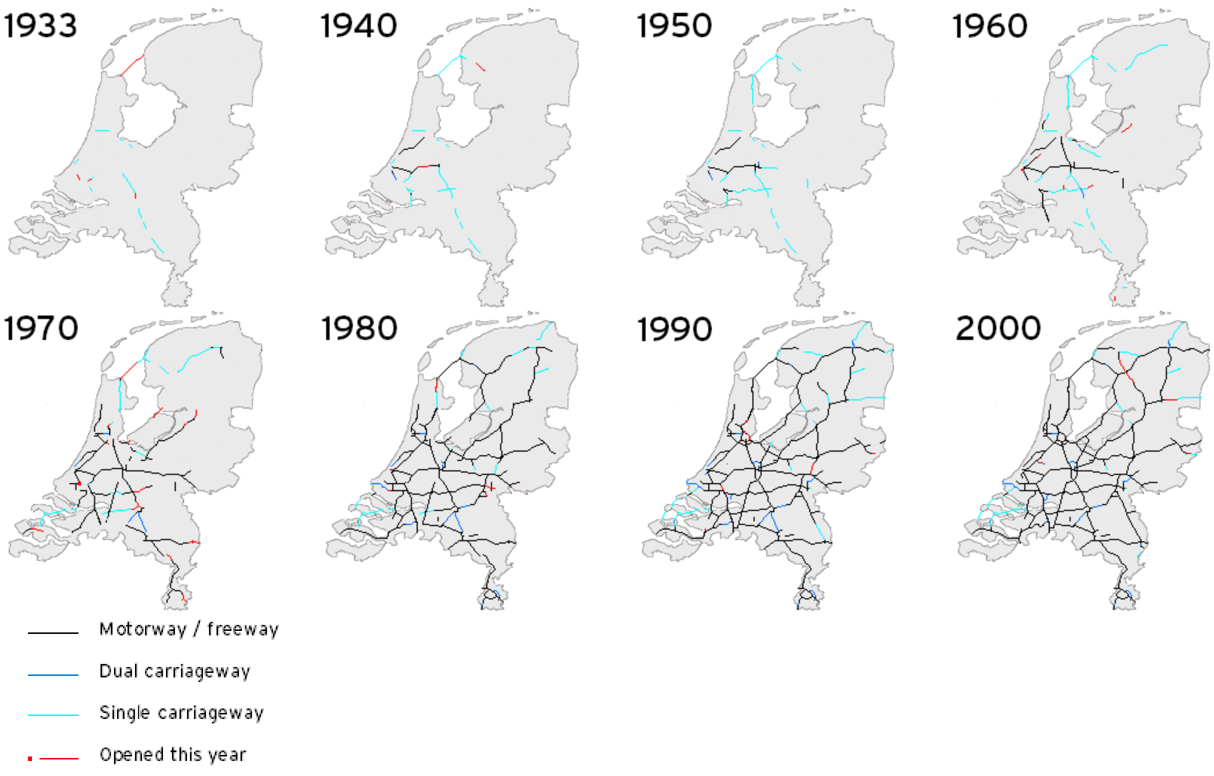


Figure 2.2: Development of the Dutch motorway network[8]

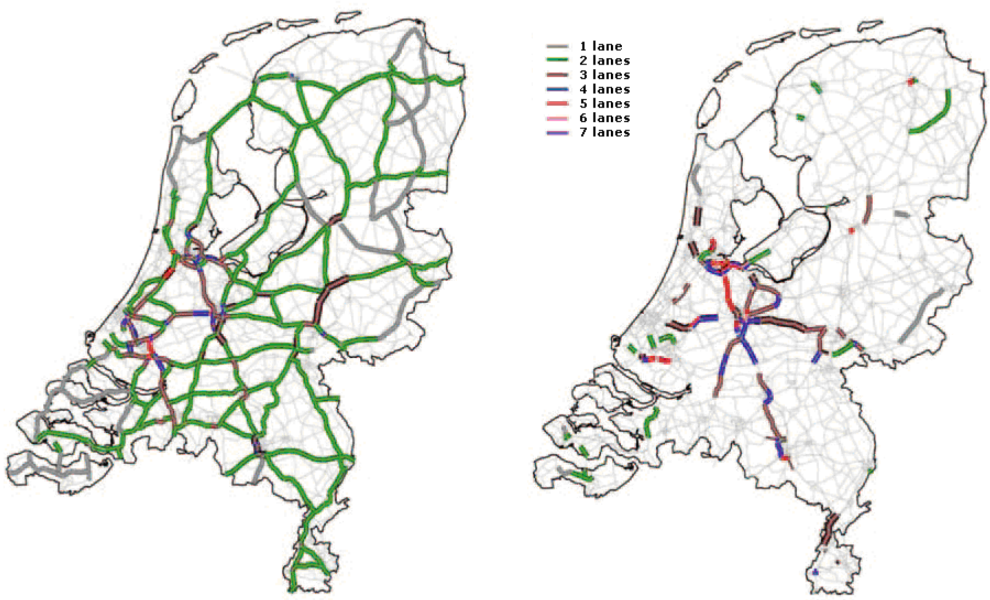


Figure 2.3: Number of motorway lanes in 2008 (left) and the changes in 2020 (right)[9]

The network development strategies in the Netherlands have been oriented towards a maximal usage of the motorways. Martens et al. (2004) compared the road networks of the Randstad, the 'Rijn-Ruhrgebied' in Germany, the 'Vlaamse ruit' in Belgium, and 'North-west UK'. The

[8] Source: www.autosnelwegen.nl
[9] Source: (Dutch Ministry of Transport, Public Works and Water Management and VROM, 2004)

Randstad is a part of the Netherlands in which the four biggest cities Amsterdam, Rotterdam, The Hague, and Utrecht are located. It was shown that the secondary road network of the Randstad is less developed compared to regions abroad. The accessibility (number of on ramps and off ramps) of the motorway network is higher. Of course, there are good reasons for this policy, of which the main arguments are probably that the motorways are safer than the secondary network and that the Netherlands has a high population density, which leaves less space for road networks compared to other countries. However, the consequence is that a large percentage of the traffic on the motorways in and around cities is short distance traffic (<30 kilometres). Persons drive there mainly because there are no good alternatives available. This implies that short and long distance traffic interfere with each other on the motorway network. When congestion occurs, due to expected or unexpected circumstances, this becomes a problem – for instance, because of the fact that for long distance traffic a high speed is usually much more important than for short distance traffic. A second consequence is that there are hardly any fallback options in case an accident occurs.

Development of the number of vehicle kilometres driven (trip demand)
In the Netherlands, scenario studies are carried out in order to forecast the long-term effects of the current policy on for example traffic and transport. The most recent scenarios are the four 'Welfare, Prosperity and Quality of the Living Environment' (WLO) scenarios (Van Beek et al., 2006). These scenarios vary in the extent to which nations and international trade blocks will cooperate and exchange and the balance between market forces and a strong public sector:

- Global Economy: emphasis on international cooperation and private responsibilities.
- Strong Europe: emphasis on international cooperation and public responsibilities.
- Transatlantic Markets: emphasis on national sovereignty and private responsibilities.
- Regional Communities: emphasis on national sovereignty and public responsibilities.

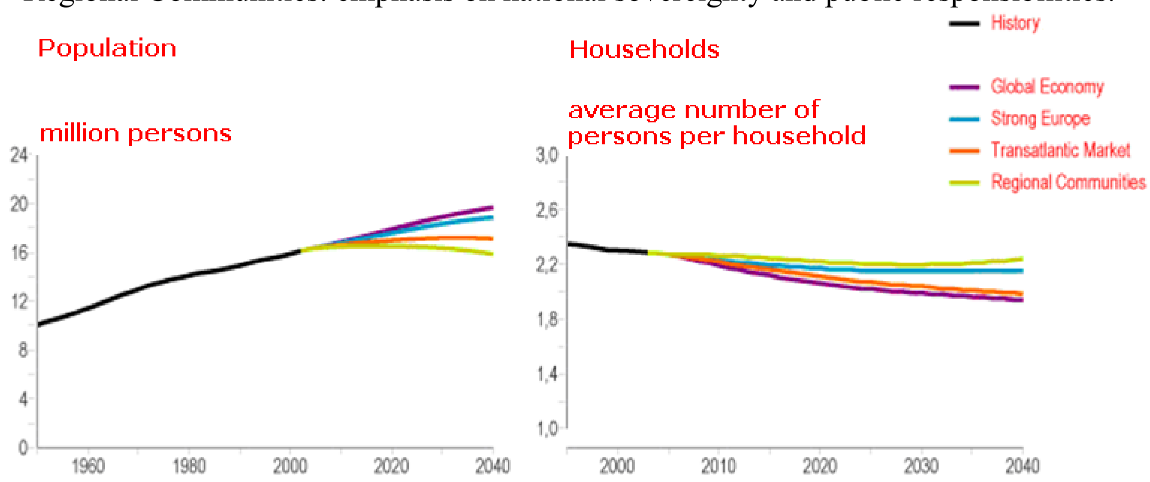


Figure 2.4: Development of population size (left) and the average number of persons per household (right)[10]

In two out of the four WLO scenarios, the total population continues growing (Figure 2.4). In the Transatlantic Markets scenario, the population stabilizes, and in the Regional Communities scenario, the population decreases. Figure 2.5 shows how the growth/decrease in population is expected to be divided over the country. From this figure it can be concluded that the population in and around the largest cities in the Netherlands is expected to continue

[10] Source: Van Beek et al., 2006

growing. These are so called agglomeration effects. Finally, the picture on the right of Figure 2.4 shows that the average household size decreases in two out of the four scenarios.

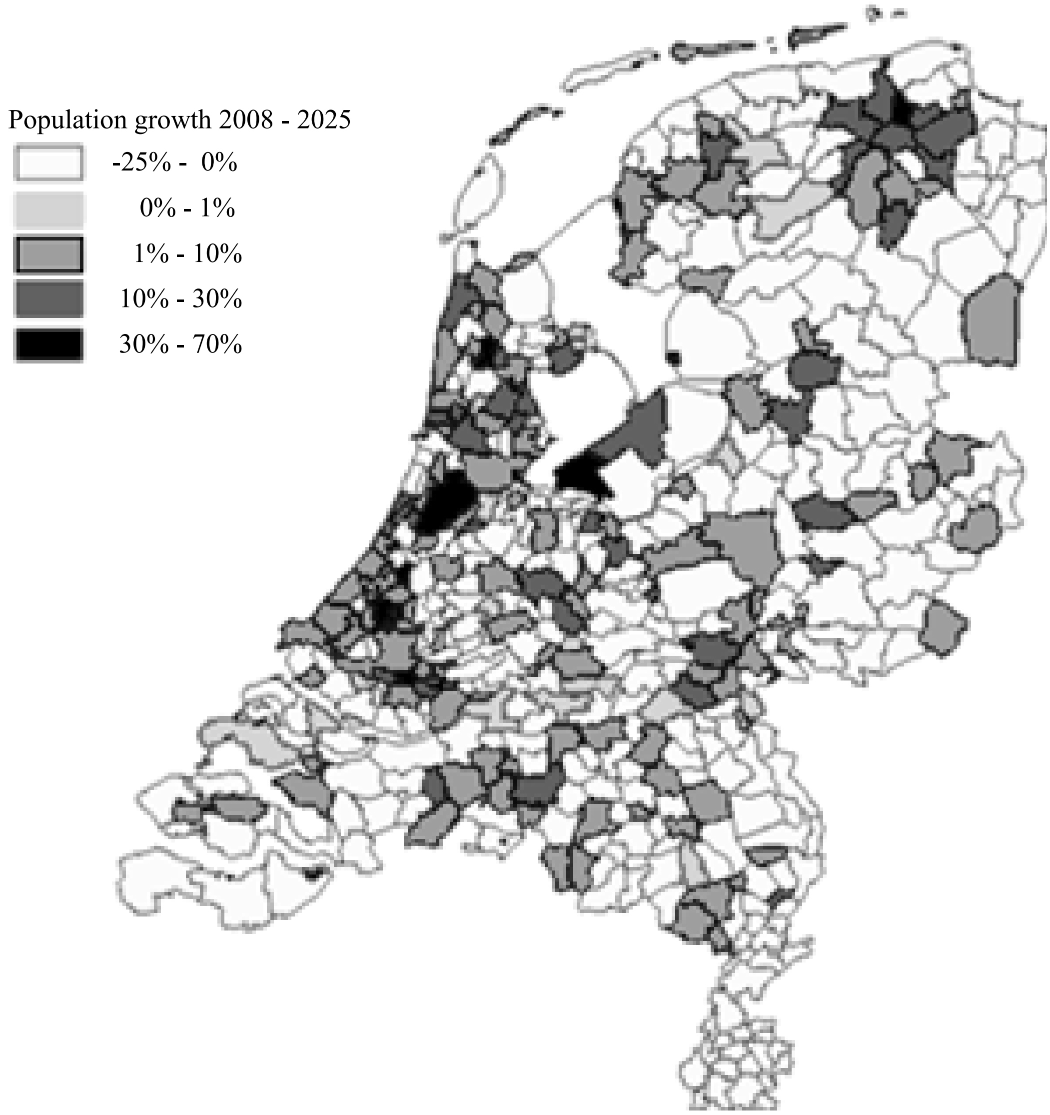


Figure 2.5: Forecast population growth per region between 2008 and 2025[11]

In the Netherlands, the number of cars grew from 4.2 million cars in 1980 to 6.2 million cars in the year 2000, and is expected to grow to 8.8 million cars in 2020. Since the population size, the average household size, and car ownership are important drivers for mobility, it can be expected that the total number of vehicle kilometres driven will continue growing in the future. This is confirmed in Figure 2.6. This figure shows that, if road pricing is not introduced, this number is expected to grow by between 25% and 85% in the period 2002-2040. In the situation with road pricing, this increase is expected to be between 15% and 55%.

[11] based on data from www.cbs.nl

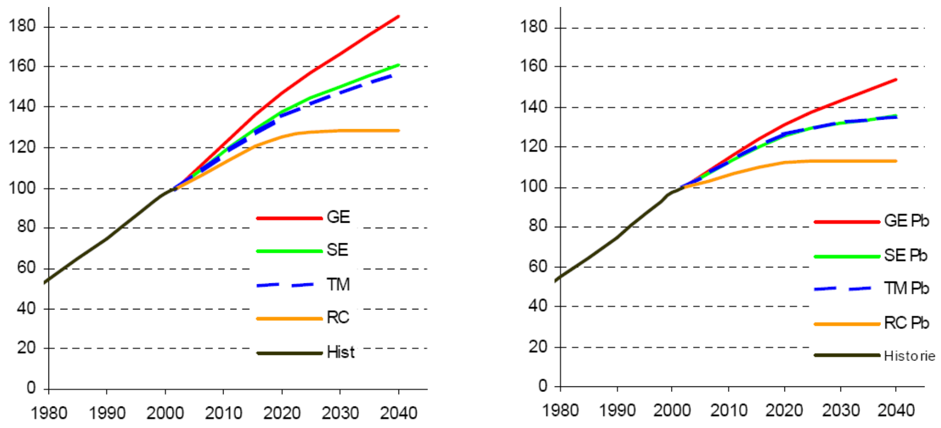


Figure 2.6: Development in the number of vehicle kilometres driven on the main road network, scenarios without road pricing (left) and scenarios with road pricing (right), vehicle kilometres (index 2002 = 100)[12]

Growth in congestion and increase in unreliability

The fact that the number of lane kilometres in the Netherlands grows slower than the number of vehicle kilometres explains why the congestion in the Netherlands is becoming more and more of a problem. In Figure 2.7, the daily congestion in the Netherlands in 2009 is shown.

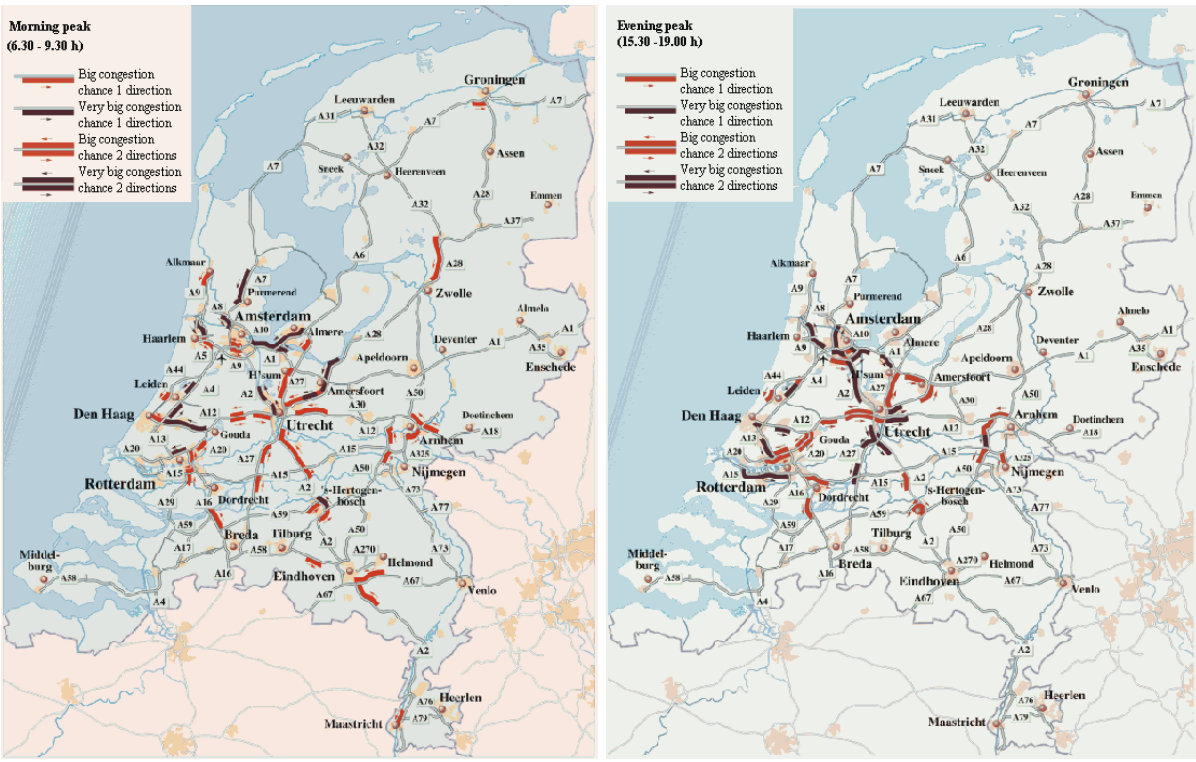


Figure 2.7: Daily congestion in the morning (left) and evening peak (right) in the Netherlands in 2009[13]

[12] Source: (CPB, 2006)
[13] Source: www.anwb.nl

Besides increasing congestion, travel times are becoming more unreliable. In Figure 2.8, the development of the number of vehicle loss hours and the reliability of travel times is shown. Reliability is measured as the percentage of trips in which the traveller arrives at his or her destination on time. For trips shorter than 50 kilometres, 'on time' means that the absolute travel time may not deviate more than 10 minutes from the expected travel time. For longer distances, the absolute deviation of the expected travel time may be at most 20 percent. The expected travel time is the median of the travel times. In Figure 2.8, reliability was measured on 106 trajectories. Since these trajectories are located on the most congested locations in the Netherlands, the numbers are biased. Nevertheless, the figure shows that the peak period reliability of the motorways for which reliability was measured decreased in the period 2000 – 2007 from 94% to 90%.

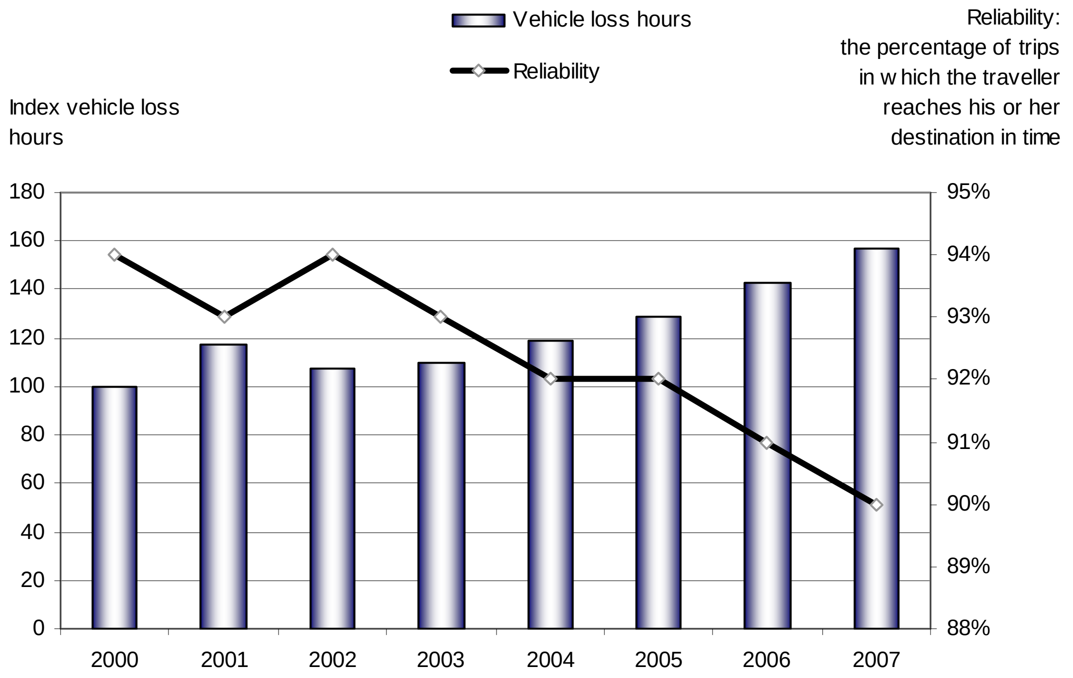


Figure 2.8: Development of congestion and reliability, 2000-2007[14]

The increase in unreliability is partly explained by the fact that the road network is becoming more vulnerable (= not robust). Below, we give an example that shows the existing road network is vulnerable.

Example: the existing road network is vulnerable

On Tuesday, 11 September 2007, an incident occurred around 7.00 AM at the off ramp of Voorburg in the Netherlands. This location is marked with a star in Figure 2.9a. Figure 2.9 a-f are based on traffic counts. From Figure 2.9a, it can be seen that at 7.15 AM the congestion spills back over Prins Clausplein ─ a big intersection of the A12 and the A4. In the period thereafter, the congestion spills back over a large part of the A12 and onto the A13. At 8.00 AM, the A13 is completely blocked and the traffic on the A12 has come to a complete standstill up to Gouda (the crossing between the A12 and the A20). At 8.45 AM, the traffic at the A13 also came to a complete standstill. Of course, this situation is made worse by the fact that the accident happened just before the start of the regular peak period. In the Netherlands, the peak period lasts on average until about 9.00 AM; but, in this case, the network as shown

[14] Based on (KIM, 2009)

in the figure remained completely congested until about 9.50 AM. Figure 2.9e shows that congestion starts to decrease on the A12 first. A short while later, the congestion on the A13 starts to decrease, and by 11.00 AM the situation is almost back to normal. This implies that the effects of the incident lasted until four hours after the occurrence of the incident.

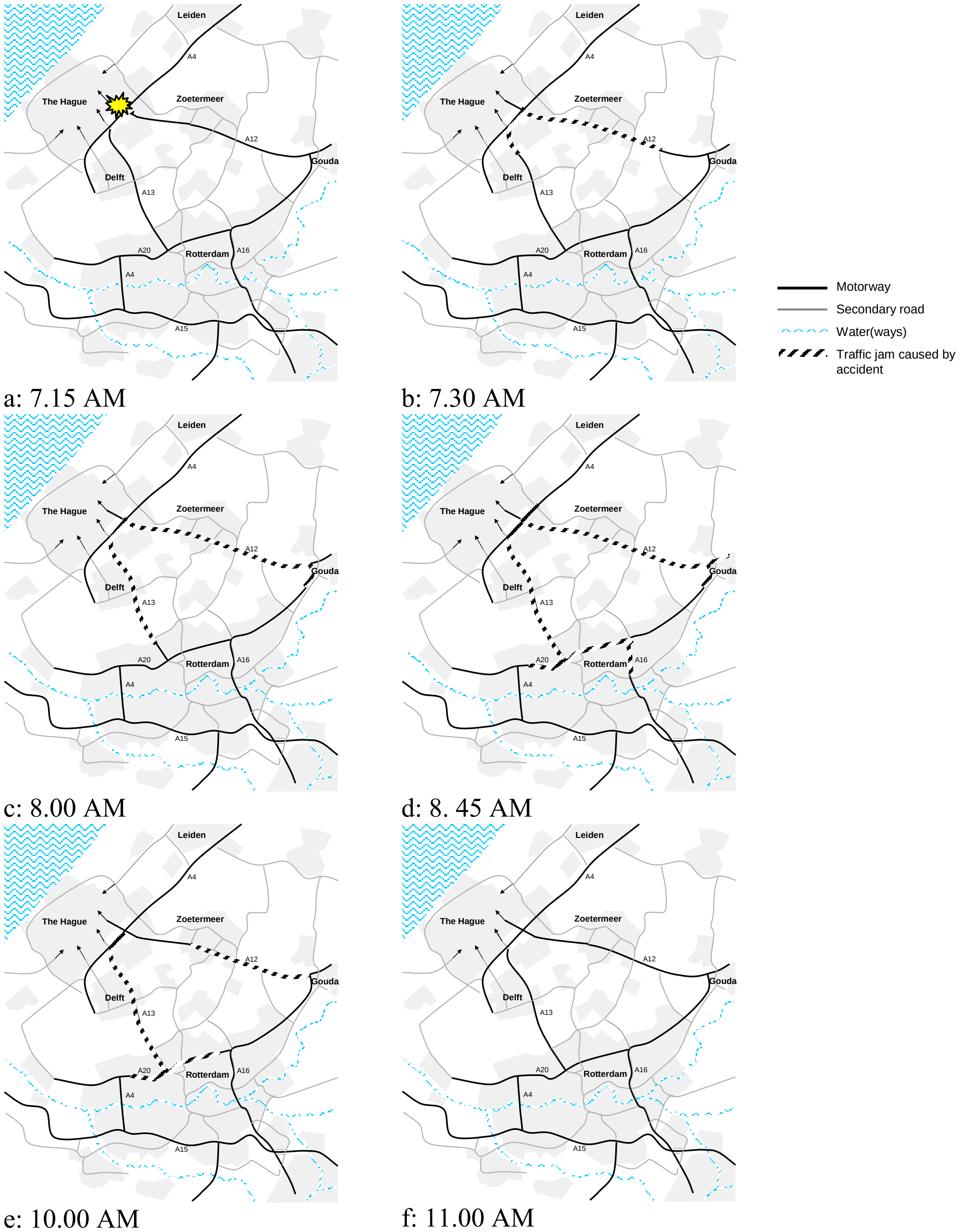
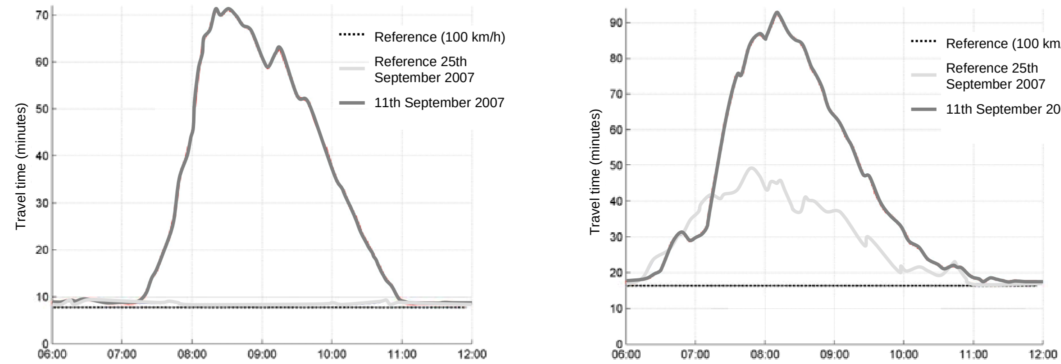


Figure 2.9: Congestion that is caused by an incident on the off ramp of Voorburg[15]

[15] Based on Regiolab

In Figure 2.10a and Figure 2.10b, the travel time that is needed to travel over the complete A13 and the A12 between The Hague and Gouda is shown. The travel time on 11 September 2007, the travel time on a regular Tuesday (without incidents) in September 2007, and the free-flow travel time with a speed of 100 km/hour are shown. These figures show that individual travellers experienced a delay of up to one hour on the A13 and up to 40 minutes on the A12.



a: b:

Figure 2.10: Travel time on the A13 (a) and the A12 (b) on 11 September 2007

The total number of extra vehicle loss hours caused by this incident on the A13 and the A12 was about 10 thousand. This number multiplied by an average value of time of 15 euro/hour results in an economic damage of about 156,000 euro. The real economic damage of this incident on an off ramp is much higher, because several effects are not included. For instance, the extra travel time of taking longer alternative routes is not included in the vehicle loss hours. Delays on the local roads, the costs of emergency services, repair costs, medical costs, and environmental costs are also not included. Finally, unexpected delays may be valued at more than 15 euro/hour. For instance, Eliasson (2004) showed that, based on a large stated preference survey, unexpected delays are valued 3 to 5 times higher than the value of time under regular conditions.

Increasing vulnerability

The example above illustrates that the road network is vulnerable. Small disturbances can cause major disruptions on large parts of the network. Numerous recent examples of accidents with a similar impact could be given. In Figure 2.11 the number of registered accidents and their durations in the Randstad area in the Netherlands in 2007 is shown. About 1750 out of the 12,000 registered accidents have a duration of more than 60 minutes. Besides accidents, about 15,000 car breakdowns and about 5500 objects on the roads were registered. Snelder et al. (2008) showed that the total cost of travel time losses of the accidents with duration of more than 60 minutes was about €1 billion in 2008.

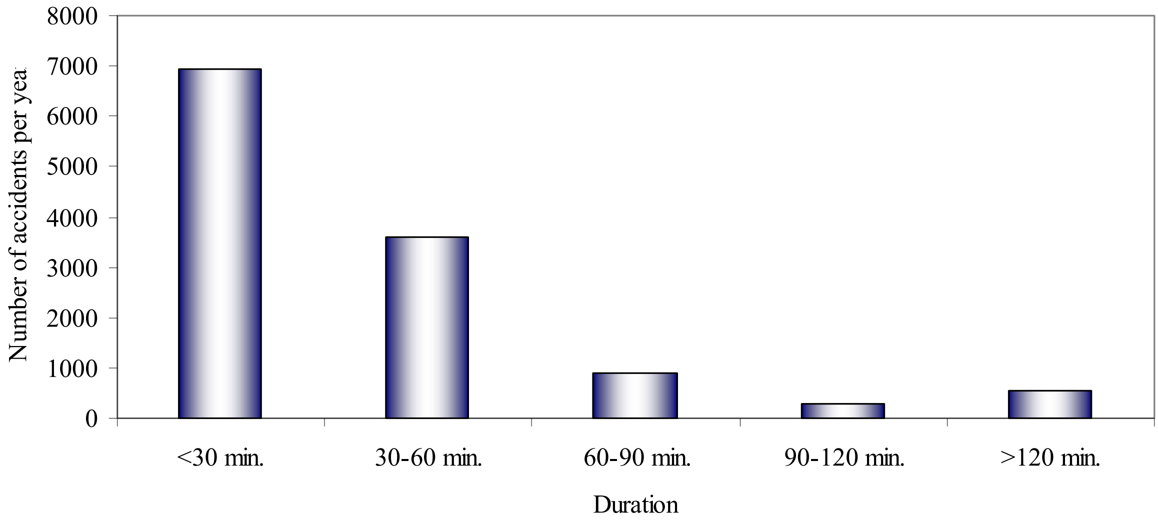


Figure 2.11: Number of accidents in the Randstad in 2007[16]

Because the level of congestion keeps growing, the motorway network, especially in and around major urban areas, is becoming more and more vulnerable to unforeseen disturbances, such as incidents. Furthermore, it will be more difficult to recover from unforeseen disturbances, since spare capacity in the network is being reduced due to more spatially concentrated flows around cities. This spatial concentration of flows is occurring due to concentrated spatial development, new land use policies (bundling) and market-led agglomeration forces (Figure 2.5), combined with an expected widening of congestion periods due to pricing measures and a gradual reorganization of household and working schedules. In Figure 2.12, the reduction of spare capacity in time-space is shown. When the peak period gets longer and more intensive, the spare capacity at each time step *t* is reduced. Furthermore, the spare capacity after *t* becomes reduced. The spare capacity after *t* can be used to recover from an incident.

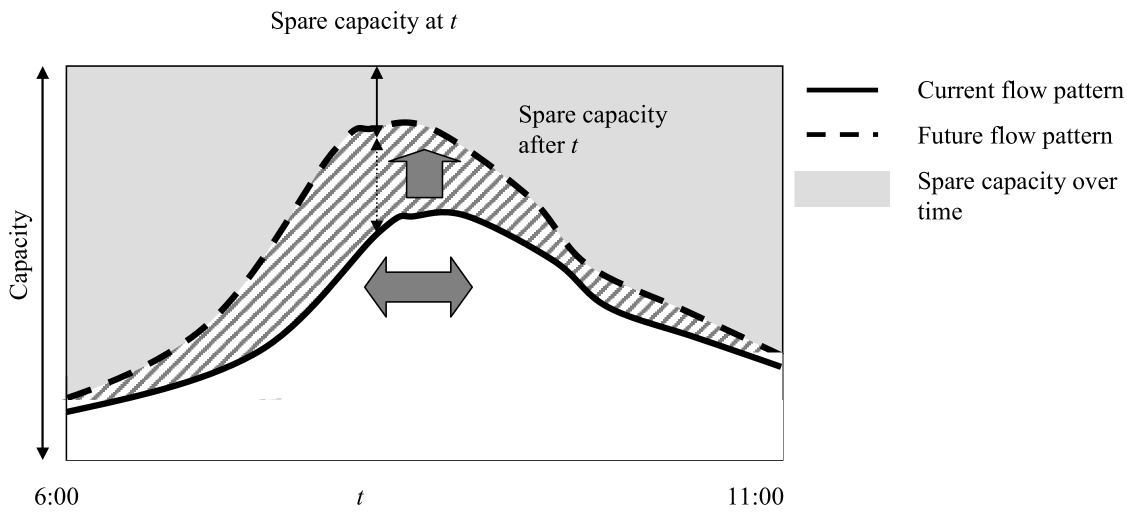


Figure 2.12: Reduction of spare capacity in time-space

[16] Based on “Program monitoring incidents from DVS”

This reduction of spare capacity in time-space leads to increasing risks of sudden drops in network performance, as a result of unforeseen events like incidents. The fact that networks in major urban areas are strongly connected implies both an additional risk (as one link can obstruct another) and an opportunity to mitigate the risk (as spare network capacity can be brought to use).

The example below illustrates how much more vulnerable road networks can become in the future if no measures are taken. Berdica and Mattsson (2007) did a similar study for the Stockholm area by using a static model. They showed, among other things, that there is a faint tendency of non-linear effects from variations in traffic demand. Increasing traffic by 16%, results in more than double the extra travel time caused by increasing traffic by 8%.

Example: The vulnerability of the network of Rotterdam in case of increasing demand
In this example, the importance of having a robust road network is shown by simulating the effects of incidents on the road network of Rotterdam using the dynamic assignment model Indy (see section 4.3.1; Bliemer, 2005; Bliemer, 2007; Yperman, 2007). The example illustrates how the costs of vulnerability can add up in the future if no measures are taken to improve the network. It is unlikely that the costs will become this high, because some network improvements have already been announced.

Rotterdam is the second largest city of the Netherlands. It has about 590,000 inhabitants. The city is surrounded by 4 motorways, which are called 'Rotterdamse ruit'. The road network of Rotterdam that is used in the model is shown in Figure 2.13.

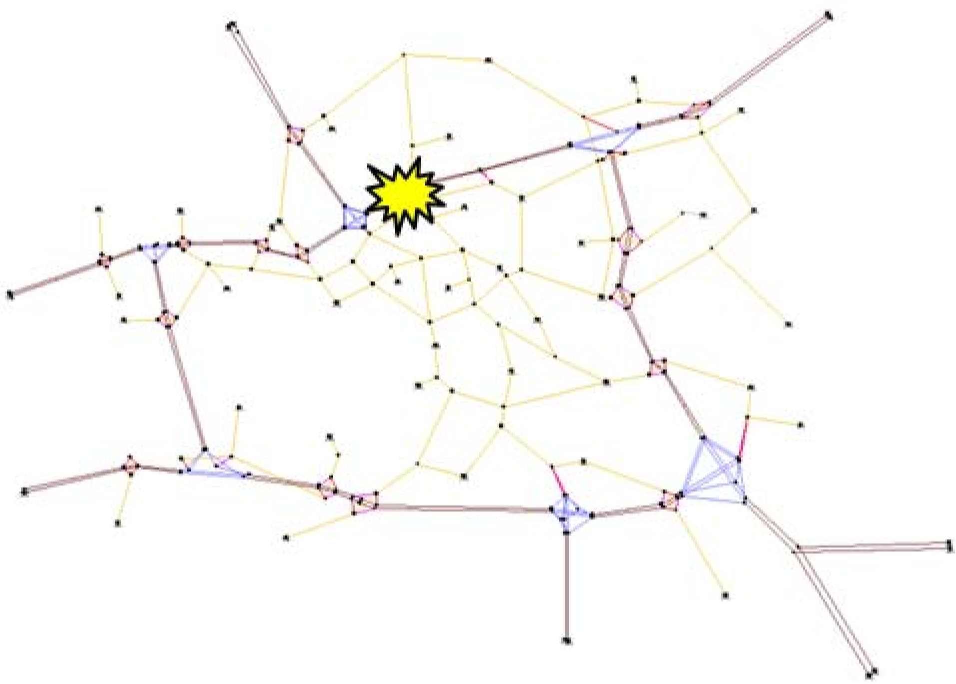


Figure 2.13: Road network of Rotterdam, the Netherlands

The network contains 51 centroïdes. In total there are 1890 OD-pairs (not all combinations of centroïdes are used), and 7674 paths were generated. Furthermore, the network contains 239 nodes and 570 directed links, of which 468 links are regular links and 102 links are feeder links. The total demand period is 24 hours. Figure 2.14 shows the departures per hour. The matrices for the period 0.00 – 12.00 PM are based on the transposed matrix of the demand matrix for the period 0.00 – 12.00 AM. It is likely that, in the future, the peak periods will be longer and more intense. In practice, this will depend on several factors, such as the increase in traffic volumes. However, for sake of simplicity, in our model runs, we used one future demand pattern compared to the existing (2008) demand pattern.

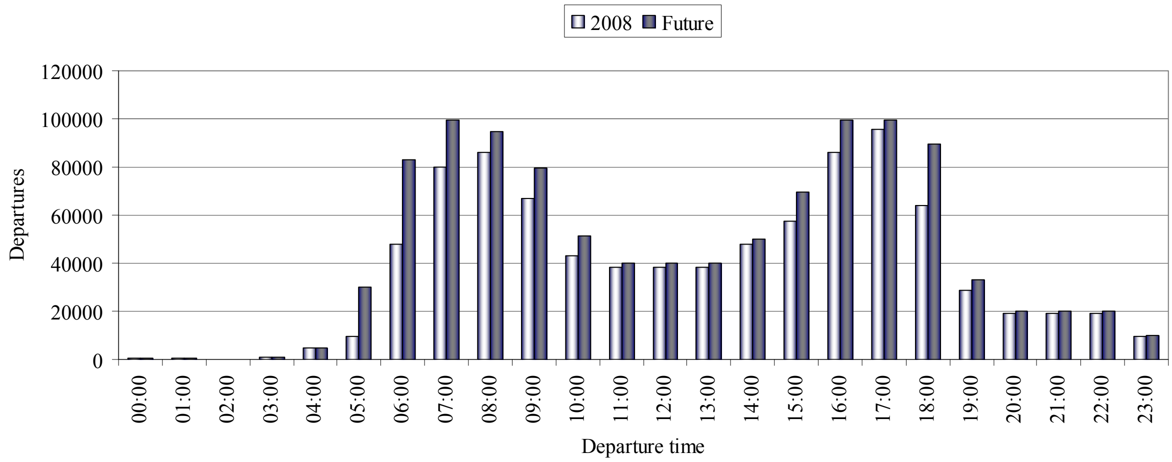


Figure 2.14: Departure time fractions, future vs. 2008

In order to show the vulnerability of the network, an incident is simulated at a location that is usually congested during the peak period, and is therefore likely to be vulnerable. The link is indicated by a star in Figure 2.13. It is an incident on the northern side of the A20, just before the crossing with the A13. The incident occurs at 8.00AM and blocks all three lanes for 1/2 hour. Thereafter, one lane is blocked for 15 minutes. In Indy, this incident is simulated by reducing the number of lanes and the capacity. In the first half hour after the incident, the capacity is reduced from 6600 passenger car units (pcu) per hour to 0 pcu/hour and the number of lanes is reduced from 3 lanes to 0 lanes. In the 15 minutes thereafter, 2 lanes are in use again and the capacity is reduced to 3300 pcu/hour ─ half of the capacity under normal circumstances. In the period thereafter, the road is assumed to function normally (capacity 6600 pcu/hour and 3 lanes). In practice, a capacity drop might occur, which would result in a lower capacity. However, this is not considered in this example.

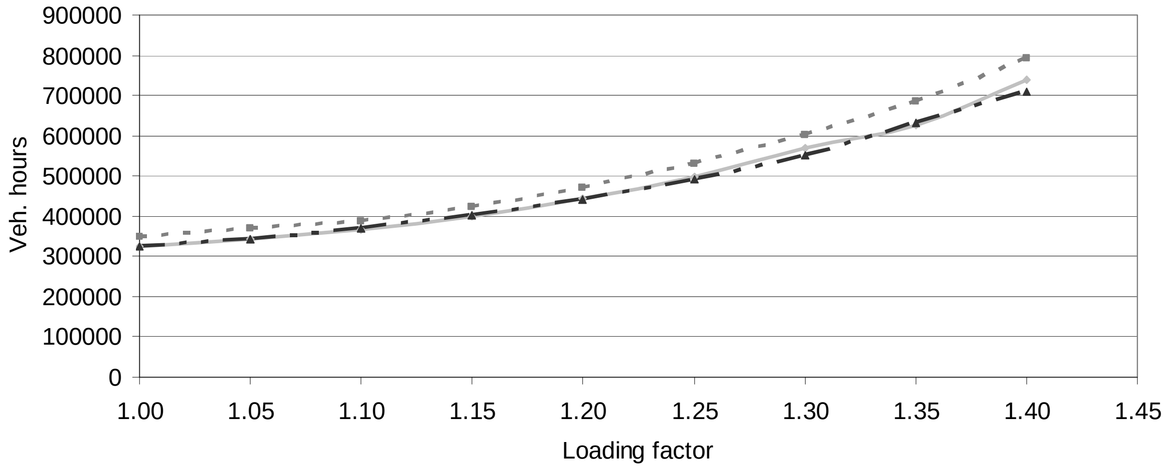
The effects of an incident depend among others on the route choice behaviour of drivers. Not much is known about the percentage of users that deviate from their original routes in incident situations. Therefore, we simulated two extreme cases: no information (fixed route choice) and maximum information (new equilibrium). The case with fixed route choice resembles a situation in which drivers have no information at all about the incident. When they see the incident or the congestion that is caused by the incident, they do not have the opportunity to deviate from their route, they do not have information about the availability and quality of alternative routes, and they therefore stick to their original route or they just do not want to change their route. The other modelling extreme is the situation in which a new equilibrium arises. This situation is similar to the situation in which everybody has complete information about the incident and the alternative routes. Both are unrealistic since, in practice, some drivers will always deviate from their original routes. This can be modelled only by models in which en-route route choice is included, for instance as is done in (Li, 2009). Indy does not have this option. On the other hand, if en-route route choice is modelled, it is most likely to be wrong as well, because not much information is available about the route choice behaviour of people during incidents. Therefore, the decision was made to use the two most extreme scenarios to show what might happen during incidents.

For this case study, a simple calibration procedure was used. This means that we checked whether the congestion locations and moments on which congestion occurs match the average daily situation on the network of Rotterdam. A forecast of the vulnerability of the road network of the future was made by doing 54 model runs. In these runs the incident situation, level of demand, and departure pattern was varied. There are three types of incident situations: a run without an incident, a run with the incident and a fixed route choice, and a run with the incident and a new equilibrium. Nine demand levels were modelled, by multiplying the OD-matrix by the following factors: 1.00, 1.05, 1.10, 1.15, 1.20, 1.25, 1.30, 1.35 and 1.40. Finally, we used two demand patterns, which are shown in Figure 2.14.

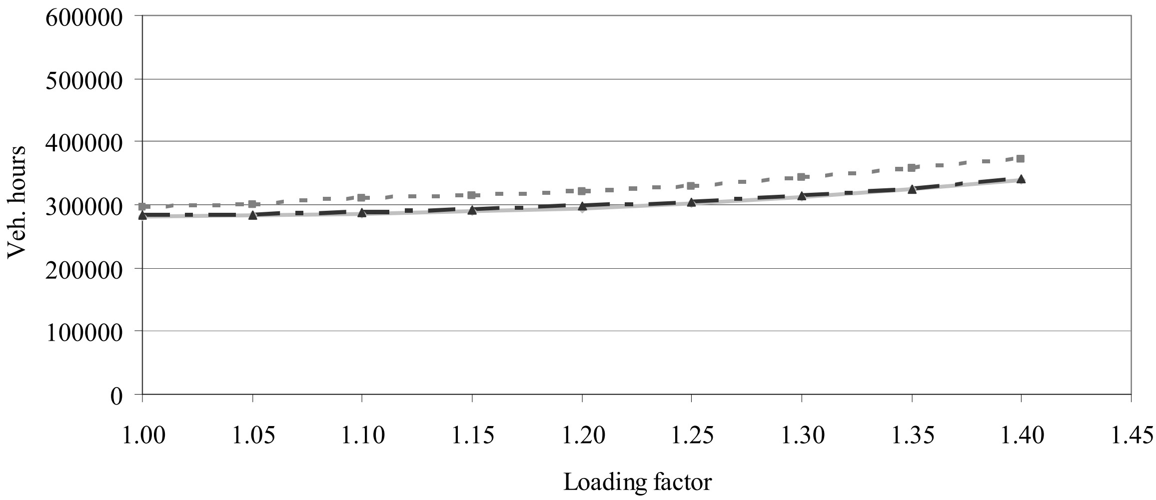
Forecasts of the total future travel time are presented in Figure 2.15a for the base year demand pattern. This figure shows that the total travel time first increases slowly as the total demand increases (loading factor). Thereafter, from the loading factor 1.2 onwards, the total travel time increases faster. When the demand increases by 40%, the total travel time increases by 68%. The additional increase of 28 percentage points (68% - 40%) is caused by the congestion effects. In Figure 2.15c, the same is shown. However, in this figure the total travel time is corrected for the increase in demand (total travel time/loading factor). Therefore, this figure shows the effects of extra congestion caused by an increase in demand and by incidents. In these figures, the distance between the lines of the incident situations (with fixed route choice and with an equilibrium route choice) and the situation without an incident indicate the number of vehicle loss hours that are caused by the incident. From the figure, it can be seen that, in general, the distances between the lines increase when the demand increases. This implies that the impact of an incident in heavily used networks is larger than in networks that are used less intensively. In the case of fixed route choice, an incident could cause a delay of 5%-10% of the total travel time of 1 day. This is quite a large number, given the fact that the duration of the incident was only 45 minutes. In the case of equilibrium route choice during the incident, the total travel time increases by only 0%-1%.

For the future demand pattern, Figures 2.15b and 2.15d show that when the demand pattern changes (the peak period gets longer and more intense), the effects are even larger compared to the situation with the demand pattern of 2008. In this case, when the demand increases by 40%, the total travel time increase by 128%. In absolute numbers, the vehicle loss hours in case of an incident are much higher compared to the base year demand pattern. However, the vehicle loss hours expressed relative to the total travel time without an incident are less remarkable. In the case of fixed route choice, the change in vehicle loss hours is +4% to +8%, and in the case with equilibrium route choice the change is -3% to +1%. This implies that, in the case of equilibrium route choice, the total travel time sometimes decreases when an incident occurs. Theoretically, it is possible that a capacity reduction results in a more efficient usage of the network. This is called the Braess paradox (e.g. Braess et al., 2005). Another explanation could be that an incident on a strategic location could result in an improved traffic flow downstream of the incident, which improves the travel times for certain drivers. However, in this case these results are more likely to be explained by the level of convergence of the model. For the equilibrium runs, five iterations were used in order to keep the computation time within acceptable ranges. This resulted in a fairly good equilibrium, with duality gaps that are in most cases less than 3%. However, this could still result in variations in the outcomes when different equilibrium runs are compared. Therefore, these results should be interpreted with care. Despite this inaccuracy, it can safely be concluded that when drivers have full information about the accident and the possible route alternatives (even before the accident occurred in practice), there is enough spare capacity in the network to make sure that the travel times on average do not change much. For individual drivers, a large delay could still occur. Besides that, many drivers take the routes through the city centre, which is not preferred from a liveability perspective.

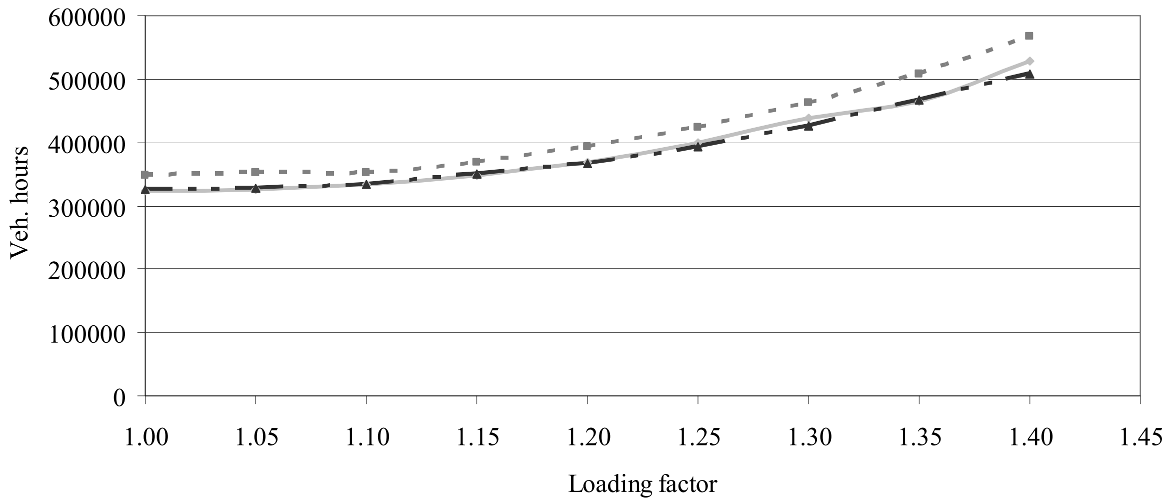
a:



b:



c:



d:

Figure 2.15: Total travel time in the network: a) uncorrected 2008 demand, b) uncorrected future demand, c) corrected 2008 demand, d) corrected future demand pattern

In order to get an impression of the total costs of vulnerability, the vehicle loss hours for the single incident on the road network of Rotterdam were extrapolated to a yearly average for the Netherlands. The vehicle loss hours are multiplied by an average value of time of €15 per hour and an average number of 2000 incidents with a comparable impact that take place on motorways in the peak period in a year (based on Meeuwisssen et al., 2004). In addition, the costs are multiplied by a correction factor for the location. This correction factor indicates the difference in vehicle loss hours between the simulated incident on the ring road of Rotterdam and the average number of vehicle loss hours of incidents on all the links of the network (motorways and other roads that are included in the network) weighted by the expected number of incidents on the link. This factor is determined by simulating an incident on all the links of the network of Rotterdam for demand with loading factor 1. For the situation with fixed route choice, this correction factor is -34.3%, and for the situation with equilibrium route choice, this factor is -14.2%. For the case in which nobody changes routes, the costs of vulnerability increase from €275 million to €900 million if the demand increases by 40%. If the future demand pattern occurs, the costs will vary between €40 million and €900 million.

The costs presented above are only an indication. The real costs could be **lower**, because:
- an improved level of information could result in a better route choice in incident situations in which enough spare capacity is available
- network improvements could lead to a lower level of congestion in situations without an incident. If an incident occurs, congestion spreads slower through the network
- the network of Rotterdam is quite congested compared to other regions in the Netherlands. Therefore, Rotterdam is not representative for the whole of the Netherlands. The applied correction factor partly compensates for this.

The costs could also be **higher**, because:
- the value of time in unexpected situations might be higher than €15 per hour
- besides the 2000 incidents with a large impact, many small incidents occur that also add to the costs of vulnerability

- incidents on the local and regional network also cause congestion. These costs have not been included yet.
- the number of incidents is likely to increase with the traffic volumes
- different disturbances at the same time could worsen the travel time losses
- the capacity drop is not considered

Since this network is rather typical for a large part of the motorway network of the Netherlands, the conclusion can most likely be generalized for a large part of the Netherlands. In Snelder et al. (2008), a similar approach has been used for a larger part of the Netherlands (the Randstad). In that case, it was shown that the expected costs of vulnerability in the Netherlands might increase from about €1.2 billion in 2008 to €4.1 billion in 2030 if no measures are taken. The reason that these costs are higher is that congestion that is caused by incidents can spread over a larger part of the network, and therefore causes more delays.

The conclusions presented above are based on the occurrence of a single incident. In the remainder of this thesis we will assess the robustness of a network for different types of incidents on different locations. This enables us to determine an optimal investment strategy to improve the robustness of the network.

Summary: vulnerability of the road network
In this section it was shown that many parts of the Dutch road network are vulnerable because there is not much spare capacity and there are hardly any route alternatives to the motorways in the secondary network. The direct costs of vulnerability are significant. Indicative simulations have shown that the costs of vulnerability in 2008 range between €275 million and €1.2 billion per year. If no measures are taken these costs can increase by 2030 to between €900 million and €4.1 billion per year. The simulations with different route choice settings showed that one of the measures that can be taken to reduce these costs is to provide appropriate information in time to the relevant drivers. However, since this could result in an unwanted increase of traffic that take the roads through city centres, it is necessary and most likely beneficial to improve the network structure as well. Finally, there are many costs of vulnerability that have not been mentioned in this section, such as indirect economic effects that give extra motivation for making the road network more robust.

2.3 Other transport networks[17]

Rail network
Figure 2.16 shows the railway network of the Netherlands. In 2008, the Netherlands had about 2800 kilometres of rail, of which a large part is single-track rail (in each direction). On the single-track rail lines, fast trains cannot pass slow passenger trains and freight trains. The largest part of the railway network is managed by ProRail. Daily, the passenger train operator (NS) operates about 4800 train services and transports about 1.1 million people over one of the busiest railway networks in Europe. There are about 380 stations (www.ns.nl). As opposed to road networks, timetables play an important role in the rail network. These timetables show how often trains travel on which corridors.

[17] This section is based on www.prorail.nl and the Veiligheidsbalans 2007, http://en.wikibooks.org/wiki/Embedded_Control_Systems_Design/Aviation.

Figure 2.16: Railway network in the Netherlands, showing number of tracks[18]

ProRail indicates that about 3000 calamities occur annually on the rail network. A *calamity* is a situation in which the train services on a certain corridor are disturbed for more than half an hour, and trains have a delay of 5 minutes or more. The following disturbances can occur on the railways: disturbances of railroad switches, broken rails (34 times in 2006), material problems with for instance locomotives and doors that do not close, objects on the rail, accidents: crashes on road crossings (47 times in 2006), crashes between two trains (4 to 5 per year), suicide attempts (about 180 per year), accidents with rail workers, people that walk on or next to the rails, weather conditions: strokes of lightening, freezing rain (a problem once every 3 years), trees on the rails, and snow.

At the railways a lot of preventive measures are taken, like putting fences next to the rail line in order to prevent people from walking on the rails. Furthermore, preventive measures are taken against bad weather conditions and slippery rails as a result of leaves on the rails in autumn.

In case of disturbances, some trains (like the Thalys and freight trains) can take alternative routes. Therefore, there is some redundancy in the system. However, in most cases it is not important that the trains can take an alternative route, but it is more important that the traveller can take an alternative route. For the traveller, this is inconvenient, because the detours are often much longer than the original route and he has to make a transfer at one or more stations, which causes delays and waiting times. Furthermore, in case of disturbances, the NS sometimes uses buses to transport travellers, which causes delays as well.

The fact that corridors are used is a way to prevent all trains from being delayed in case of a disturbance. Trains that do not have to pass by the location of the disturbance are not delayed. This can be seen as a form of compartmentalization. However, it can take a long time before the train services are back to normal on the disturbed corridor. Finally, a lot of incident

[18] source: www.wikipedia.nl

management scenarios are available that can be used to minimize the delays caused by disturbances.

Comparison of the rail network with the road network
The fact that trains have to make use of rails makes them less flexible than cars and trucks. However, the traveller is not restricted to his vehicle/train. If a train is damaged, the traveller can use other train services if they are offered. This makes the traveller slightly more flexible. Furthermore, the railway network is less fine-meshed than the road network. As a consequence, detours that have to be taken by trains or travellers in case of disturbances are longer and more inconvenient than on the road network. Furthermore, trains drive on certain corridors. Therefore, the traveller sometimes has to take multiple trains to reach his/her destination. This causes some inconvenience because the traveller has to make transfers. On the other hand, this reduces the chance that disturbances on one part of the network cause delays on other parts of the network. In road networks, these kinds of effects are more likely to occur. Since the consequences of rail accidents can be quite large, more preventive measures are taken compared to road networks. At the same time, it is probably easier to take preventive measures on railway networks, since there are fewer rail links, and the surface of a rail is much smaller than the surface of a road.

Inland waterway network
Figure 2.17 shows the inland waterway network of the Netherlands. This network contains 500 kilometres of main broad transport waterways, which are used for transport between the ports of Rotterdam and Amsterdam and the hinterlands of Germany and Belgium. The Netherlands also has main waterways with a total length of nearly 900 kilometres which interconnect provinces and are used for national and international transport purposes. The other waterways constitute a network within the Dutch provinces (Visser, 2009).

Figure 2.17: Dutch inland waterways network 2009[19]

[19] source: (Visser, 2009)

The *Nota Mobiliteit* (Dutch Ministry of Transport, Public Works and Water Management and VROM, 2004) indicates that dams, locks, bridges and waterways have not been well maintained. This causes unreliability in the inland waterways network. Furthermore, in the future the capacity of certain locks and bridges might not be enough. Another possible future cause of extra delay and uncertainty in travel times is the fact that ship sizes are increasing, which means that the number of ships that can go through a lock is decreasing. Finally, in the future, climate change may result in more high- and low-water levels, which could restrict inland waterway traffic.

The *Veiligheidsbalans 2007* (Inspectie Verkeer en Waterstaat and Rijkswaterstaat, 2007) indicates that, in 2006, 731 accidents were registered on the inland waterways, of which 118 were significant (at least one victim, and/or infrastructure or ship damage of more than €50,000, and/or a damaged load of 10 tonnes or 1 container or more, and/or environmental damage of phase 1, 2 or 3, and/or a complete blocking of an inland waterway). Most accidents are caused by a crash between two ships or crashes between a ship and the infrastructure.

In general the impact of accident on the inland waterway network is not large, because often other ships can sail around, since the waterway is broad enough. If the accident occurs on a narrow waterway, the consequences are larger, because taking a detour is usually not feasible. Compared to the road network, the inland waterway network is much less fine-meshed. Because of this, and the fact that the speed of ships is much lower than the speed of cars, taking detours takes a very long time. Choosing other modes is not a good option either. Depending on the ship type, ships can carry volumes between 350 and 11,000 tonnes. Container ships can carry between 24 and 470 TEU[20]. A train can carry between 800 and 2200 tonnes, and trucks can only carry between 3.5 and 50 tonnes (or a few TEU). This implies that multiple trains and trucks have to be used in order to carry a shipload. Furthermore, it takes a lot of time to arrange this, and will therefore not happen very often.

Comparison of the inland waterway network with the road network.

In general, the number of accidents on the inland waterways network is lower than on the road network. The impact is usually lower as well, because the inland waterways are often wide enough to be able to bypass the accident. On road networks there is often not enough spare capacity to bypass the incident location without experiencing a delay. If bypassing is not possible, the impact is larger, because it takes a long time to make use of detours. The number of ships that experience this delay is lower than the number of cars and trucks that experience a delay due to a big accident on the road.

Aviation networks

Aviation networks are three dimensional. The networks consist of airports (nodes) and there are no physical links in the network. The airplanes can fly, restricted by some rules and given some corridors, at different heights in the air and in different horizontal directions as well. If a plain crashes on an ‘air-link’ this has hardly any implications for other flights. However, if an airport is closed (for instance due to an accident), the consequences are large. This is especially the case for large hub-airports. Furthermore, rerouting has great implications for travellers, since airports are usually far apart.

Opposed to the road network, for civil aviation the aircraft, the air traffic control, the pilots and crew, and the airports are more important than the links. Passengers board an aircraft at

[20] TEU is an indicator for the size of containers; it stands for Twenty feet Equivalent Unit.

an airport. The aircraft, which is controlled by pilots, transports them in the air to another airport, and air traffic control is used for guidance of the aircraft.

There are a lot of embedded control systems in an airplane. One of them is the flight control system, which controls the flight trajectory and the stability of the airplane. The actuators of this system are the engines and the movable devices of the main wing and tail. They are actuated by the pilot and by the embedded control system of the airplane.

An *aviation accident* is defined in the Convention on International Civil Aviation Annex 13 as an occurrence associated with the operation of an aircraft that takes place between the time any person boards the aircraft with the intention of flight and all persons have disembarked, in which a person is fatally or seriously injured, the aircraft sustains damage or structural failure, and/or the aircraft is missing or is completely inaccessible. An *aviation incident* is defined as an occurrence other than an accident, associated with the operation of an aircraft, that affects or could affect the safety of operations. An accident in which the damage to the aircraft is such that it must be written off, or in which the plane is destroyed, is called a *hull loss accident.*

According to the Aircraft Crashes Record Office, in the period 1999 – 2008, between 136 and 198 accidents occurred annually involving aircraft that are capable of carrying more than six passengers, not including helicopters, balloons, or fighter airplanes, which resulted in between 766 and 1567 fatalities. As can be seen from Table 2.1, air transport is the safest mode of transport per kilometre. However, when measured by journeys, buses are the safest form of transportation, and the number of air transport fatalities per journey is surpassed by only that of bicycle and motorcycle journeys[21].

Table 2.1: Risk per mode[21]

	Deaths per billion journeys	**Deaths per billion hours**	**Deaths per billion kilometres**
Bus	4.3	11.1	0.4
Rail	20	30	0.6
Van	20	60	1.2
Car	40	130	3.1
Foot	40	220	54.2
Water	90	50	2.6
Air	117	30.8	0.05
Bicycle	170	550	44.6
Motorcycle	1640	4840	108.9

There are many factors that could cause accidents or incidents, such as lightening, ice and snow, engine failure, metal fatigue, fire, bird strike, ground damage, volcanic ash, and human factors, including pilot errors and terrorism.

There are many safety improvement initiatives in aviation involving regulators, manufacturers, operators, professional unions, research organizations, and international organizations. The

[21] http://en.wikipedia.org/wiki/Air_safety#cite_note-22

major safety initiatives worldwide are the Commercial Aviation Safety Team (CAST) in the US and the European Strategic Safety Initiative (ESSI).

For aircrafts, dependability (reliability, maintainability, and availability) is very important for safety. The flight control system must be available at all times, unlike cars or trucks where simply stopping is an option. A lot of systems include redundancy, in order to increase their reliability. Airplanes also undergo maintenance frequently.

Of all the different system levels, air traffic control is at the top level. This is a special level, because it does not exist in other transport modes, like road transport, where there is no global communication among vehicles. The primary goal of air traffic control is collision prevention, which requires communication and radar systems. To provide redundancy, most aircrafts are equipped with a Traffic alert and Collision Avoidance System (TCAS).

In larger airplanes with autopilots, the pilot can be seen as a redundant system to fly the plane if the autopilot fails. Even the pilots themselves can replace each other. Even though they are supposed to do different tasks, all pilots can fly the plane in case of an emergency. The autopilot is a system that serves to diminish the workload of the pilot. After long hours in the air, the pilot must be able to concentrate enough to be able to land safely. Another reason to use a control system in larger planes is the fact that the control surfaces of these planes require large control forces, which are impossible for a human pilot. Furthermore, a computer controlled system is more accurate. Sensors will produce a signal according to the direction the aircraft is flying. The control system will compare this information with the desired direction.

In the air industry there are ‘fail-safe’ and ‘safe-fail’ strategies (Möller and Hansson, 2008). Fail-safe systems or networks, remain functioning if a part of the system fails. If for example an engine of an aircraft fails, the aircraft can still fly. In safe-fail networks or systems, failure of the system is accepted as long as it happens in a safe way or it results in an acceptable performance reduction. An example of this is the landing gear of an aircraft. There is no second landing gear available. However, the aircraft is designed in such a way that it can make relatively safe crash landings. In fail-safe strategies there is a lot of redundancy in the system, as is explained above.

Comparison of the aviation network with the road network.
Opposed to road networks (mainly 2D), aviation networks are three dimensional. For road networks the links (the roads) are very important, whereas the air network does not have any physical links and only a limited number of nodes, i.e. airports. If there is an accident along an “air link”, the crashed plane does not block the network, unless it happens in a node and causes a closed airport. However, the consequences for the passengers that are involved in a plane crash are in most cases, much larger than the consequences for the passengers that are involved in a car crash. Futhermore, more passengers are involved at the same time. Therefore, it is more important to take preventive measures for airplains that it is to take preventive measures for cars.

In general, there is much more redundancy in aircrafts than in cars. There is, for instance, redundancy in parts of the system (for instance the engine), in the pilots (autopilot, pilot, and co-pilot). Furthermore, pilots are better trained than car drivers, a lot of control systems are available, and airplanes are maintained on a regular basis. This is all meant to prevent accidents or incidents from occurring.

2.4 Other non-transport networks

In order to learn which elements make a network robust, networks other than passenger/freight transport networks can be considered. In Appendix A, it is explained which elements make the cardiovascular system and the nervous system, both networks within the human body, robust. Furthermore, telecommunication networks are considered in Appendix A. Below the most important findings from Appendix A are summarized.

Comparison of the cardiovascular system with the road network

- If we compare the different types of blood vessels to the different road types, the most appropriate analogy would be that the capillaries are local roads, the veins are the secondary network, and the arteries are the motorways. With respect to flow speeds, this analogy is appropriate. However, with respect to the order in which the types are used, there is not a 100% match. This would for instance imply that blood flows from motorways (arteries) to local road (capillaries) to secondary roads (veins) and back again to motorways, etc. In road networks the order is usually local road, secondary road, motorway, secondary road, local road. However, sometimes some road types are skipped or used in another order. In the cardiovascular system, the order is always maintained.
- The road network is designed to transport passengers and freight in cars and trucks (and other modes of transport). These passengers and freight products have specific origins and destinations whereas blood (which is transported within the cardiovascular system) does not have a specific origin and destination. The only restriction is that all parts of the body should be oxygenated by blood cells (it is not important which blood cell oxygenates which part of the body), which implies that the blood should pass by the lungs and all parts of the body. The fact that blood cells do not have a specific origin and destination makes it easier to circulate the blood through different routes. Furthermore, blood cells are flexible and can stick to each other, whereas cars and trucks cannot bump into each other and cannot reshape in order to pass by small passages. This is also a reason why transport of blood is easier, and more homogeneous flows occur. Besides this, human behaviour plays an important role in driving cars and trucks. This is not the case with blood cells, which also make it easier to obtain system optimal flows instead of user optimal flows.
- In the cardiovascular system, congestion is less likely to occur than in road networks. This is not only because of the factors described above, but also because of the fact the blood is circulated under high pressure, which forces the blood to find its own way past bottlenecks. As a consequence, spillback effects that occur on the road when queues spill back to other roads do not occur in the cardiovascular system.
- If disturbances occur in the cardiovascular system, there is an immediate response. The baroreceptors in the blood vessels immediately detect changes in pressure and an automatic nervous signal is send to the blood vessels to adjust their diameter in order to correct the pressure. The heart rate can be adjusted as well. In road networks the detection of incidents is not instantaneous, but can go relatively quickly by means of monitoring systems and phone calls. The response takes more time. Cars and trucks have to be removed from the road, and it is not possible to increase the road capacity (of alternative roads) as is done in blood vessels. There are some mechanisms in the road network by which the capacity of certain routes could be increased ─ e.g., by changing the traffic signal priorities and by opening roads whose use is discouraged by barriers that are raised and lowered according to the time of day. However, in practice these systems are not (or hardly ever) adjusted in case of disruptions. Furthermore, it is not possible to reroute all vehicles optimally as is done in veins, since humans cannot be guided in the way that blood cells can.

- When partial blockings of blood vessels occur, buffers are created by extending the diameter of the blood vessels. In road networks, buffers cannot be flexibly created.
- Bypasses can be created surgically (and new blood vessels can be created by the human body itself under specific circumstances), which is by far less expensive than creating new bypassing roads.

Comparison of the nervous system with the road network
Compared to the road network, the nervous system has no route alternatives, which makes the system more vulnerable than the road network. On the other hand, the nervous system is much better protected than the road network, which makes it more robust.

Comparison of telecommunication networks with the road network
Compared to the road network, telecommunication networks are often made more redundant. A lot of spare capacity is available. Furthermore, this spare capacity can more easily be used, since information packages can easily be rerouted, which is much more difficult compared to rerouting drivers. It is easier to make a telecommunication network redundant than to make a road network redundant, because the construction costs are much lower. Laying down a cable costs on average about €25–35 thousand per kilometre, whereas constructing a road costs on average about €8–10 million per lane kilometre.

Furthermore, the speeds at which packages are sent, disturbances are recognized, and packages are rerouted are much higher than in the road network. Where a disturbance on the road network can easily cause a delay of several minutes up to more than an hour for drivers, disturbances on the Internet usually do not lead to more than 30 seconds delay for users. Of course, in exceptional situations the delay on the Internet (mainly at the application side) can also take much longer.

2.5 Summary: the vulnerability of networks

In this chapter, the first research question has been answered: What is the importance of robustness for the road network? Furthermore, the comparison of the road network with other networks showed which elements are likely to make a network robust. This, therefore, contributes the first part of the answer to the fourth research question: Which elements determine the robustness of a road network?

The importance of robustness for the road network
In section 2.1, we explained that drivers increasingly face unexpected delays of more than one hour. This results in costs, either because travellers have to take into account a buffer and therefore might arrive early if nothing happens during the trip, or because they arrive late, which causes economic damage (depending on the purpose of the trip). The same is true for freight transport. The question is how large these costs are. The answer to this question illustrates how large the problem of vulnerability is, and suggests the amounts that can be invested in order to make the network more robust. It was shown that the direct costs of vulnerability are significant. Indicative simulations have shown that the costs of vulnerability in 2008 range between €275 million and €1.2 billion. If no measures are taken, these costs might increase by 2030 to between €900 million and €4.1 billion. For individual travellers, accidents can cause unexpected delays of more than one hour. The simulations with different route choice settings showed that one of the measures that can be taken is providing appropriate information in time to the relevant drivers. However, since this could result in an

unwanted increase of the traffic that takes the roads through city centres, it is necessary and most likely beneficial to improve the network structure as well. Finally, there are many other costs of vulnerability, such as indirect economic effects, which give extra motivation for making the road network more robust.

Elements of a robust network
It can be concluded that a network becomes more robust if there are alternative routes available and there is enough spare capacity on all the routes (redundancy). Having this redundancy is not enough; the transported elements must also be rerouted to these alternative routes as fast as possible, which requires some flexibility in the network and some traffic and incident management strategies (resilience). Furthermore, in a robust network, spillback effects must be reduced to a minimum level (compartmentalization). Finally, preventive measures make a network more robust as well. In the next chapter these elements are worked out in more detail.

The conclusion that prevention, redundancy, compartmentalization, flexibility and resilience are elements that make a network robust is supported by the fact that these elements are also found in other transport and non-transport network. The most important lessons learned from the comparison between the road network and other networks are:

- Compared to other non-transport networks, the road network is unique, since passengers are transported and cars are controlled by humans. Humans are not interchangeable, whereas blood cells are. Therefore, human behaviour has to be considered in any robustness analysis of the road network.
- Building roads is more expensive than building infrastructure for most of the other non-transport networks. Not only is it more expensive, but it also has a higher impact on the environment, which makes constructing road networks extra complex. This explains why it is more difficult to make road networks redundant than other non-transport networks. In general, it can be concluded that the competition for space in land use is a main factor that makes it complex to make transport network robust.
- The road network is more susceptible to disturbances than most other networks, because it is a less controlled network. Therefore, in other networks, such as the railway network and the nervous system, more structure related preventive measures can be taken (and are taken).
- The road network is less flexible and less resilient than, for instance, the cardiovascular system and the Internet. This is mainly because blood and information packages can more easily be rerouted and because there is an instantaneous response to disturbances. The diameter of the blood vessels of the cardiovascular system is flexible enough to cope with sudden changes in the volumes transported. Compared to the rail network, the road network offers more flexibility, because it is easier for cars and, therewith for passengers, to switch to other routes (if they are available).
- Spillback effects are more likely to occur in road networks than in other networks, because they are used more heavily than most of the other networks and because rerouting is more difficult than in most of the other networks. This is probably a result of a lack of alternative routes in some parts of the network and limited access points to alternative routes that are available, a lack of accurate information about the available alternative routes, and the fact that not all drivers are willing to take alternatives routes.

3 Definitions and indicators for road network robustness

3.1 Introduction

In the previous chapter it was shown that the road network is vulnerable and will become more vulnerable in the future if no measures are taken. Furthermore, a comparison was made with other networks. This comparison is input for explaining the concept of robustness in relation to the road network in this chapter. Questions like "What is robustness?", "Which elements make a network robust", "How is robustness related to reliability?" and "Which indicators are used for robustness?" will be answered. By doing this, the chapter gives policy makers and transportation analysts a common framework to discuss issues that are related to road network robustness and vulnerability.

In section 3.2, we define the terms reliability and robustness and explain the relationship between them. Section 3.3 describes the disturbances against which the road network can be made robust. Section 3.4 explains the elements that make a road network robust. In section 3.5, we make choices among the different indicators for robustness. Section 3.6 summarizes the chapter's results.

3.2 Relation between reliability and robustness

Robustness and reliability are often mentioned at the same time. This is because there is a strong relation between both concepts. However, they are not the same. In this section the relation between reliability and robustness and the differences between them are explained. First reliability and robustness are defined.

The most accepted definition of network reliability is given by both Billington and Allan (1992) and Wakabayashi and Iida (1992):

Reliability is the probability of a road network performing its proposed service level adequately for the period of time intended under the operating conditions encountered.

Reliability is often divided into three categories: connectivity reliability, capacity reliability, and travel time reliability:

- *Connectivity reliability* relates to the probability that the network nodes remain connected (Chen et al., 2002). Terminal reliability can be seen as a special case of connectivity reliability. It concerns the existence of a path between a specific origin-destination (OD) pair (Iida and Wakabayashi, 1989). Connectivity reliability has also been described by many others, including Nicholson et al. (2001), Berdica (2002), and Clark and Watling (2005).
- *Capacity reliability* is defined as the probability that the maximum network capacity is greater than or equal to a specified demand level when arc capacity is subject to random variations (Chen et al., 2002).
- *Travel time reliability* relates to the probability that a trip between a given OD pair can be successfully made within a specified interval of time (Chen et al., 2002). In contrast with terminal reliability, travel time reliability identifies the impact caused by network degradation (Nicholson et al., 2001).

Reliability indicators are often categorized as follows (Texas Transportation Institute and Cambridge Systems Inc., 2006; Lomax et al., 2003):

1. Statistical range measures, such as the standard deviation or x% percentiles.
2. Buffer time measures: the extra percentage travel time due to travel time variability on a trip that a traveller should take into account in order to arrive on time.
3. ‘‘Tardy-trip’’ measures, such as the misery index, which takes the difference between the average travel time of the 20% worst trips and the overall travel time average.
4. Probabilistic measures, such as the probability that a trip can be made in time.

Van Lint et al., (2008) argue that there clearly is inconsistency among these indicators. They showed that there are arguments that are empirically underpinned to prefer measures that include the skew of the travel time distribution.

Most reliability studies rely on the presumption that information on the probability of degradation events is available. However, this information is difficult to obtain in practice, particularly for events that occur rarely. These rare events often result in large-scale disruption with severe social and economic impacts. D’Este and Taylor (2003) suggest that network reliability based on link choice probabilities may not be appropriate to evaluate the impacts of such rare events. Like Luathep et al. (2010) state: “Vulnerability analysis is a proactive approach for identifying weak spots in a network and evaluating the adverse consequences from network failure.” Vulnerability analysis emphasizes the consequences from network degradation or failure and avoids using event probability information in the analysis. Vulnerability analysis is thus vital in strategic transport network planning for dealing with the impacts from natural or malevolent events. The result of vulnerability analysis can be used to recommend proactive remedial countermeasures, e.g. improving the performance of the vulnerable (or weak) links to be more robust, or adding new links to provide more alternative routes.

In comparison with research into reliability, research into robustness and vulnerability is less extensive. The terms robustness and vulnerability have a strong relation, but they are actually each other's opposites. Vulnerability describes the weakness of a network and robustness describes the strength of a network. Berdica (2002) has done leading research into road vulnerability. She defines vulnerability in the following way: "*Vulnerability* in the road transportation system is a susceptibility to incidents that can result in considerable reduction in road network serviceability." In this definition the *serviceability* of a link/route/road network describes the possibility to use that link/route/road network during a given time period. Others also describe the vulnerability of road networks. For example, Taylor and D'Este (2003) relate vulnerability to the degree of accessibility of a given node in the network, where accessibility is expressed as the travel cost needed to access the particular node, comparing optimal and alternative routes or detours. D'Este and Taylor (2003) define vulnerability to be the likelihood of severe adverse consequences if a small number of links (or possibly a single link) is degraded. They distinguish between *connectivity vulnerability* and *access vulnerability*. *Connectivity vulnerability* considers a pair of nodes and the generalised cost of travel between them. If the loss or substantial degradation of one or more network links leads to a substantial increase in the cost, then the connection between those nodes is vulnerable. *Access vulnerability* considers a single node and the overall quality of access from that node to all other parts of the network. A node is vulnerable if the loss of substantial degradation of a small number of links results in a significant reduction in the accessibility of that node, as measuredThis by a standard index of accessibility. It should be noted that the second definition of vulnerability ignores probability; this vulnerability is really a measure of the consequence of degradation (Nicholson et al., 2001).

Based on these definitions we came to the following general definition of robustness that considers the performance of a complete network and allows considering all kinds of disturbances on links, nodes, and routes that lead to a partial degradation of those elements or a complete loss of function of those elements:

Robustness is the extent to which, under pre-specified circumstances, a network is able to maintain the function for which it was originally designed.

Vulnerability is the opposite of robustness. A network that is vulnerable is not robust, and vice versa.

Three elements of the above definition require more explanation:

- *"Function":* The most general function of a road network is to enable trips from origins to destinations. Achieving an adequate road network design requires knowing for which kind of trip a network link or network node has a function. Trips can be categorised by their length, whether they are for passenger or freight transport, and by their purpose. The level at which a network has to function is usually specified by a government organisation or defined by design standards.
- *"Pre-specified circumstances":* In practice the demand and supply are not constant. There are for instance variations in demand and supply due to irregular, unexpected, and exceptional events, such as natural disasters (e.g. earthquakes, hurricanes, floods, landslides), extreme weather, incidents, roadworks, social events (e.g. football matches, big fairs), malicious attacks, and signal failures. In principle, a network can be made robust against all these disturbances. However, in practice, choices have to be made by policy makers about the disturbances on which they want to focus. In section 3.3, the different circumstances are explained in more detail. Furthermore, in that section it is explained that the focus of this thesis is on incidents.

- *"The extent to which":* The definition includes the words "the extent to which". This implies that, in cases of disturbances, the network does not have to function just as well as it would without disturbances. From an economic perspective it is not advisable to make a network 100% robust against all disturbances. The costs of creating such a network would exceed the benefits. However, this does raise the question of the extent to which the network should maintain its function. In *Nota Mobiliteit* (Dutch Ministry of Transport, Public Works and Water Management and the Ministry of VROM, 2004), norms are specified for the reliability of travel times in the Netherlands. These norms focus on average situations and the spread around the average. These norms must be evaluated over multiple days. In the Netherlands there are no norms for robustness. Therefore, we asked a group of experts from the Advisory Council for Transport, Public Works and Water Management to give an indication of delays (additional to delays caused by daily congestion) that would be acceptable from the perspective of a road administrator. In Table 3.1 the average of the acceptable delays is shown for a trip that would take 30 minutes under regular conditions.

Table 3.1: Acceptable increase in travel time as a result of a disturbance (minutes)

	Peak (minutes delay)	**Off-peak (minutes delay)**
Fluctuations in demand		
Extreme peak hours	11	8
Holiday traffic	14	11
Events	14	13
Roadworks	9	6
Weather conditions		
Rain	4	3
Heavy rain	8	5
Fog	19	9
Freezing rain	20	13
Snow	15	11
Incidents		
Road closure	21	16
One motorway lane closed	11	8
One lane closed regional road	16	8
Car breakdown	7	5

Table 3.1 shows that during peak hours higher delays are more acceptable than in off-peak periods. Furthermore, the table shows that it is acceptable that incidents and heavy weather conditions cause higher delays than other conditions. The same is true for events. However, the experts made a distinction between travellers that go to the event and travellers that do not go to the event. For travellers that do not go to the event, the acceptable delay caused by the event is much lower. For roadworks and rain, only relatively short delays are acceptable. This seems remarkable, because during rain there is often much more congestion than during dry conditions. However, since rain occurs often and is more or less predictable, the acceptable delays are relatively low. Some experts did not distinguish among the different disturbances. They said that for the traveller it is irrelevant what causes a delay. Therefore, it should be irrelevant for the road authorities as well. Finally, they said that advance information could make higher delays acceptable.

Besides, the above mentioned acceptable delays, it is important to indicate how long these delays may occur. Table 3.2 shows how long after clearing the road the experts indicate that congestion caused by an incident may exist.

Note: the numbers presented in Table 3.1 and Table 3.2 are expert opinions and not norms that are used in policy.

Table 3.2: Maximum acceptable duration of congestion after the road is cleared (minutes)

Incidents	**Peak**	**Off-peak**
Road closure	27	24
One motorway lane closed	13	11
One lane closed regional road	16	14
Car breakdown	10	9

A framework for reliability and robustness

Figure 3.1 shows the relation between network characteristics and robustness and between robustness and reliable travel times. The numbers in the figure refer to the order in which the figure should be read:

(1) Under regular circumstances (no disturbances), the network performance is determined by the regular demand and supply pattern.

(2) Disturbances, such as accidents, special weather conditions, roadwork, events, and seasonalities, lead to short term variations in demand and supply. These disturbances occur with a certain probability and have a primary effect on the capacity (capacity reduction) and/or the demand (increase or decrease in demand).

(3) The primary effect of the disturbances combined with the regular demand and supply pattern results in a new level of demand and supply.

(4) The effect on the network performance (5) of this new level of demand and supply depends on the robustness of the network, the measures taken by network managers (e.g. information provision) and the response of drivers (e.g. route choice) to that. For instance, in a robust network deviations from the regular demand and supply pattern will result in less variation in travel time compared to a network with a lower robustness level. Robustness can be subdivided in five components (prevention, redundancy, compartmentalization, resilience and flexibility) as is explained in section 3.4. Furthermore, as is explained by Nicholson et al. (2001), if the user has a high level of information (i.e. information is provided well in advance, and route guidance is available once the trip has commenced), the range of available options is greater and the consequence of degradation is reduced.

(6) The variation in travel time that results from the disturbances determines the objective travel time reliability. The variations in travel time can be expressed by a probability density function of travel time. Stability is the degree to which the travel time changes as the intensity rises and/or the capacity falls. Ideally, the change will remain limited; after all, a sudden large increase in travel times should be avoided, if possible.

(7) The way in which reliability (or unreliability) is experienced (8) depends on the characteristics of the driver (e.g. level of risk averseness), the trip purpose, the information that is received by the driver about disturbances, and the alternatives that are available to the driver. If the traveller is informed about delays and if route alternatives are available, then the travel times are less unreliable in the perception of the driver then in a more insecure situation.

Besides these relations, there are some other relations: The redundancy/spare capacity (component of robustness) depends on the regular demand and supply pattern. Furthermore, the response of the drivers to disturbances and measures taken by network managers depend on the characteristics of the driver, the trip purpose, the received information, and the available alternatives. Finally, the available level of information and the driver's route alternatives depend on the information offered by network managers (or information received through other sources, such as navigation systems or the radio) and the robustness of the network.

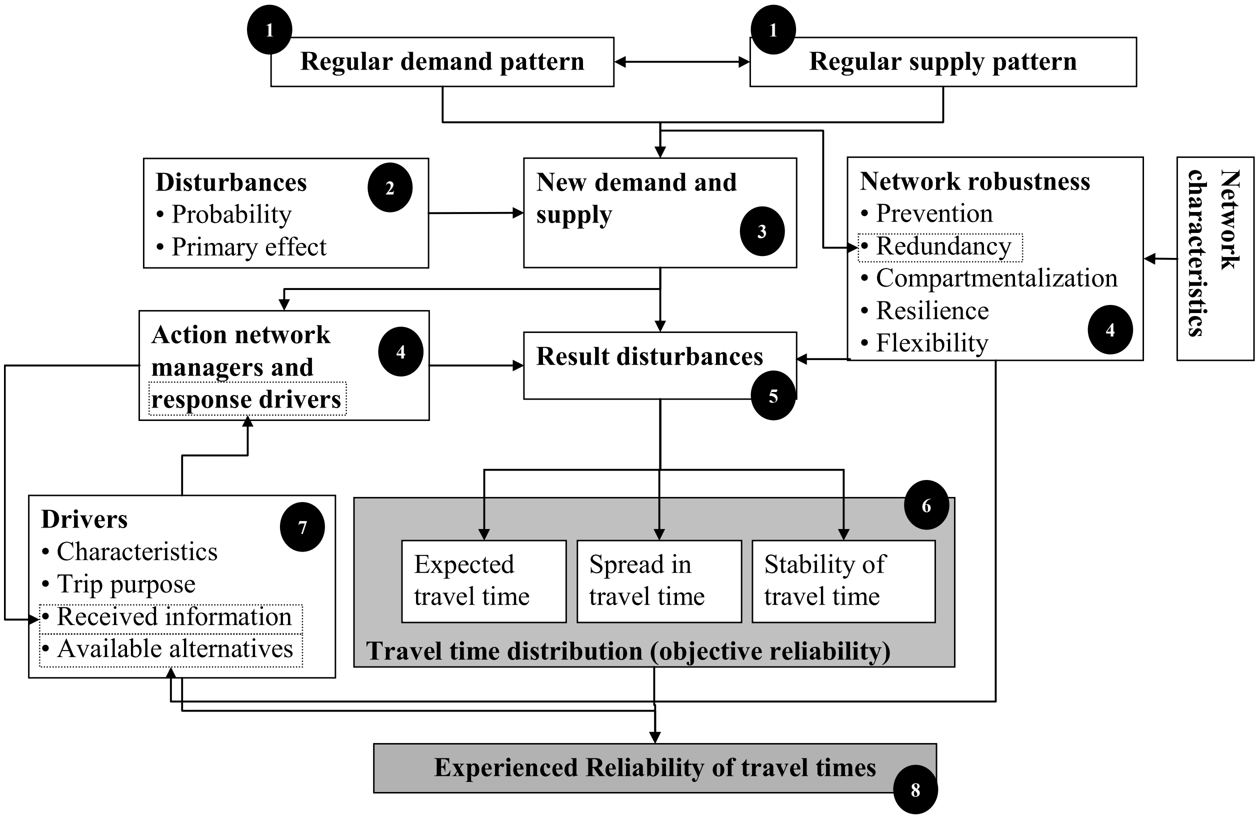


Figure 3.1: Factors that influence the reliability of travel times

In Figure 3.1, the relation between reliability of travel times and a robust road network is illustrated. Although reliability and robustness have a strong relation, they are not identical. In Table 3.1 the differences between the concepts are summarised.

Table 3.1: Differences between reliability and robustness

Reliability	Robustness
User-oriented quality of the transportation system	Characteristic of the system itself
Focus on expected disturbances	Focus on unexpected disturbances
Focus on long period uncertainty	Focus on instantaneous uncertainty
Focus on probability	Focus on effects

It is clear that robustness is a property of the system. By contrast, the reliability of the travel time is something that the traveller experiences. Immers et al. (2004c) expressed this as follows: reliability is a user-oriented quality, while robustness is a characteristic of the system itself. In addition to this distinction, there are three other distinctions:

- Where reliability is concerned, the emphasis lies in disruptions that occur at regular intervals, whereas with robustness the emphasis lies in disruptions that occur unexpectedly and that have a large impact. In the figure below, a travel time distribution is shown in which the foci of reliability and robustness are indicated. A strict distinction cannot be made, because unexpected disturbances can have a large effect (focus robustness) as well as a small effect. The same is true for regularly occurring disturbances (focus reliability). Therefore, the tail of the distribution also has a small weight in the determination of the reliability of travel times. However, because very high travel times do not often occur, they are relatively unimportant for the reliability of travel times. On the other hand, road networks can be made robust against all kinds of disturbances (disturbances with small and large effects). Because robustness focuses on unexpected disturbances, the tail of the distribution function becomes more important (unexpected disturbances have a higher effect than expected disturbances, in general).

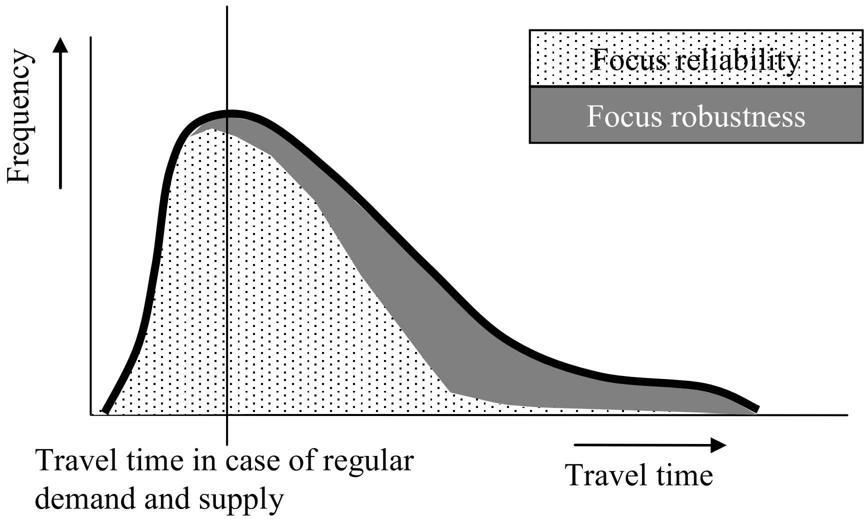


Figure 3.2: Focus reliability and robustness

- Reliability is geared towards an average spread in the travel time and must therefore be determined over a longer period (ranging from several days to a year). With robustness, the emphasis lies on the period in which the effect of a specific disruption is noticeable. It focuses on the impact of single disturbances and not so much on the probabilities that these disturbances occur.
- In the case of reliability, the emphasis lies on the probability that a specific disruption occurs, and with robustness the emphasis lies on the effect, as was also noted by others. Husdal (2004) noted that probability or predictability is a major concern in network reliability studies. The impacts or consequences of disruptions are the focus of vulnerability studies. D'Este and Taylor (2003) note that vulnerability and reliability are two related concepts, but emphasize that network vulnerability relates to network weaknesses and the economic and social consequences of network failure, not so much to the probability of failure. This distinction is, of course, related to the two previous points taken together. This does not mean that the probability of a disruption is unimportant. For the functioning of the whole system, this probability is of great importance. If the probability that disruptions occur can be reduced, this will have a great effect on the average travel time and the reliability of the travel time. In this case, the robustness of the system is of less importance.

Finally, the discussion above gives rise to the following important question, which we then answer:
Does a more robust road network always lead to more reliable travel times (or the other way around)? Husdal (2004) states that there may be a non-linear or reciprocal relationship between robustness and reliability, which implies that vulnerable may not mean non-reliable and that reliable may not mean non-vulnerable. So the answer to the above question is: *no*. In general improving the robustness of road network should lead to more reliable travel times, but some examples can be thought of in which this is not true or at least not true for all travellers. We mention two examples:

- Adding lanes to a road can make this road more reliable[22] because there is more space to drive around an incident and also all kind of small disturbance can be dealt with more easily. This sounds more robust as well, but it is not more robust in the case of a big incident in which, for instance, the complete road is blocked. The fact that the road has more lanes usually leads to more travellers making use of it. This implies that also more travellers experience a delay as a result of the big incident.
- If an alternative route is built or improved, there are more back-up options (redundancy) in the network, which makes the network as a whole more robust for disturbances. However, the original users of the alternative route experience delays if their route is being used as a backup option in case of a disruption, which makes their travel times less reliable.

These counter examples indicate that it is important to evaluate network improvements both on the network level and on the OD-level, and to look at the entire travel time distribution.

3.3 Disturbances

In traffic and transport, many disturbances can occur that result in travel times that deviate from the travel times under regular conditions, such as natural disasters (e.g. earthquakes, hurricanes, floods, landslides), extreme weather, incidents, roadworks, social events (e.g. football matches, big fairs), malicious attacks and signal failures. There are many ways to classify these disturbances. In the literature about reliability, a distinction is often made between recurrent (such as weekday peak hour congestion) and non-recurrent (such as floods and other events of nature) disturbances. The essence of the degree of recurrence is that it provides information about the predictability of the event. In Wilmink et al. (2003), a distinction is made between predictable and non-predictable conditions and between regular and non-regular conditions. In Table 3.3, examples of the different situations are summarized. Examples of regular and predictable situations are morning and evening peak hour congestion. Small incidents can be classified as regular non-predictable disturbances. Examples of non-regular predictable situations are holiday traffic, big events, and extreme weather conditions. Finally, in the class of non-regular non-predictable disturbances, we can mention calamities, big incidents, etc. Some of these disturbances influence the supply and others influence the demand.

[22] Adding lanes does not always have to lead to more reliable travel times. If a road has more lanes, it is likely to be used by more travellers. The probability of a disturbance increases as a result, since this probability depends on the number of vehicle kilometres driven on a link. An increased chance and a decreased effect could, in combination, still lead to more unreliable travel times. A second argument is that the extra lanes could reduce the congestion to a level that is somewhere in between congestion and free-flow on average. This implies that on one day there could be free-flow and on the other day there could be congestion, which is very unreliable. However, a contra argument is that congestion states are also unreliable and that reducing the congestion should, therefore, lead to more reliable travel times.

Table 3.3: Classification of disturbances[23]

	Predictable	Non-predictable
Regular	Peak hour congestion, off peak hours, weekend traffic, bridge openings, small maintenance activities.	Small incidents
Non-regular	Holiday traffic, big events, large maintenance activities, special transport, extreme weather conditions.	Calamities, big accidents, defective infrastructures, crisis, terrorist attacks

Husdal (2004) classifies disturbances by their nature. He describes structure-related vulnerability, nature-related vulnerability, and traffic-related vulnerability. Structure-related or structure-generated vulnerability pertains to the way the road is built and attributes of the road network itself, not only in terms of topology, and connectivity, but also in terms of the physical, body of the road, geometry, width, curvature, gradient, tunnels, bridges, weight restrictions for certain vehicle types, etc. Nature-related or nature-generated vulnerability pertains to attributes of the natural environment, the topography and the terrain that the road traverses, and to nature-given incidents, such as flash floods, avalanches, rock fall, snow and ice, fog, earthquakes, tsunamis, and consequences of climate change, to mention but a few. Traffic-related or traffic-generated vulnerability pertains to attributes describing the generic flow of traffic and attributes resulting in flow decrements, such as daily rush hour and weekend highs, as well as maintenance operations, snow clearing, accident clear-up, and on-going construction works. Besides these unintentional disturbances there are intentional disturbances, such as terrorist attacks.

Furthermore, a distinction can be made about the impact of the disturbances. Most disturbances have a temporarily impact/effect, which can vary from small to large. However, there are also disturbances that have a permanent impact. Another distinction is between within day and between day variations. Finally, the location of the disturbances can vary. Some disturbances have a network-wide effect and others have a local effect.

The above mentioned table matches with risk theory: risk = probability x effect. Regular disturbances have a higher probability than non-regular disturbances. Furthermore, in general, non-regular disturbances have a higher effect than regular disturbances, and non-predictable disturbances have a higher effect than predictable disturbances, because preventive measures can be taken for predictable disturbances. In network design, both the probability and the effect are important. Of course, the disturbances that should get the biggest attention are those with both a high probability and a high effect. However, fortunately there are not so many of those disturbances.

If robustness issues are discussed, it is advisable to clearly indicate against which disturbances a network is to be made robust. As already explained, the robustness of a road network focuses more on effects than on probabilities. Nevertheless, for network design, probabilities are an important factor for investment decisions. Therefore, in this thesis we focus on the class of non-predictable regularly occurring (relatively high probability) disturbances: incidents. We also consider some incidents, such as road closures, which belong in the class of non-regular non-predictable disturbances.

[23] Source: (Wilmink et al., 2003).

The effect of an incident depends upon the level of demand and supply, the structure of the network, the characteristics of the incident itself (number of lanes closed, duration of incident, location of incident, damage caused to road infrastructure, injuries of people, time of incident), the level of the incident and traffic management, the level of information offered, and the response of drivers. Beside these factors, there exist two phenomena that increase the effect of an incident – ‘blocking back’ and ‘capacity drop’ (Immers and Van Koningsbruggen, 2004a):

- Blocking back (spill-back): Congestion can cause vehicles upstream of a bottleneck/incident to come to a standstill. The tail of the queue can also block routes that do not pass the incident location. This phenomenon occurs, for example, if incidents occur at locations downstream of an intersection. The ‘blocking back’ effect can cause sudden large variations in travel time. In extreme cases, it can cause congestion on large parts of the network, and can even cause grid lock (the is a situation in which the tail of the queue reaches the head of the same queue).
- Capacity drop: If the density on the road (vehicles per kilometre) exceeds the critical density, a significant reduction of the maximum capacity can occur. This is called a capacity drop. As a result of a capacity drop, the capacity of a motorway lane can decrease from 2200 ─ 2400 pcu/hour to, for instance, 1800 pcu/hour. This reduced capacity exists as long as there is congestion. As a consequence, the congestion reduces more slowly (Hall and Agyemang-Duah, 1991).

3.4 Elements of robustness

To get a better understanding of robustness we can ask ourselves the question: “What makes a network robust?” or the other way around: “What makes a network vulnerable?” Answering these questions not only clarifies the term robustness, but also gives direction for specifying indicators for robustness and for the measures that need to be taken to make the network more robust. In order to answer these questions, we made an analysis of all the incidents that occurred in the region South Holland in the Netherlands during the period January 1st – April 15th 2007. In total, 3484 incidents were considered. Of these 3484 incidents, 1046 were accidents and the other 2438 were car or truck breakdowns. The incident information (location, start time, end time, date, and incident type) comes from the “Program monitoring incidents from DVS”. We combined a database with incident information with traffic counts on the motorways. From the flow and speed data, the vehicle loss hours can be estimated by using the “piece-wise linear speed-based algorithm (PLSB)” (Van Lint and Van der Zijpp, 2003). Since congestion can also occur when there are no incidents, we computed the vehicle loss hours that occur during four reference days. The reference days were four days in the same week (for weekdays). For Saturdays and Sundays, four weekend days were used as a reference. The vehicle loss hours caused by an incident are computed by subtracting the vehicle loss hours of the reference days from the vehicle loss hours that occurred during the period in which the effects of the incident were noticeable. The vehicle loss hours of the incident and the reference days are computed on the road on which the incident occurred. Of course, the traffic jam can also spill back to other roads. However, these spillback effects are not considered explicitly. Furthermore, the congestion can spill back to local roads. The congestion on these roads is not measured either, since data for these roads was not available. Finally, we cannot be sure that incidents did not occur on the reference days. These are three reasons why the vehicle loss hours of incidents have been underestimated. Nevertheless, the results give a good indication of vulnerable road sections. Jonkers et al. (2009) explain the method that was used in more detail for truck incidents. A similar approach was used for all incident types.

In Figure 3.3, the number of incidents per kilometre that occurred on various road segments in South Holland is displayed. Figure 3.4 shows the average effect, expressed in vehicle hours lost per incident on these road segments. Figure 3.5 shows the total number of vehicle loss hours per kilometre as a result of incidents that occurred on these road segments Although, quite a lot of incidents were included in the analysis, when interpreting these figures it has to be taken into account that the incidents vary in severity. Therefore, some incident locations might by chance be indicated more or less vulnerable than we would have found had we analysed more incidents over a longer period of time.

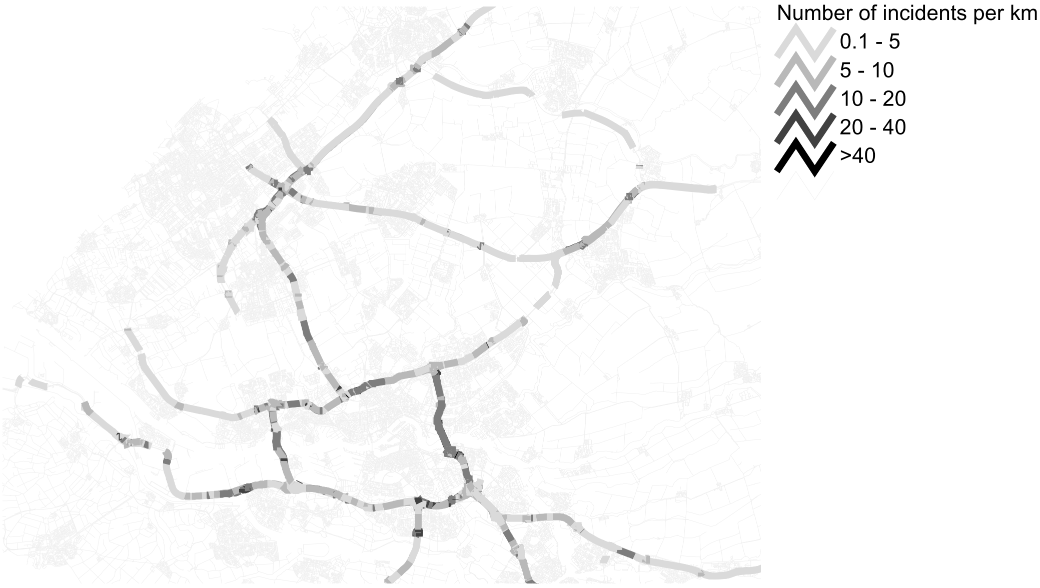


Figure 3.3: Number of incidents per kilometre on the motorway network

Figure 3.4: Vehicle loss hours per incident on the motorway network

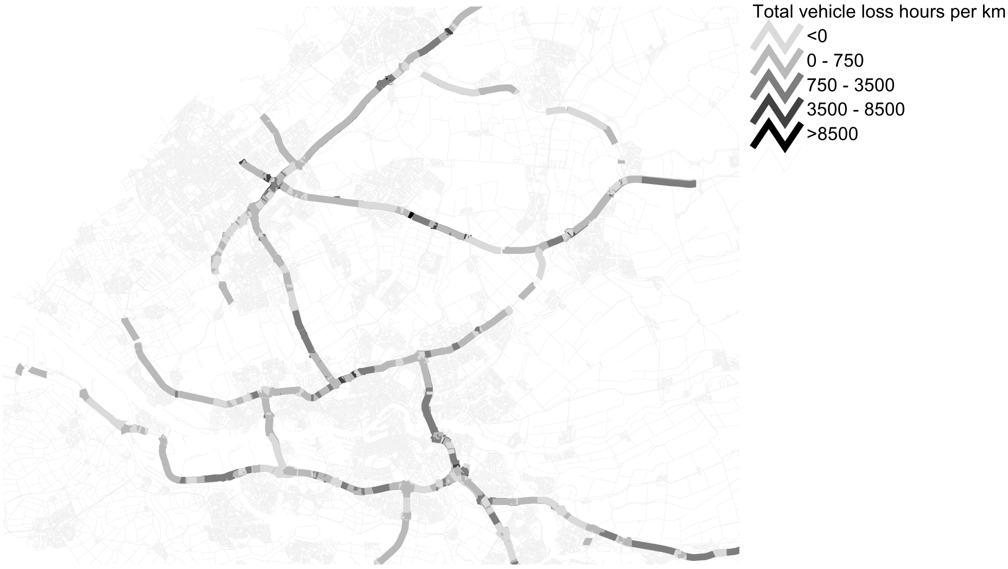


Figure 3.5: Total vehicle loss hours per kilometre on the motorway network

In the mapping of the vehicle loss hours, we had to deal with the following limitations:

- An incident is characterized by a road number and a hectometre number. However, multiple road segments in the GIS-layer can have the same road number and hectometre number. This happens, for instance, on motorways with parallel road structures or at junctions. Therefore, the same incident can be plotted on multiple road segments.
- One road segment can be longer than one hectometre. Therefore, incidents that occurred on different locations may have been plotted assigned to the same road segment.

We combined the figures above with our knowledge about the network (especially the secondary network) and the locations on which regular congestion occurs (Figure 2.7). This resulted in the following hypothesis:

1. Incidents that occur on roads with a high intensity-capacity ratio (I/C-ratio) or a high intensity-spare capacity ratio (I/(C-I)-ratio), and thus with little spare capacity, have a higher effect (expressed in vehicle loss hours) than incidents that occur on roads with low I/C-ratios.
2. Incidents that occur on roads with high flow have a higher effect (expressed in vehicle loss hours) than incidents that occur on roads with lower flow, if the I/C-ratio is equal.
3. Incidents that occur at locations where good alternative routes are available have a lower effect (expressed in vehicle loss hours) than incidents that occur at locations where good route alternatives are not available.
4. Incidents that occur near intersections are likely to have a larger effect (expressed in vehicle loss hours) than other incidents.
5. Merge locations and bridges have a higher chance of incidents than other locations.

Furthermore, an analysis of the relation between the duration of the incidents and the effects of the incidents (expressed in vehicle loss hours) shows that a significant linear relation between duration and effect does not exist ($R^2 = 0.11$) (see Figure 3.6). This can be explained by the fact that the impact of an incident depends not only on the duration, but also on other factors, such as spare capacity on the link where the incident occurs and on alternative routes, spillback effects, etc. With respect to the duration, Knoop (2009) analytically derived a formula for the total delay when all factors other than the duration are kept constant. He showed that if spillback effects are not considered, the delay is proportional to the square of the blocking duration. In case spillback occurs, the delay grows faster than proportional to the duration squared. This results in the sixth hypothesis:

6. Incidents with a long duration are more likely to have large effects than incidents with a short duration.

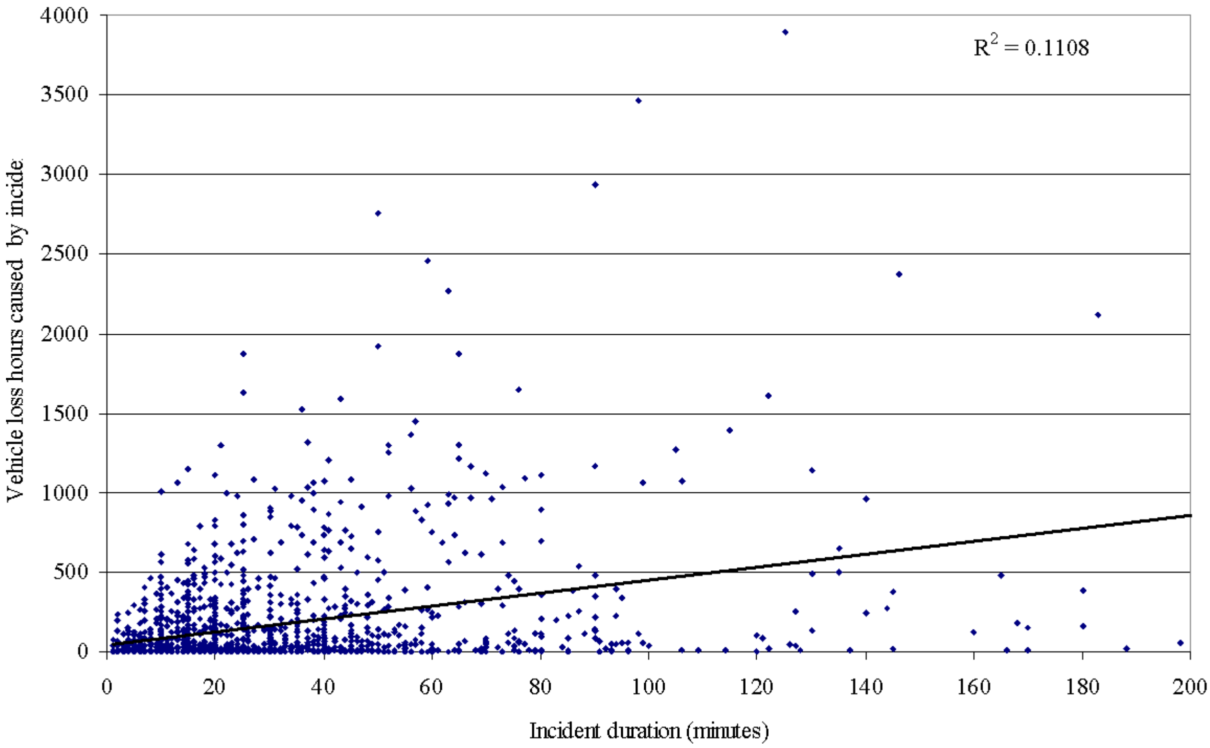


Figure 3.6: Relation between incident duration and vehicle loss hours

Elements of a robust network

The data about incidents is not detailed enough to test these hypotheses statistically. Although, the analysis of incidents does not prove that the hypotheses are true, we do generalize the hypotheses to the following five elements that are likely to make a network more robust:

1. *Prevention:* the road system will continue to function well if disruptions are prevented. However, the concept of prevention is not used here in relation to the robustness of the network in the sense of the prevention of disruptions, but the prevention of congestion *due to* disruptions. For example, if the road surface is heated, then snow and freezing rain are less likely to cause problems on the road. Furthermore, better driver training could result in a better response of drivers in case of disruptions, and therewith to less vehicle loss hours as a result of the disruption. (This element is related to preventive measures on rail infrastructure networks and air networks.)
2. *Redundancy*: the robustness of a system can be increased by introducing a certain spare capacity into the system. This spare capacity is often referred to by the term 'redundancy'. When disruptions occur, this spare capacity can be made available so that the system continues to function better. There are actually two types of redundancy: active and passive redundancy. Active redundancy, like alternative routes, is redundancy in the network that can also be used in regular situations. Passive redundancy refers to back-up

options that are only used in case of disturbances. An example of this are ferries that can be used in case of bridge failures. (This element relates to hypotheses 1 – 3.)
3. *Compartmentalization*: this is the degree to which traffic congestion remains restricted to the relevant link or a small section of the network. If there are less interdependencies in the network, congestion at a centrally located link or node will not cause a series of cascading failures disrupting traffic on large parts of the networks. (This element relates to hypothesis 4.)
4. *Resilience*: Resilience is the capability of the transport system to recover, preferably within a short time period, from a temporary overload. (This element relates to hypothesis 6.)
5. *Flexibility*: the robustness of the system can partly be measured by the degree to which the system is able to fulfil more and different functions than the functions for which the system was originally designed. In other words, flexibility is a property that enables the system to expand in line with new requirements that are demanded of the system.

Four of these components have already been mentioned by Immers et al. (2004b). Furthermore, as was shown in the previous chapter, in other disciplines these elements can be recognized as well, which strengthens the conclusion that a road network is more robust if it includes the above mentioned elements. In railway networks, many preventive measures (such as fences and heating parts of the rails) are taken in order to prevent disturbances (like snow) from having large effects. In the nervous system, the nerves are very well protected, which is an example of a preventive measure. In the Internet, there is a lot of redundancy in the cable network (many alternative routes are available). In airplanes, there is redundancy in the number of engines. Fire doors in buildings are a good example of compartmentalization. The cardiovascular system is an example of a very resilient network. The diameter of the blood vessels can instantaneously be adjusted is such a way that the brains and heart will keep receiving enough oxygen to function properly. And blood vessels are an example of flexibility as well, since the capacity of the blood vessels can be varied.

3.5 Robustness indicators

In order to measure robustness, we need to define one or more robustness indicators. Of course, the quality of the indicator(s) is very important. Below, a set of questions are presented that can be used to score the quality of an indicator:
1. Does the indicator describe the concept of robustness in a complete (all elements of robustness) and logical way? In other words: is there face validity? If the indicator increases, is the network more vulnerable (=less robust), and vice versa?
2. Are there data available for monitoring the indicator?
3. Can the indicator be estimated inside a computer model?
4. Can the indicator be estimated inside a computer model within an acceptable computation time?
5. Can the indicator be explained to policy makers and other people who are not robustness experts?
6. Can the indicator be calculated on a network, route, and link level?
7. Can the indicator be evaluated in a cost-benefit analysis?

The more of these questions that can be answered positively for a specific indicator, the better the indicator is suited for robustness analysis and robustness optimization.

Up to now, no generally accepted indicator for robustness exists. The list below contains indicators that can be used to determine the robustness of a network. This list of indicators is mainly based on (Murray-Tuite and Mahmassani, 2004), (Tampère et al., 2007), (Li, 2009), and (Jamakovic et al., 2008). Of course, others use similar indicators. Furthermore, there might be other robustness indicators that are not included in this list. We classified the indicators in the following way. The *static* (i.e. independent of the traffic flow) indicators refer directly to the properties of a network, and therefore to the robustness of the network. The *dynamic* (i.e. dependent on the traffic flow) indicators refer directly to the robustness of a network. Finally, the *indirect* indicators refer to the travel time and to the stability of the travel time. Between parentheses, the elements of robustness to which the indicators are related are mentioned.

Static indicators

1. The availability and quality of alternative routes (redundancy). If allowance is made for the traffic intensity on the alternative routes, this indicator becomes a dynamic one. An example of such an indicator is the vulnerability index. An aggregation of the vulnerability index across all origins and destinations results in the 'disruption index'. The disruption index accounts for the availability of alternate paths, travel times, marginal costs, and link capacity (Murray-Tuite and Mahmassani, 2004).

   Furthermore, we propose an additional indicator for alternative routes. This indicator is presented in equation 3.1.

$$\text{Altroutes}_a = \frac{cap_a}{\sum_{a \in A_a} (cap_{aa} * \varsigma 1^{dist_{a,aa}})} = \frac{cap_a}{\sum_{a \in A_a} (cap_{aa} * e^{\varsigma 2 * dist_{a,aa}})} \tag{3.1}$$

   In this formula, a is the link where the disruption occurs, aa is a link from the collection A_a of links that form an alternative for link a, cap is the link capacity, $\varsigma 1$ and $\varsigma 2$ are parameters that represent the importance of the distance from alternative routes, and $dist_{a,aa}$ is the shortest distance over the network between link a and link aa. The set A_a is determined by taking a line perpendicular to link a. In Figure 3.7, an example of such a line is shown.

Figure 3.7: Vulnerability indicator alternative routes

The links that cross the black line are considered to be an alternative for link *a* if they meet the following requirements:

- The absolute angle between the original link and the alternative link must be smaller than 60 degrees. Of course, this parameter can be varied. The choice for 60 degrees was made in such a way that routes that do not run more or less parallel are not considered as alternatives. The figure below gives an example of three valid and one invalid route alternative according to this criterion.

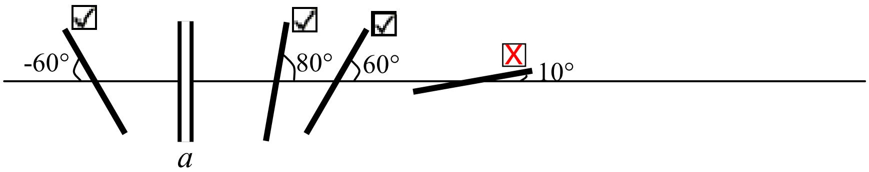


Figure 3.8: Valid and invalid route alternatives

- The direction of the original link and the alternative link must the same.

By multiplying the capacity of the alternative link by the parameter *ς1* with the distance between the two links as the exponent, nearby links are considered more important than more distant links. *ς1* must have a value between 0 and 1. We chose to set *ς1* to 0.8, which implies that links up to about 10 kilometres are considered to be valid alternatives (*ς2* = ln(*ς1*)). This parameter could also be varied. In areas where people are used to making longer trips, this parameter could be set a bit higher. In an alternative formulation, *ς1* is replaced by exp(*ς2*). Another choice that could be made is to make *ς1* trip distance specific. For long distance trips, route alternatives that are far away from each might still be good route alternatives, whereas for short distance trips, far away route alternatives are not good alternatives, since they would result in large detours. In the literature, we could not find any statistical evidence about which detours are acceptable to drivers in incident situations. We chose a maximum of 10 kilometres, because this seemed to be a reasonable distance given the fact that almost 50% of the traffic in the Randstad on the motorways has a trip distance shorter than 20 kilometres (4Cast, 2005). We do not make a distinction between distance classes, because for both short- and long- distance travellers, the far away options are only alternatives when they have not past the point where the original and the alternative route split at the time when they are informed about the incident.

If an alternative link has a higher capacity it is considered to be a more useful alternative. The flows on the alternative route are not considered. Including the capacity of link *a* ensures that links with a higher capacity are given a higher score. The higher the score, the more vulnerable the link is. Links with a score higher than 1 are considered vulnerable. The indicator defined above gives an impression of the vulnerable links in the network solely based on the network structure and can be computed within a short computation time.

This indicator could easily be extended to a time dependent indicator that is based on flows and spare capacity in time and space. In equation 3.2, an extension of this indicator is shown. Here, v_a is the flow on link *a* and rc_j is the spare capacity on link *aa.* This indicator can be calculated for different time intervals or for a longer period. Finally, care should be taken when the spare capacity is computed by subtracting the intensities from the capacities, since low intensities can refer to congested and non-congested states. For instance, if a road is completely congested, the intensity is 0 pcu/hour, which indicates that there is a lot of spare capacity in the network, whereas in practice there is not even space for one extra car. Therefore, a correction should be applied. This can, for instance, be done by setting the spare capacity to 0 pcu/hour if the speed ratio drops below a certain threshold.

$$\text{Altroutes2}_a = \frac{v_a}{\sum_{a \in A_a}(rc_{aa} * e^{\varsigma 2*dist_{a,aa}})} \qquad (3.2)$$

2. Graph theoretical measures, such as:
   - The degree (distribution): a node's degree describes the number of neighbours a node has. The nodes' degree distribution is the probability that a randomly selected node has a given degree (compartmentalization, flexibility).
   - The distance (distribution): The distance distribution *P(dist)* is the probability that the length of the shortest path between a random pair of nodes is *dist.*
   - The clustering coefficient: The clustering coefficient of a node is the proportion of links between nodes within the neighbourhood of a node, divided by the maximum number of links that could possibly exist between those neighbours (redundancy).
   - The connectivity in a network: the link connectivity is the minimum number of links whose removal would disconnect a graph. The node connectivity is defined analogously (nodes together with adjacent links are removed) (redundancy).
   - The centrality/the betweenness: betweenness is a centrality measure of a node (link) within a graph: nodes (links) that occur on many shortest paths between other node pairs have higher node (link) betweenness than those that do not (redundancy).
   - The coreness: The *k*-core of a graph is a subgraph that is obtained from the original graph by the recursive removal of all nodes of degree less than or equal to *k*. The node coreness of a given node is the maximum *k* such that this node is still present in the *k*-core but removed from the (*k* + 1)-core.
3. The distance between on ramps and off ramps (compartmentalization, flexibility).

Direct dynamic indicators

4. Spare capacity. This is the capacity that is not used in normal circumstances (redundancy).
5. The total length of the roads on which the consequences of a disruption are noticeable (compartmentalization, redundancy).
6. The average time before an incident is resolved (resilience).
7. The total number of vehicles on the roads on which the consequences of a disruption are noticeable (compartmentalization, redundancy).
8. The total distance covered by all vehicles over a whole period in the situation with an incident compared with that situation without an incident (redundancy).
9. The total number of arrivals in a specific period in the situation with an incident compared with that situation without an incident (all components).
10. The number of vehicles on the network in a period in the situation with an incident compared with that situation without an incident (all components).

Indirect dynamic indicators

11. Total travel time of all vehicles per time interval in the situation with an incident compared with that situation without an incident (travel time, stability).

12. The extra travel time caused by an incident (travel time).
13. The average speed per time interval in the situation with an incident compared with that situation without an incident (travel time, resilience, and stability).

Most of these indicators can be determined at the *road section level*, *route level*, and *network level*. Indicators at network level tell us something about the functioning of the whole network. Indicators at network level are of particular importance to the network administrator. With indicators that are calculated at network level, it must be taken into account that a small change in, for example, the total travel time of travellers that make use of a network can still have a big impact for individual travellers. A local incident will, for example, reduce the total network performance by only a small percentage (for example 0.5%). This is due to the fact that the incident does not have any effect on a large part of the network, and because the incident causes congestion for only part of the time. However, the individual traveller that finds himself in the middle of the congestion due to that incident can easily suffer an increase in travel time of perhaps 50%. By calculating the indicators at route level, it is possible to determine which road users suffer from an incident (or other disruption). By selecting all routes that run past that incident location, the losses in travel time for the travellers affected by the incident can be made transparent. This is done by zooming in on a part of the travellers and a part of the travel time. Therefore, the effect of a disruption is magnified. By focusing on the road section level, we can determine road sections on which the effect of incidents is noticeable. This means that the effect of a disruption is even more magnified.

In addition to the distinction made between road sections, routes, and networks, we can also make a distinction between *periods*. It is possible to determine some indicators for whole days, parts of days, or time intervals (hours or even minutes). This last-mentioned aspect also makes it possible, for example, to determine how quickly the network 'collapses' after disruption, and how quickly it recovers again.

Finally, for indicators where traffic flow plays a role, a distinction can be made between *passenger transport* and *goods transport* and between *short distance travellers* and *long distance travellers*. Long distance travellers have, for instance, more route alternatives available than short distance travellers, if they are informed about disturbances in time and if they are not too far on their route to be able to switch to the other route. Therefore, it might be needed to use a different value for the parameter *ς2* in equation 3.2 for long distance traffic and for short distance traffic.

As indicated above, a single universal indicator for robustness does not exist. The various indicators defined above all provide a picture of some aspect of robustness. Indicators 1 to 10 refer directly to robustness, and are therefore of importance in the design phase. Indicators 11 to 13 are related to travel time, and are therefore easier to use in cost-benefit analyses, because evaluations exist for travel time. Of course it is ultimately of importance how measures that have been formulated in the design phase to improve robustness affect travel time, reliability, and the loss in travel time as a result of disruptions.

In Appendix B, a selection of these indicators is applied to an example network. This analysis is done in order to show differences among the indicators and to analyse whether or not it is possible to approximate the more complex dynamic indicators with static indicators in order to save computation time. This turned out not to be possible, as is discussed in the next chapter. Jacimovic (2010) did additional research in order to learn whether or not more advanced graph theoretical metrics could approximate the more complex dynamic indicators.

His conclusions were negative as well. Based on these analyses, we chose to use the indicator 'vehicle loss hours caused by incidents' (indicator 12) as our primary indicator of robustness throughout this thesis. This indicator scores as follows on the seven criteria that were mentioned at the beginning of this section:

- This indicator describes the different elements of robustness. Depending on the way it is computed, the indicator can consider spillback effects (compartmentalization), route alternatives (redundancy), and resilience. If flexibility is included in the network, these flexible infrastructures can be considered (depending on the type of model that is chosen). The element 'prevention' relates to the capacity reduction that occurs as a result of incidents. This can be included in the input of the model that is used (criterion 1).
- Monitoring this indicator is difficult. It requires a link between incident registration databases and traffic counts, as is done in the previous section. For a complete picture, traffic counts on the secondary network are needed as well. The most difficult part is the data about the reference day, because reference days do not exist (criterion 2).
- Modelling this indicator is also difficult. Ideally a dynamic model is needed to model spillback effects of incidents properly (Knoop, 2008 and 2009). There are many different disturbances that can occur on many different locations of which ideally the effects should all be modelled by using this dynamic model. Finally, there is a lot of uncertainty about the choice behaviour (route choice, departure time choice, and mode choice) of travellers in case of disturbances. The amount of information given to the traveller can vary, as well as their responses to it. The next chapter describes the way in which this indicator can be modelled in the best possible way. In the next chapter it is explained that we use the alternative route indicator (equation 3.2) in the computation of the vehicle loss hours caused by incidents (criterion 3).
- The indicator is relatively easy to explain, because policy and decision makers are used to thinking in terms of travel time losses. However, it must be clearly stated that not only the travel time delays of people that are in the queue caused by the incident should be considered. The travel time losses as a result of taking detours (of the people that take the detour and of the people that were on the route of the detour) should be considered as well (criterion 4).
- In order to compute the indicator on the network level, the chance of disturbances is needed. Multiplying the chance of all disturbances that are relevant for robustness by the effects of those disturbances results in the total expected travel time loss on the network in a certain period. If different networks (for instance, without and with a robustness measure) are compared, it must be taken into account that measures could for instance result in extra vehicle kilometres driven as a result of distribution effects that could result in lower robustness scores. The construction of an alternative route is, for instance, expected to improve the robustness of the network. However, if this route is fully used in the regular situation, it does not offer spare capacity in case of disturbances on other routes. Because the number of vehicle kilometres driven on the network increased, more people are delayed in case of disturbances, and the chance of a disturbance is larger (criterion 5).
- The indicator can be evaluated by using values of time. However, a value of time in case of unexpected disturbances is not yet available. There is a lot of research into values of reliability. Therefore, it is likely that in the future there will be value of time for travel time losses caused by unexpected disturbances (criterion 6).

As can be seen from the list above, there are still some problems in monitoring and modeling this indicator. Nevertheless, this indicator was chosen because the indicator scores well on the other criteria. Since, there is not one indicator that scores well on all the criteria, a choice was made to prefer an indicator that well describes the concept of robustness, is explainable, can be used to obtain a network wide indicator and can be used in cost-benefit analysis, but has some technical (data and model wise) challenges, over indicators that can easily be measured and modeled, but does not have explanatory power.

Choosing for travel time losses as a result of disturbances as our primary indicator of robustness does not imply that the costs of extra vehicle kilometres driven in case of disturbances should not be considered. However, the costs of extra vehicle kilometres driven are expected to be much lower than the costs of travel time losses, because only a small percentage of the travellers take a detour and because time-related costs are usually higher than distance-related costs (when road pricing is introduced this might change). The indicator chosen can be extended to generalized cost (a composite distance and time related cost). However, the method presented in the next chapter does not allow for this extension.

3.6 Summary: robustness definition and indicators

In this chapter, a common framework was presented that gives policy makers and transportation analysts the possibility to discuss issues that are related to road network robustness and vulnerability. In this chapter research s 2, 3, 4, 5 and 6 have been answered:

Research question 2: How is robustness defined?
In this chapter the following definition for robustness was given:

Robustness is the extent to which, under pre-specified circumstances, a network is able to maintain the function for which it was originally designed.

Vulnerability is the opposite of robustness. A network that is vulnerable is not robust, and vice versa.

Research question 3: Against which disturbances should the network be made robust?
In the definition of robustness, the phrase ‘pre-specified circumstances’ is included. This phrase is included because, in principle, a network can be made robust against all kinds of disturbances ranging from predictable to non-predictable disturbances, from regular to irregular disturbances, from disturbances with a temporarily small or large impact/effect to disturbances that have a permanent impact, and from disturbances that have a network-wide effect to disturbances that have a local effect. Policy makers have to decide how robust the road network is to be made against which disturbances. In any case, if robustness issues are discussed, it is advisable to clearly indicate against which disturbances a network is to be made robust. The robustness of a road network focuses more on effects than on probabilities. Nevertheless, for network design, probabilities are an important factor for investment decisions. Therefore, in this thesis we focus on the class of non-predictable regularly occurring (relatively high probability) disturbances: incidents. We also consider some incidents, such as road closures, that belong in the class of non-regular non-predictable disturbances.

Research question 4: Which elements determine the robustness of a road network?
In the previous chapter, we concluded that a network becomes more robust if there are alternative routes available and there is enough spare capacity on all the routes. Having this redundancy is not enough; the transported elements must also be rerouted to these alternative routes as fast as possible, which requires some flexibility in the network and some traffic and incident management strategies. Furthermore, in a robust network, spillback effects must be

reduced to a minimum level. Finally, preventive measures make a network more robust as well.

In addition to this, in this chapter an analysis of the incidents that happened in a part of the Netherlands was carried out that underlined that five elements (prevention, redundancy, compartmentalization, resilience, and flexibility) are likely to make a network more robust. This is important to realize, because this gives direction to the indicators that need to be chosen for robustness analysis and the measures that can be taken in order to improve the robustness of the road network.

Research question 5: What is the relationship between robustness, travel times, and travel time reliability?
In the end the traveller wants to have reliable travel times and does not want to be confronted with unexpected high travel times as a result of incidents. We showed that robustness and reliability are two related concepts that differ in focus. The focus of robustness is on the network structure and the effects of disturbances with an unexpected effect on a specific moment in time, whereas the focus of reliability is on the traveller and the probability of expected disturbances measured over a longer time period. In a robust road network, disturbances lead on average to smaller variations in travel times, and therewith to more reliable travel times. However, this is not a one-on-one relation. Making the network more robust does not have to result in more reliable travel times for all travellers.

Research question 6: Which indicators can be used to measure robustness?
In this chapter, a list of indicators was presented that refer to one or more elements of robustness. We chose to use the indicator ‘vehicle loss hours caused by incidents’ throughout this thesis, mainly because this indicator describes the different elements of robustness best. Depending on the way it is computed, the indicator can consider spillback effects (compartmentalization), route alternatives (redundancy), and resilience. If flexibility is included in the network, these flexible infrastructures can be considered (depending on the type of model that is chosen). The element ‘prevention’ relates to the capacity reduction that occurs as a result of incidents. This can be included in the input of the model that is used. Furthermore, the chosen indicator is relatively easy to explain and can be assigned a value by using a value of time. Monitoring the indicator (based on data) is complex, but possible to a certain extent. The most complex part of this indicator is to find a model by which this indicator can be computed in an acceptable computation time. The next chapter deals with this problem.

4 Evaluating the robustness of a road network

4.1 Introduction

In this chapter, we propose a method to evaluate the robustness of a road network (or the vulnerability of links) by using the indicator 'vehicle loss hours caused by incidents'. For practitioners and policy makers, it is important to know what the most vulnerable places in the road networks are. Those are the locations where robustness measures need to be taken. On the short term, it is for instance important to quickly remove the wreckage from an incident at those locations. In case of a high risk of an attack, these vulnerable links are the links that should be protected. On the long term, a quick assessment of the vulnerability and the vulnerable parts of a network is also needed for the design of robust road networks with network design models. As explained before, the problem of network design is very complex and computationally intensive, even without the robustness aspect. A very long computation time for the robustness assessment would increase the computation time of the 'robustness network design problem' to an unacceptable level. Finally, robustness is not the only design objective in road network design. Therefore, the method that is used must make it possible to compare robustness with other design objectives.

This chapter starts by presenting the choices that have to be made with respect to the method that is used for assessing the robustness of a road network. The section thereafter presents a framework that is used for assessing the robustness throughout this thesis. The last section of the chapter explains how robustness can be evaluated for the purpose of cost-benefit analysis.

4.2 Literature review: identifying vulnerable links

In the literature, several methods are used to determine the indicators that were specified in the previous chapter. These methods can be distinguished according to the following characteristics:

Whether or not a traffic model is used: to determine the static indicators, a traffic model is not necessary because they can be derived from the structure of the network.

Model type: micro-, meso- or macroscopic; static or dynamic. The distinction among micro-, meso- and macroscopic models is of importance for the level of detail at which robustness is studied. Because robustness is a network property, most microscopic models are not suitable, since these models can be used to study only a relatively small part of the network. Meso- and macroscopic models can both be used. The distinction between static and dynamic models is of importance for robustness because it is important to know how quickly a network 'collapses' and how quickly it 'recovers'. The development over time is, therefore, important. The traffic flow over time can be modelled using dynamic models. It is not possible with static models. Additionally, dynamic models often have model congestion better, which makes them more suited to determine the congestion spillback to other links. This is of great importance, because the degree of congestion spillback (interdependence) is an important factor for the robustness of a network, as was indicated in the previous chapter.

Route choice type: it is a well-known fact that, when disruptions such as incidents occur, some road users change their route. However, not much is known about how much road users change their routes under which incident and information circumstances. A few stated and revealed preference studies have been done into this topic. For example, Koo and Yim (1998) found by means of stated preference analysis that, even if travellers are informed about the traffic situation, 70% of them still do not change their departure time and route choice. In a similar way, Jou et al. (2005) found that when travel times remain within a certain band width, travellers do not change their routes. By analysing loop detector data of 5 different incidents, Kraaijeveld (2008) and Knoop (2009) found that, depending on the severity of the incident, 0% up to 50% of the travellers take another route. In practice, this percentage will probably depend on the information provided, the route alternatives, and the severity and duration of the incident. However, a clear relation between the factors that influence the effects of an incident and the percentage of drivers that take an alternative route has yet to be found.

In the literature, three ways of simulating incidents have been found:

- Fixed route choice: the road users stay on the original route. This is the route that would be chosen if no disruption had taken place. With this form of congestion modelling, an element of robustness remains underexposed, because use cannot be made of the spare capacity on alternative routes.
- User equilibrium route choice: all road users can change their route. This is done in such a way that a new equilibrium is found and the travel time on all routes used between each origin-destination pair is identical. As was the case with staying on the original route, this situation is not realistic, because, in practice, only some of the road users change their route. Additionally, this approach assumes that complete information about the disruption is made available (even before the disruption takes place).
- En-route route choice: with this form of route choice modelling, the road users can change route while on the move. This form of modelling most resembles the choice behaviour that takes place in practice. The disadvantage, however, is that assumptions must be made concerning the amount of information that is available and the response to this by the road users. As we indicated earlier, little is known about these aspects.

  The importance of en-route route choice for the assessment of the impact of incidents is advocated by Li (2008). Tampère et al. (2007) argue that en-route route choice can indeed be of added value, but that it is very difficult to correctly model the en-route route choice of travellers during incidents because of the uncertainty that is inherent in human behaviour (see, for instance, Bogers at al. (2005)). This uncertainty is important especially during incidents, because it is not known how many people have information about the incident and how they will respond to that information. Of course, this is also a problem in the other two route choice methods.

Staying on the original route and user equilibrium route choice can each be seen as an approach to the two extremes that may occur. However, this does not mean that these two extreme situations are reflected in the travel times in the system. It may for example be the case that, if a few road users change their route, the total travel time in the network/system increases, because the road users who change route also adversely affect other road users. Because there is insufficient knowledge about this detour behaviour of motorists with regard to different types of disruptions, one form of route choice behaviour is not necessarily better or worse than the another forms. Applying more than one form of route choice modelling provides more insight into the bandwidth in which the results lie.

Congestion modelling: traffic models usually assume a ‘vertical traffic queue’ or a ‘horizontal traffic queue’. The method of congestion modelling is of importance for determining the robustness of a network. With vertical traffic queue modelling, a vertical traffic queue occurs at the location of the bottleneck in the road network. With a vertical traffic queue, the waiting time for the road users that wish to pass the bottleneck is calculated, but the traffic queue does not block any other flows. This vertical traffic queue is situated in the bottleneck and not before it, as is the case in practice. The fact that congestion can spill back onto other roads is one of the most important reasons why a network can be vulnerable. This principle can only be charted if the congestion is projected horizontally onto the road (in fact using vertical queues is not really congestion modelling), thereby taking capacity restrictions into account and using realistic flow propagation assumptions. Different categories exist within the group of models using traffic queue modelling. The categories are related to the form of the fundamental diagram that is incorporated in the model. The fundamental diagram determines how quickly congestion builds up and clears up, and what the congestion density is. In the majority of macroscopic dynamic models, a fundamental diagram with a constant congestion density is used. This means that the head of the congestion always stays at the same location, even if the bottleneck has already disappeared. It would appear then that the congestion clears up at the tail. In the dynamic model ‘Indy’, a link transmission model is included that works according to the first-order kinematic wave theory of Newell (1993). This incorporates a triangular fundamental diagram. As in practice, the congestion density here is dependent on the capacity reduction of the maximum outflow. If the bottleneck no longer exists, the head of the congestion moves backwards, because the cars at the front of the congestion drive off. At the tail of queue, cars are still joining, so that the congestion moves upstream. As far as we can establish, Indy is the only macroscopic model with this detailed form of congestion modelling. In microscopic models, this form (and even more detailed forms) of congestion modelling are used more often.

Intersection modelling: delays arise at intersections. Especially on the underlying road network, it is of importance to model this effectively. However, modellilng intersections requires data about the intersections. Additionally, according to Yperman the explicit simulation of traffic lights has a number of disadvantages (taken from Yperman, 2007, page 148):

1. “Explicit simulations of traffic light measures and gaps in priority flows demand a small simulation time interval. This results in a critical increase in calculation time.

2. Explicit simulations generate frequently fluctuating travel times. However, road users do not take these highly frequent fluctuations into account when determining their route. Average or ‘expected’ travel times would be more relevant.
3. Explicit simulations result in ‘possible’ travel times, which are the result of a coincidental set of circumstances. Average or ‘expected’ travel times would be more relevant.”

Generally, there are three possibilities for determining the robustness of a road network by using models and methods with different characteristics, as described above:

1. **Full computation:** one simulates all possible disturbances in a road network, which is computationally expensive. In the group of “full computation methods”, the capacity is reduced for each link separately. In order to find out which links in a network are the most vulnerable, a complete simulation could be made. That is, for each link the capacity could be reduced and an assignment could be made. The effects of the capacity reduction on the total travel time could be regarded as an indicator for the vulnerability of a link. Jenelius (2007) uses the approach of blocking each of the links in a simulator without traffic jams. Knoop et al. (2007) use the same approach computing the consequences for a blocking at each link. However, they argue that the network effects, including spillback, are significant. Hence, they use a more accurate simulator that represents the dynamics of traffic jams, including spillback. The advantage of the approach used by both ─ a full computation ─ is that it gives a complete analysis. However, the computation time of this approach is very high, which can be considered as a disadvantage; furthermore, this brute force method lacks a structure for searching for weak links.
2. **Pre-selecting vulnerable links:** one pre-selects potentially vulnerable links based on an equilibrium assignment and certain criteria, and performs an additional analysis for the selected links. Several approaches have been introduced in order to overcome the disadvantage of full computation. In this group of approaches, a first selection of links that are likely to be vulnerable is made based on certain criteria. For these links, a more detailed analysis is made by reducing the capacity and by assessing the vulnerability of these links based on more detailed simulations. In the Netherlands, the ‘Robustness scanner’ (Tamminga et al., 2005) was the first method based on a static traffic assignment model in which this approach was used. Also Tampère et al. (2007) and Li (2008) introduced their own selection criteria. These methods are still computationally intensive, because simulations for all the selected links are required. This approach raises the following questions. What is the quality of the selection criteria used in the second group? How large should the selection of possible vulnerable links be to be sure that the most vulnerable links are indeed included? And, if the selection is good, is a detailed analysis really needed, or could the vulnerability and robustness of a network (or parts of the network) also be determined by applying only the selection criteria (without reducing the capacity for a selected link)? If this is possible, then it would make the modelling of the implications of protective measures much easier. In (Knoop et al., 2010) and in a similar analysis presented in Appendix B, we concluded that different criteria proposed in the literature indicate different links as most vulnerable. Excluding freeways gives a completely different list of vulnerable links from the list when freeways are included. This implies that the freeways are usually indicated as vulnerable. The Incident Impact ratio (I/(1-I/C)) produces the best correlation with the other factors. When comparing it to the fully computed results, though, it is not better than the others. In fact, none of the selection criteria on their own give a good representation of the full consequences of the blocking of a link. It is

also insufficient to take the top-level numbers and analyse them in depth, as there is no indication that the indicated top-level vulnerable links are indeed the most vulnerable. Apart from that, they differ among the criteria. Furthermore, a combination of the criteria also did not result in a good prediction of the list of most vulnerable links. The combined selection power of the criteria in the network appeared to be minimal. In particular, the freeway junctions, the links after the junctions, and the main urban arterial roads, are not well covered by the criteria. This could imply that spillback effects are not properly included in the criteria. From these results, we conclude that the quality of these criteria is not good enough to properly identify the most vulnerable links in a network.

3. **Vulnerability analysis based on a single equilibrium assignment:** one does a single user equilibrium assignment for a situation without disturbances. Based on that assignment the effects of incidents on all links are analytically approximated. This is in fact a combination of the first and second methods. In (Luathep et al., 2010) such a sensitivity analysis based approach that deals with the implicit relationship between the input data of a traffic assignment model and the equilibrium network flows based on that data has been proposed and applied to realistic networks. In this approach a static probit-based stochastic user equilibrium has been used. In the Netherlands, we used four indicators to create a combined indicator that describes the different elements of robustness. These indicators can all be computed based on a single equilibrium assignment using the national or regional model systems. The method is applied and described in the national market and capacity analysis (NMCA) (Snelder et al., 2010a). Furthermore, the method is briefly described in Appendix C. The advantages of this method are that (1)it is relatively fast, based on model runs with the national or regional model system that have to be carried out anyway to compute the travel time under regular circumstances, and (2) it describes the different elements of robustness. A disadvantage of this method, and the method of Luathep et al. (2010) as well, is that it does not consider spillback effects and time dynamics, since a static assignment model is used. Finally, this method does not produce an estimate of travel time losses caused by incidents, but produces an indicator that can be used only to describe relative differences. In the next section, a method is described that belongs to this class of methods, but that overcomes the above-mentioned shortcomings.

Finally, there are also other methods, such as the game-theoretical approach presented by Bell (2000). However, this method has to the best of our knowledge never been applied in a dynamic simulation environment on a real-size network. In Appendix C, different methods that are used in the Netherlands for robustness/reliability analysis are mentioned.

4.3 Framework for assessing robustness

In the previous sections, we presented the advantages and disadvantages of several available methods to measure robustness. Ideally, a method that is used should meet the following, most important, requirements:

- Spillback effects should be modelled. Spillback effects are the cause of the fact that the effects of local disturbances spread all over the network. In a robust network, these effects are minimized (for instance by creating compartmentalization). If robustness is to be assessed, these effects should be captured.

- Alternative routes should be included in the route choice. A network is more robust if alternative routes are available since they offer spare capacity that can be used in case of disturbances. Furthermore, they create a balanced network.
- Time dynamics should be included. Since the speed at which network performance drops during disturbances and the speed at which the network recovers after disturbances is important for the robustness, time dynamics should be included.
- The method should be fast: since measuring robustness requires a lot of simulations of different disturbances on different locations, a method with a short computation is preferable.
- The method should be able to deal with all kind of disturbances in such a way that the complete travel time distribution is modelled.
- The method should be able to deal with intersection delays, because in a robust road network, regional (and local roads) are important elements.

The dilemma in the model choice is that a choice has to be made between accuracy (ideally using a dynamic traffic assignment model with detailed congestion modelling, including spillback effects, with multiple types of route choice behaviour during incidents, and with an accurate intersection modelling) and computation time. In general, the most accurate models have the longest computation time. Using rulse of thumb takes hardly any computation time, but is not accurate. For some applications, a rule of thumb can be good enough to get a quick impression of the robustness of a network. However, to make a well balanced decision about robustness measures in network design, it would be better to look for a method/model that deals with the above requirements in the best possible way.

To the best of our knowledge, there is not yet a model that covers all these six features completely. We chose to use the dynamic traffic assignment model ‘Indy’, because this model has an accurate network loading model that models spillback effects according to the simplified kinematic wave theory of Newell (Yperman, 2007). Furthermore, the model can compute an equilibrium route choice and can deal with fixed route choice. En-route route choice is not possible. However, this is not a problem, since Indy is used only for a basic run in a situation without disturbances, for which an equilibrium assignment is most appropriate. Indy is a dynamic model, which makes the modelling of time dynamics possible. And, since Indy is able to work with the marginal incident computation model (MIC) (Corthout et al., 2009), it is able to get an estimate of the impact of incidents very quickly. The MIC-module is currently capable of dealing only with fixed route choice during incidents, which can be considered as a disadvantage. However, to partly overcome this shortcoming, we used an approximation method for the use of alternative routes,. Finally, the MIC-module can only simulate the effects of local capacity reductions. Therefore, variations in demand and network-wide capacity variations (for instance as a result of rain) cannot be modelled with the MIC-module. If the impact of these disturbances is to be simulated, a complete run with Indy has to be done, which results in extra computation time. Finally, Indy does not have an explicit modelling of intersections and traffic signals. Therefore, it underestimates the delays at intersections. In Appendix D, it is shown that, to a certain extent, delays at intersections can be approximated by using outflow constraints that reflect the capacity restraints at intersections.

Given these arguments, we came to the method that is presented in Figure 4.1 for assessing the robustness of a road network that can be used in network design. In the first step, an equilibrium assignment is done with Indy. This results in the cumulative inflows and outflows per link (to all other next links). This is used as an input for the MIC-module. Furthermore the

MIC-module needs to know on which link what kind of incident occurs. We chose to model four types of incidents on all links. As was shown before, it is very difficult, if not impossible, to make a pre-selection of the most vulnerable links. Therefore, the choice was made to use this full computation method based on one equilibrium run (method 3 from section 4.2).

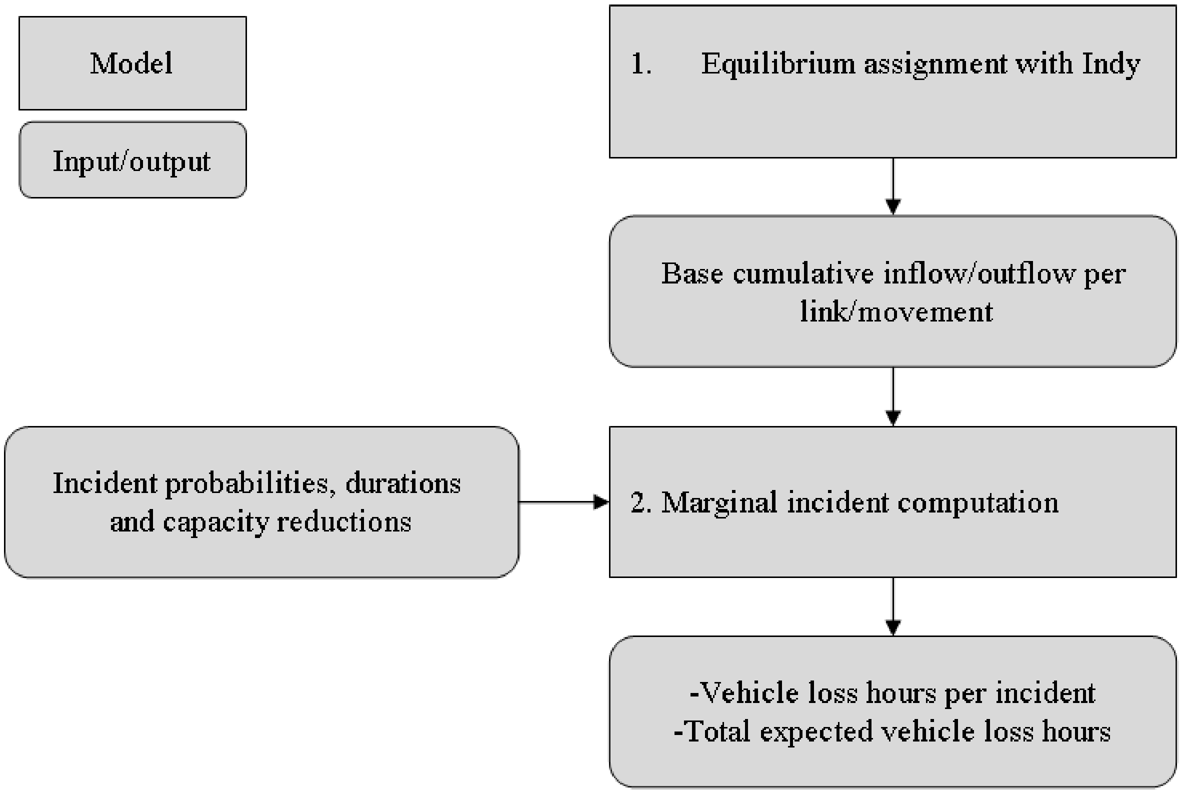


Figure 4.1: Framework for assessing robustness

We model the following four incident types:

1. Car break down with a duration of 15 minutes, a capacity reduction of 5%, and a chance of occurrence of 0.34 per 100,000 vehicle kilometres.
2. Small incident that blocks one lane, with a duration of 45 minutes, a capacity reduction of 20%, and a chance of occurrence of 0.06 per 100,000 vehicle kilometres.
3. Big incident that blocks more than one lane, with a duration of 45 minutes, a capacity reduction of 70%, and a chance of occurrence of 0.015 per 100,000 vehicle kilometres.
4. Rubbernecking, with a duration of 45 minutes, a capacity reduction of 15%, and a chance of occurrence of 0.07 per 100,000 vehicle kilometres.

These incident types, with their matching incident probabilities, durations, and capacity reductions, are based on the statistical analysis carried out in SMARA (Meeuwissen et al., 2004; Appendix E). In SMARA, however, a distinction is made among different road types, and among roads with one, two, and three or more lanes. It is, for instance, logical that the capacity reduction of the third incident type depends on the number of lanes. Since the input of the MIC-module is link specific, this distinction can be made here as well, without any problems. The numbers above are presented to give an impression how the model works. In practice, all input files should be adjusted to the specific situation.

The probabilities of the incidents are not needed to compute the effects. Strictly speaking, they are not in input to the MIC-module. However, they are needed to compute the expected vehicle loss hours in a certain period. Finally, also the start time and end time of the incidents are needed as an input. The number of vehicle loss hours of an incident is, for instance, different in the peak period than in the off-peak period. In order to keep the computation time within acceptable limits, the number of periods for which an incident is computed should be kept as limited as possible. Choosing some representative incident types (chance, duration, reduction, and start time) is therefore always advisable. It is up to the model users to do this in the most appropriate way.

From this framework it can be seen that we chose to model only a few incident situations on all links. This gives an impression of the travel time distribution, but it does not capture all possible variations in travel time. In the future, the model should be extended to make this possible.

The output of the method is the number of vehicle loss hours per incident per link. The output also indicates which links are affected by each incident. Furthermore, the vehicle loss hours are also given per route. Multiplying the vehicle loss hours by the incident probabilities gives the expected vehicle loss hours. In fact, the output is the travel time distribution. From this distribution a variety of indicators can be computed ─ for example, the standard deviation of the travel time, in addition to the indicators mentioned before.

Below we explain in a bit more detail how Indy and the MIC-module work. Furthermore, we explain that both Indy and the attached MIC-module are in their original forms, not capable of dealing with alternative route choice during incidents. However, we have implemented some approximation methods to get an indication of the effects.

4.3.1 Dynamic traffic assignment model Indy[24]

Indy is a macroscopic dynamic user equilibrium model that can be used to evaluate the impacts of congestion relief strategies, such as infrastructure expansions. It can also be used to evaluate the effects of all kinds of dynamic traffic management strategies and to determine the impacts of incidents and other disruptions. Indy shows the locations where congestion occurs and how the congestion is propagated through the network. The equilibrium approach of Indy's dynamic traffic assignment (DTA) produces chosen paths that are consistent with drivers' desire to minimize their travel costs.

In Figure 4.2 the model framework of Indy is depicted. It consists of three main modules: route generation, route choice, and dynamic network loading.

Each of these main modules can contain different kind of models. Due to the modular setup, different combinations of route generation models, route choice models and dynamic network loading models can be made.

[24] This section is based on *(Bliemer, 2005)*.

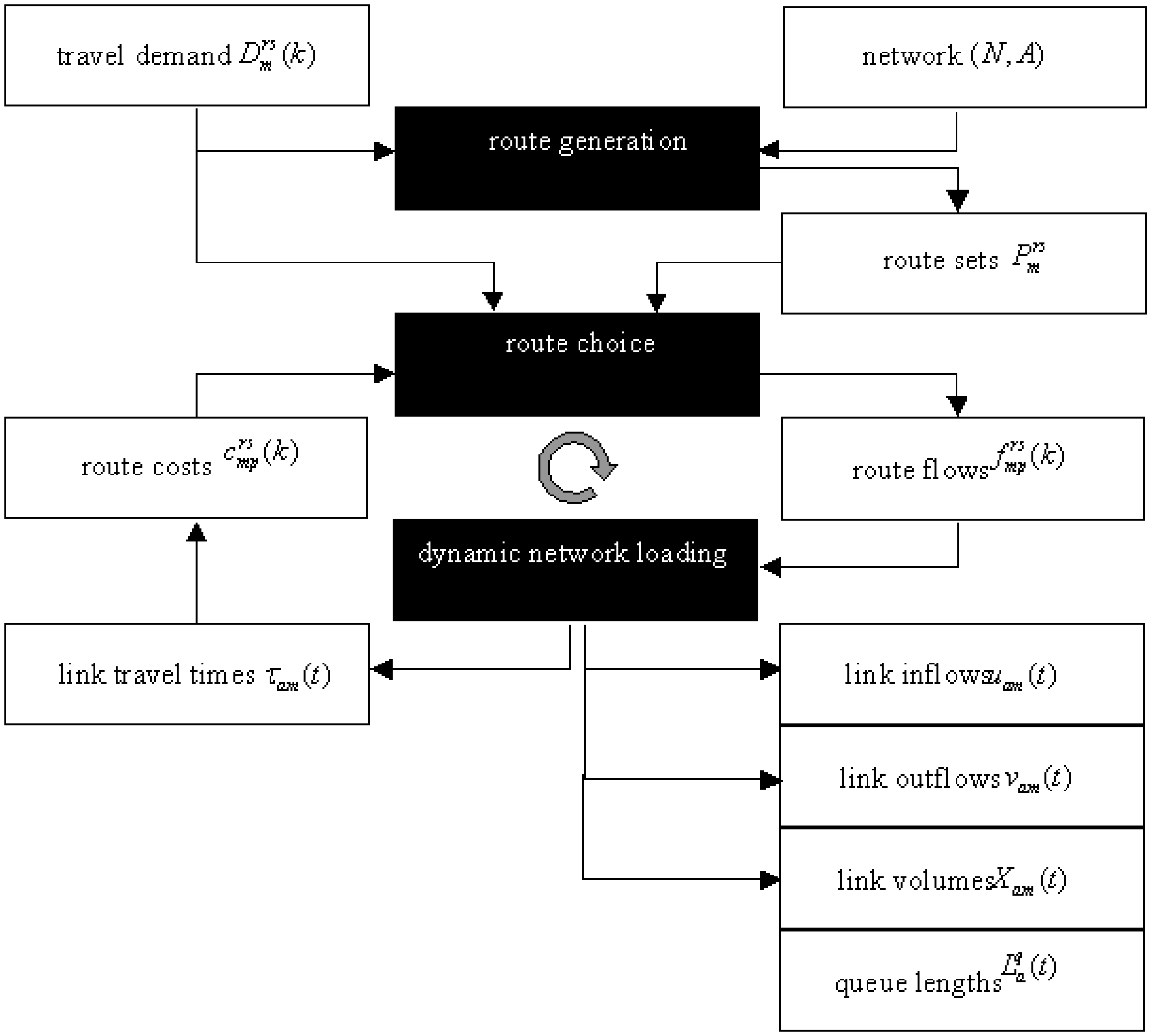


Figure 4.2: Flowchart Indy[25]

First, the route generation module determines the routes based on the network characteristics and travel demand. There are three methods implemented for generating the routes: a Monte Carlo approach, an approach using a static traffic assignment, and an approach in which a pre-specified route set is used. The output of the route generation module will be route sets for all vehicle classes that describe the available routes between each origin-destination (OD) pair.

Second, the route choice module models the behaviour of the travellers by choosing the best route for them from the set of available routes determined in the route generation module. The best route depends on the route costs for each of the alternatives and consists mainly of the route travel time, but can include other (non-additive) cost components, such as tolls. The outputs are dynamic route flow rates between each OD pair on the network.

Finally, the dynamic network loading module propagates the traffic along the chosen routes. This module is the heart of the Indy model. Its outputs are link characteristics, including link inflows, outflows, volumes, queue lengths, and travel times. The link travel times can, in turn,

[25] Source: (Bliemer, 2005)

be used to compute the route costs. There are three different network loading models implemented in Indy. The first model uses link performance functions to compute the link travel times in order to propagate the flow through the network. The second model explicitly assumes hard capacity constraints on link inflows and outflows, leading to a dynamic queuing model. The third model is the so called ‘link transmission model’. The link transmission model is the most accurate of the three, because it comes closest to the way in which queues build up and resolve in practice. A detailed description of the link transmission model can be found in (Yperman, 2007).

Figure 4.3 and Figure 4.4 show, respectively, the building up and resolving of queues according to all three models. The vehicles drive from the south to the north. The width of a link indicates the flow density. The colour indicates the speed ratio (green or light gray indicates free flow, red or black indicates congestion). At the location where two links come together (in the upper left of the figure), a bottleneck occurs because the demand gets higher than the capacity at a certain moment in time. A little later the demand decreases and the bottleneck disappears.

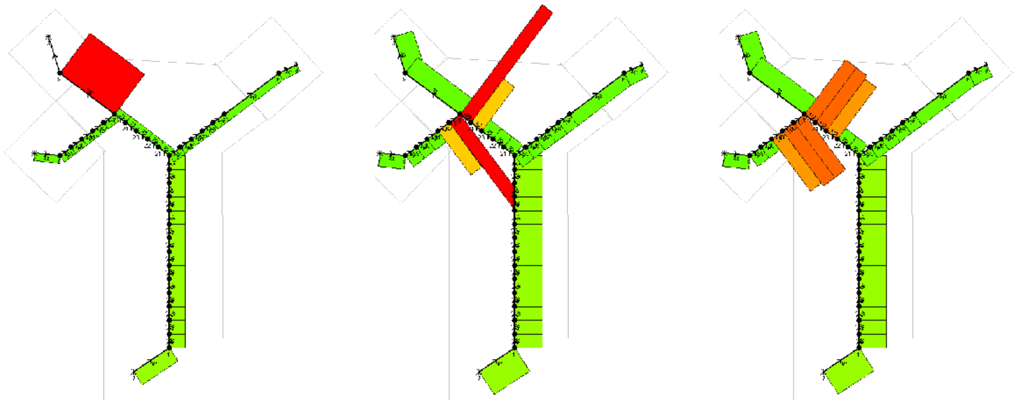


Figure 4.3: Build-up of queues in the vertical queuing model (left), the dynamic queuing model (middle), and the link transmission model (right)

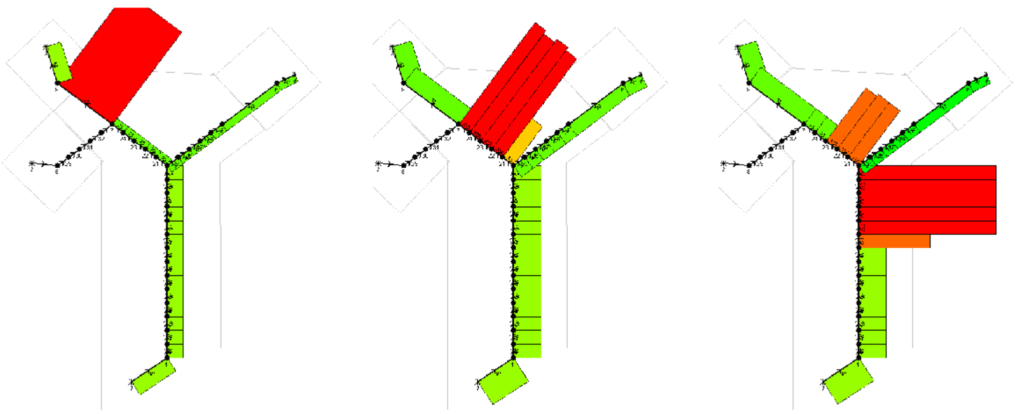


Figure 4.4: Resolving of queues in the vertical queuing model (left), the dynamic queuing model (middle), and the link transmission model (right)

In the vertical queuing model, a vertical queue occurs at the location of the bottleneck. In this way the delay of the vehicles that need to pass that point is computed, but the queue does not block other flows. In the dynamic queuing model, a horizontal queue occurs upstream of the bottleneck. This is what happens in practice. Therefore, the building of queues is more realistically modelled than in the case with vertical queues. This queue actually blocks other flows upstream. In the link transmission model, a horizontal queue occurs as well. However, in this case the building up and resolving of the queue is more realistically modelled. In the dynamic queuing model, a fixed density is assumed during congestion. This makes it seem like the congestion resolves from the tail of the queues, whereas in practice vehicles leave at the front of the queue and keep arriving at the back of the queue until the congestion is completely gone. So, in the link transmission model the congestion moves in the upstream direction just as it does in the real world, as is also indicated in Figures 4.3. and 4.4. In (Yperman and Snelder, 2010), the three types of congestion modelling are described in more detail. The paper also describes two case studies that show that the different queue propagation characteristics substantially influence the queue locations and travel times. Results for network performance deviate substantially when using different traffic models. For cases in which network performance or network reliability is evaluated in heavily loaded or congested real-world networks, the use of realistic traffic models like the link transmission model is indispensable.

In general, we always use the link transmission model, since it is the most accurate one. Therefore, when we assess the robustness of a network and when we design a robust network, we use this model as well. However, before we design a robust road network, we optimize the capacity under regular circumstances. For this particular application, we use the first network loading model, since vertical queues are very useful for finding bottlenecks. This is explained in more detail in Chapter 7.

Finally, after network loading, there is feedback from the new route costs to the route choice module, leading to new route flow rates, after which a dynamic network loading is performed again. These two modules are executed iteratively until convergence is reached. The following sections describe Indy is in more detail.

4.3.2 Marginal Incident Computation (MIC)[26]

Traditionally, incidents are simulated in an explicit way. The capacity of a link is reduced and the resulting traffic flow is calculated through the entire network. In that way, every link flow in the network is recalculated for each incident, even though many links are not influenced by the incident during all or most of the simulated time increments. Obviously, these redundant calculations are inefficient, causing high computation times. The MIC model presented in (Corthout et al., 2009) approximates (additional) congestion caused by an incident by superimposing it on a single base simulation (without incidents). Drivers are assumed to make the same journey (no changes in departure time, destination or route) in case of an incident as they would make in the base situation. Rerouting effects are thus neglected. (In future versions of MIC, this rather stringent assumption will be partially relaxed.) Consequently, the influence zone of an incident becomes relatively small. The *influence zone* is a set of 'affected' links, starting at the link on which the incident takes place and growing in the upstream direction. The affected links are links on which the incident spillback wave restricts the original link flow of the base situation and thus causes (additional) congestion. For every simulated incident, new calculations are carried out only for the affected links. Since in large networks the influence zone of an incident covers only a few percent of the entire network, a significant reduction of computation time is possible. It is shown that the results obtained with

[26] This section is based on (Corthout et al., 2009).

MIC vary only slightly from the outcomes of a complete dynamic network loading, but the gain in computation time is significant: a factor > 1100 for a case study of the Sioux Falls benchmark network.

From the base simulation run with Indy, the cumulative vehicle numbers for all link boundaries are known for the incident-free situation. The MIC algorithm is repeated for every incident of the set, using the same base simulation. It is assumed that an incident always takes place at the upstream boundary of a link. However, the module could easily be changed such that any location could be chosen, since Newell's theory (Newell, 1993) allows constructing cumulative numbers at any location within a link. From there on, the incident constraint is tracked upstream as a spillback wave. The constraint alters as the wave travels further upstream, starting as the reduced capacity of the link where the incident takes place. If the spillback wave reaches a node, the incoming links of that node are subjected to the constraint imposed by the spillback wave. If the spillback constraint restricts the flow of an incoming link, this link is affected by the incident. Every affected link is checked to see whether or not congestion reaches the upstream link boundary. If it does, the spillback constraint is passed on to the upstream node and the influence zone of the incident expands further. The cumulative vehicle numbers from the base simulation are reduced for every affected link, according to the spillback constraint. After clearance of the incident, capacity at the incident location is restored. The acceleration wave proceeds through the affected links in a way similar to the spillback wave and finally catches up with the spillback wave.

The MIC Algorithm is a simplification of the real effects that occur during incidents. We already mentioned that alternative route choice is not modelled. The two most important other simplifications are mentioned below. (A complete overview of the algorithm and the simplifications can be found in the paper of (Corthout et al., 2009)).

- *Neglecting downstream effects*: the MIC algorithm proceeds in the upstream direction, starting from the incident location. Links downstream of the incident are not examined, since only a shift in the cumulative vehicle numbers is expected there, not an additional delay. Secondary downstream effects, such as relieving or overloading of downstream bottlenecks, are thus not taken into account.
- *Neglecting delayed spillback*: When a spillback wave reaches a node (and flow over this node thus decreases), the downstream effects of this decreased flow are not incorporated into the algorithm. An important consequence is that 'delayed spillback' is not accounted for. Delayed spillback occurs if multiple spillback waves travel towards the same node via different routes. By neglecting the delayed spillback effect in the MIC algorithm, the arrival time of the second spillback wave at node n is not correctly calculated. Tests showed that this can lead to a significant error on isolated circuits, while on a realistic traffic network, notable errors due to delayed spillback are expected to be quite rare. Because of this approximation, multiple spillback waves over the same affected link are not simulated, which means that gridlock effects are implicitly neglected.

4.3.3 Alternative routes

The fact that neither the Indy or the MIC model are suitable for assessing the usefulness of alternative routes can be considered as serious shortcomings, because alternative routes are an important part of robustness. Therefore, for both models we implemented a simple algorithm that gives an approximation of what will happen if alternative routes are used in incident situations. Both methods require as input the maximum percentage (x%) of drivers that choose an alternative route during incidents.

For Indy, we implemented two path shift algorithms. In the first, an equilibrium assignment is done for the case without disturbances. From this, we know the available paths, path departure flows, and path travel times. Drivers only choose this alternative path if it does not have a travel time larger than their original travel time + *y*%. Furthermore, they only choose an alternative path if it makes sense given the time of the incident and the time at which they pass the splitting point between their original path and the alternative path. The car drivers should be before the point at which their original path and the new path split, or they should still be at their departure location when the incident occurs and the information about this reaches the driver (the time for spreading information is a parameter in the model). From all the drivers that can choose an alternative path, *x*% switch to suitable alternative paths. Based on this algorithm, the path departure flows are changed and a new simulation with the revised path departure flows is carried out (only 1 iteration with Indy).

The second version of the path shift algorithm starts with an equilibrium assignment and, thereafter, simulates the effects of an incident by assuming that everybody sticks to their original path (only 1 iteration with Indy). The resulting path travel times can be compared with the travel times of the case without disturbance. The people with a delay larger *z*% will choose an alternative path if they meet the other criteria above. This module requires extra calculations and an extra simulation with Indy (1 iteration). It is, therefore, slower than the first algorithm. However, it is more accurate, since, in this algorithm, people that do not pass the incident location can change their route if they experience delays as a result of spillback effects of incidents. In the first algorithm, only the people that actually pass the incident location can choose alternative routes.

Another possible extension of the model could be a refinement of which alternative paths are actually chosen. In the current version all alternative paths get assigned an equal amount of drivers from the blocked path. However, some alternative paths will always be preferred over others, since they have shorter travel times.

The above mentioned algorithms illustrate how difficult it is to correctly model the behaviour of drivers. Many assumptions have to be made by the modeller. Therefore, it is advisable to use the model with different parameter settings. This could be seen as different scenarios with respect to the information that is given to drivers and the way they respond to the information. In section 4.3.4, we will vary the parameters to get an idea which parameters give the closest fit to available data on real world incidents.

For the MIC-module, we use an even more simplified algorithm. In this algorithm, the maximum available spare capacity on the alternative routes[27] and the percentage of drivers that choose an alternative route (*x*%) is added to the capacity of the incident location (The jam density and the capacity reduction are adjusted as well.) If, for instance, 15% of the traffic that passes an incident location chooses an alternative route, the intensity on the original link is 15% lower. The effects of this can be approximated by not reducing the intensity by 15%, but by increasing the capacity by the same number of vehicles as the 15% reduction. In this way, the number of travellers that experience a delay as a consequence of the incident is the same, and the spillback effects are the same as long as the congestion does not spill back over a

[27] The available spare capacity on alternative routes is determined similar to the procedure used for calculating the altroutes index (section 3.5.): $rc_{\mathrm{a}} = \sum_{a \in AA_a} (rc_a * e^{\varsigma 2 * dist_{a,aa}})$

node. If the congestion does spill back over a node, the capacity reduction is passed on to the following links in a different way. This is shown in the example below.

Example:
A node has one incoming link and two outgoing links – one that goes to the incident (route A) and one that goes to the alternative route (route B). In the regular situation, the flow is 3450 in the direction of the incident route and 550 in the direction of the alternative.

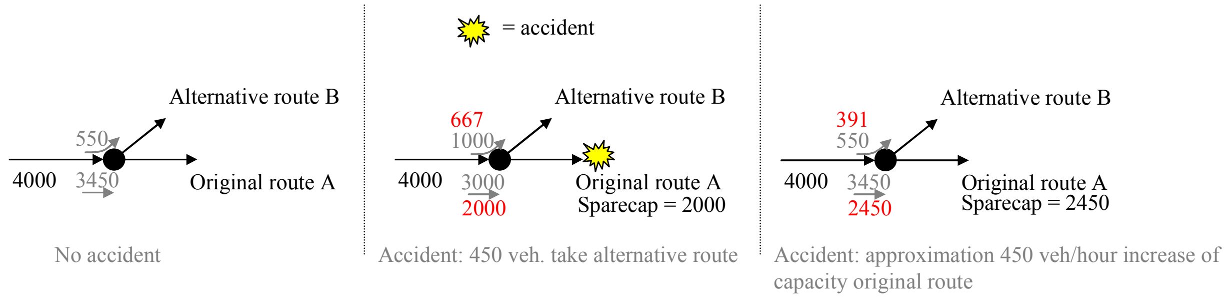


Figure 4.5: Approximating alternative routes

In case of an incident, a capacity restriction of 2000 pcu/hour due to an incident could result in a new distribution of flows: 3000 on the first route and 1000 on the second. The outflow in the direction of the incident is then restricted to 2000 pcu/hour due to the incident, and the outflow to the other direction is restricted by the same ratio to 667 pcu/hour in the node model of Indy. In the approximation that we use, the same incident leads to a capacity restriction of 2450 vehicles per hour (2000 + 450 vehicles that choose an alternative route). The incident restricts the flow in one direction to 2450 pcu/hour and in the other to 391 pcu/hour (= 2450/3000 * 550). The flow further upstream in this approximation method is therefore restricted to 2841 (2450 + 391) pcu/hour instead of 2667 pcu/hour.

This example illustrates that the spillback effects are not modelled 100% correctly in the approximation that we use. Furthermore, possible delays on the detours are not considered either. These are not only delays as a result of possible extra congestion on the alternative routes for the original and new users, but also extra travel time as a result of taking an alternative route that has (also without the extra congestion) a higher travel time than the original route, otherwise the alternative route would have been chosen in the first place. In the future, the MIC-module will be improved in such way that it can better deal with alternative route choice. For now, this approximation is used, since it accounts to a certain extent for additional benefits if alternative routes are available. Furthermore, we believe that the extra error that is introduced is not bigger than the error that we make anyway by not knowing how people behave during incidents.

4.3.4 Validation of Indy for incident situations

It is important to know what the quality of the modelled effects of incidents is if the framework mentioned above is used. Indy was, therefore, calibrated for the regular situation on the network of the area Rotterdam-The Hague. Then, five incidents were modelled by using different percentages of travellers that choose an alternative route (using the first path shift algorithm) and the effects were compared with the available data about the incidents. An extensive description of this comparison and the results is given in (Muller, 2009). In this section, a short summary is given:

First, a reference model was calibrated for the area Rotterdam – The Hague in the Netherlands, based on loop detector data of nine Tuesday mornings in March, April, and May 2007, for the period 5.00 – 10.00 AM. The travel times of eight trajectories and the speeds on the roads were taken as indicators.

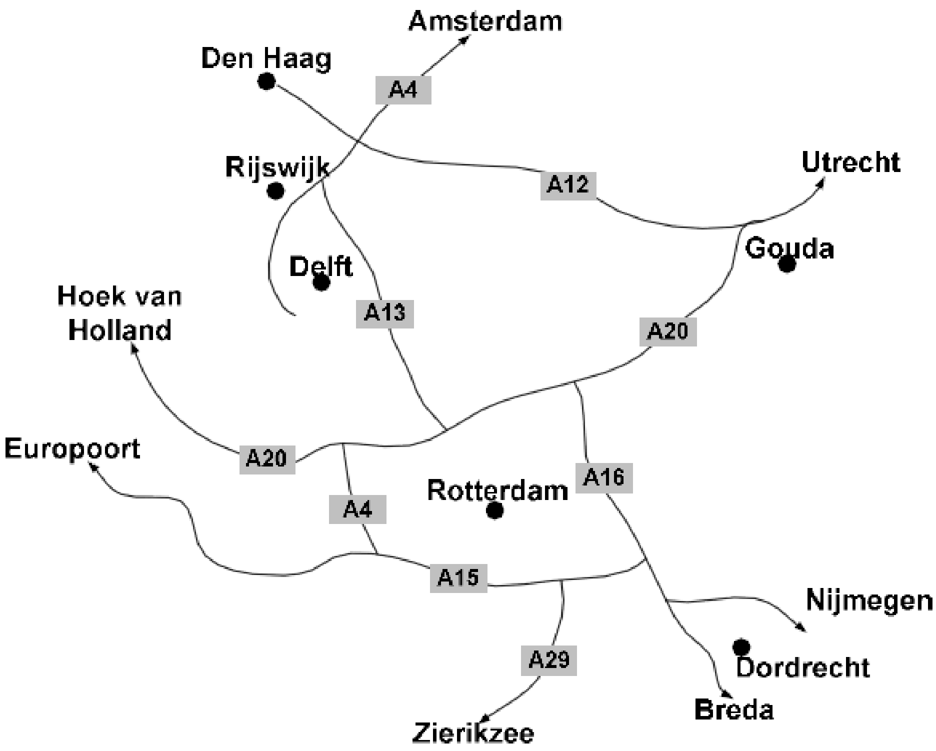
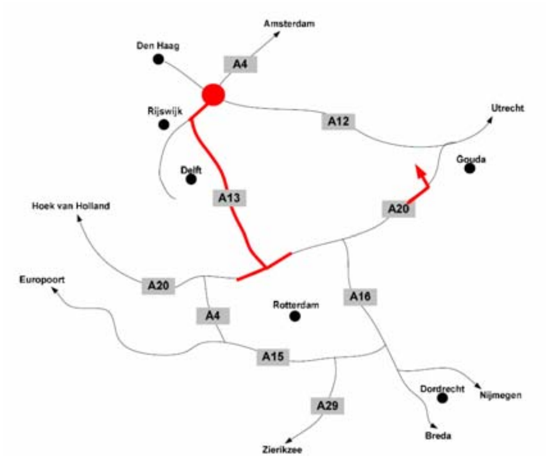


Figure 4.6: Network Rotterdam – The Hague (Den Haag in figure), which was used for validating Indy for incident situations

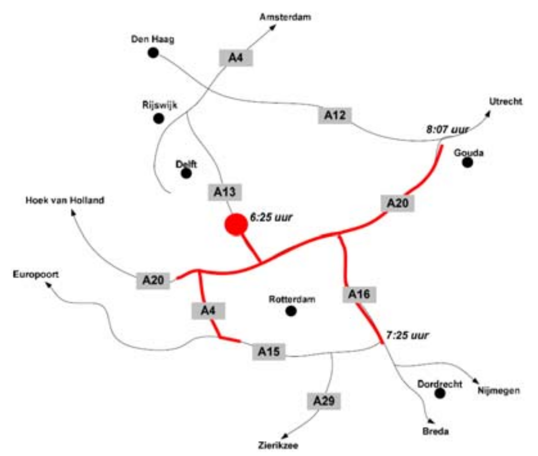
For five different incident cases, the travel times and speed on the same eight trajectories were collected from loop detector data:

- Case 1: Accident A4 Prins Clausplein (Tuesday 23rd January 2007) ─ an accident on the A4 in the direction of Amsterdam occurred that blocked two lanes between 5.48 AM and 8.25 AM.
- Case 2: Accident A13 (Thursday 1st March 2007) ─ an accident occurred on the A13 in the direction of The Hague. Two out of three lanes were blocked from 6.25 AM to 7.20 AM. The traffic could use the lane that was left, and the hard shoulder.
- Case 3: Accident on the intersection A16/A20 in the direction A20 -> A16 shortly before 7.00 AM (25th September 2007).
- Case 4: Accident A20 (19th November 2007) ─ an accident on the A20 in the direction of Gouda at 7.45 AM.
- Case 5: Accident A13 (Wednesday 12th December 2007) ─ an accident occurred at about the same location as Case 2. This accident blocked two lanes between 6.32 AM and 8.04 AM. However, in the months between these two incidents, a peak hour lane had been opened that could be used during the incident.

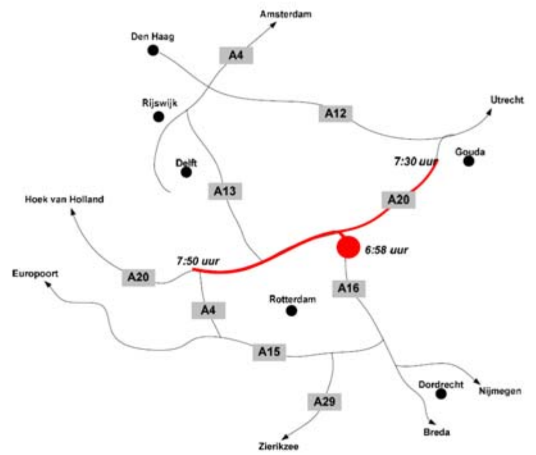
Figure 4.7 shows the spillback effects that occurred as a result of these five accidents.



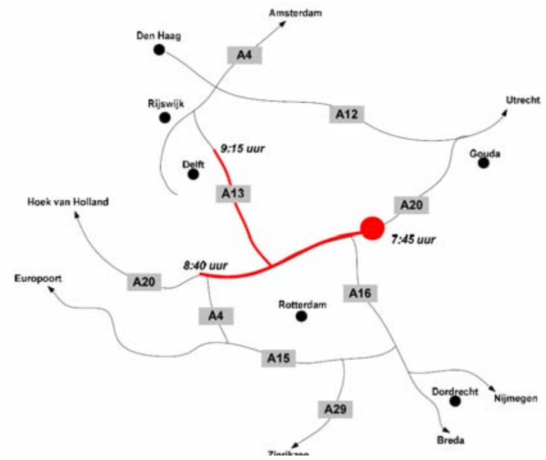
Case 1



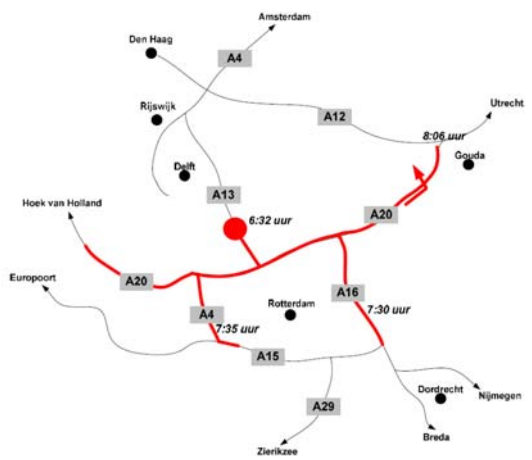
Case 2



Case 3



Case 4



Case 5

Figure 4.7: Locations where the effects of the incidents could be measured

The model was calibrated for incidents by the use of three incident cases (Cases 1, 3, and 5). In the calibration, the percentage of users that change routes as result of an incident was varied to find out which percentage matches best with the loop detector data that were collected for the incident situations. We found that every incident is different. Forecasting the impacts is, therefore, extremely difficult. Nevertheless, the results showed that the model is a reasonable representation of reality (root-mean-square percentage (RMSP) error between 0.2 and 0.3). The best fit was found when 15% of the road users changed their route as a result of an incident. As an example, in Figure 4.8 the model fit for the regular situation and an incident for one of the motorways is shown. It can be seen that, when 15% of the road users change their routes, the line of modelled travel times follows the line of the accident data very well for the A20 in the direction of Gouda and reasonably well for the A20 in the direction of Hoek van Holland. In the latter case, the travel times are first underestimated and then overestimated. For the A13, the model seems to be better when a percentage of 0% or 7% is used. Nevertheless, the two lines are not that far away from each other.

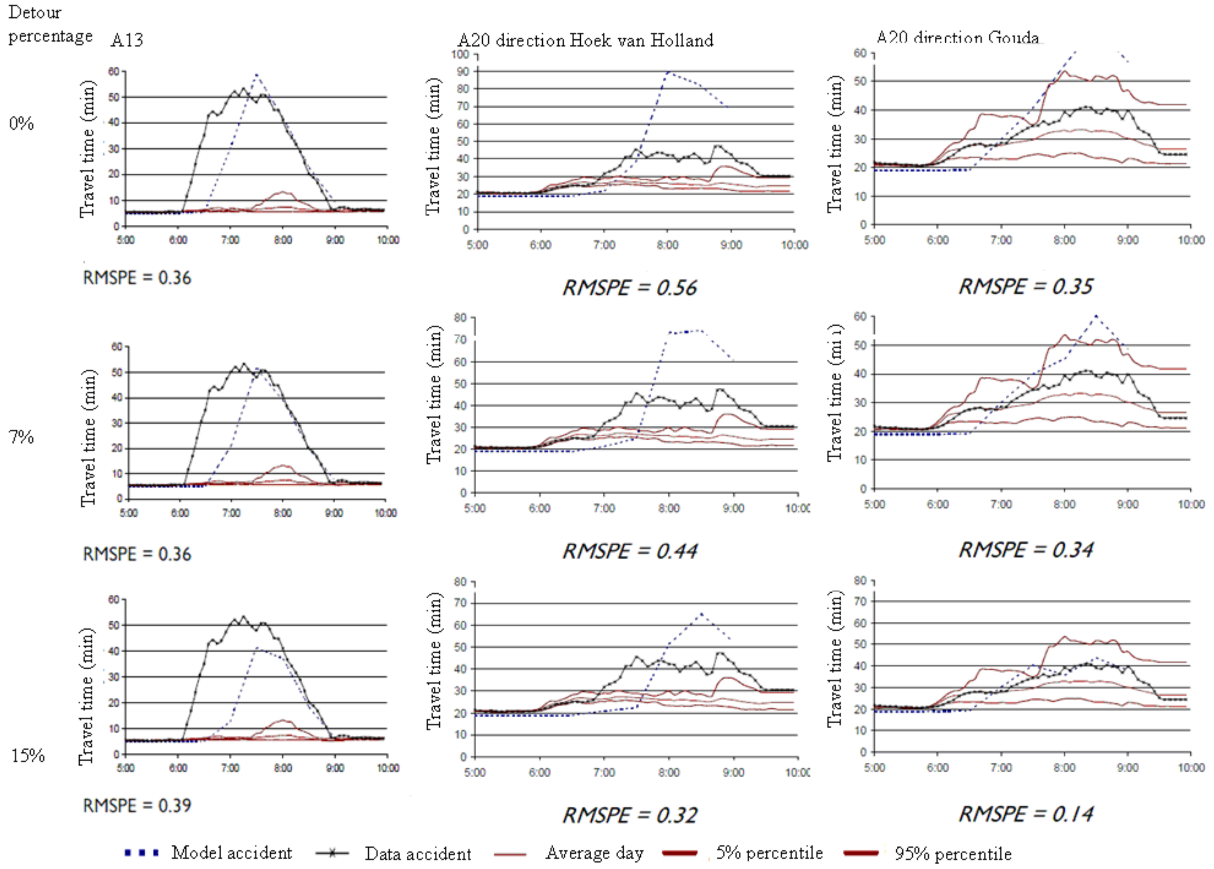


Figure 4.8: Model fit for the travel times on the A20 in the regular situation and during an accident (case 1)

The validation consisted of the comparison of the model data with the real-world data for the two other incident cases. In this comparison, a detour percentage of 15% was used. The results for Case 2 are shown in Figure 4.9 and the results for Case 4 are shown in Figure 4.10.

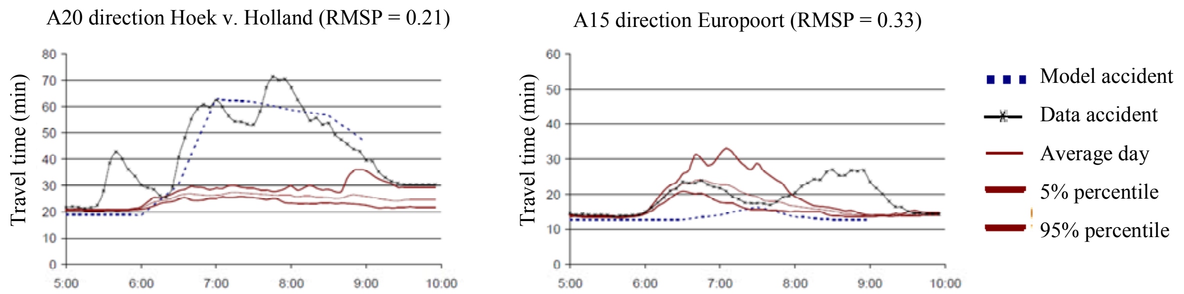


Figure 4.9: Model fit for the travel times on the A20/ A15 during an accident (case 2)

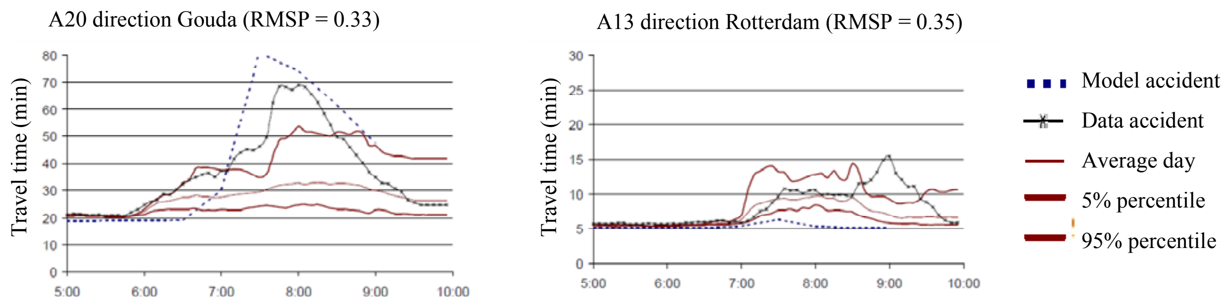


Figure 4.10: Model fit for the travel times on the A13/ A20 during an accident (case 4)

These figures show that the model predicts the delays on the A20 quite well. The delays on the A15 and the A13 are, however, underestimated. The explanation for this is unknown, but must have something to do with the specific circumstances of the accident and the responses of drivers in practice, which might deviate from the model input. Muller (2009) showed that the modelled locations of the queues come quite close to the situation that could be seen in practice. We, therefore, concluded that the spillback effects are quite well represented by the model.

The performance of the MIC-model has not been tested for the same incident situations. However, for the Sioux Network it was shown that the average deviation in vehicle hours lost is only 0.9% compared to incident simulations with fixed route choice in Indy (Corthout et al., 2009).

We can conclude that it is possible to get reasonably good forecasts of the effects of incidents with Indy as long as we know the percentage of travellers that choose an alternative route during incidents. For three incidents that were used for calibration, it turned out that the assumption that 15% travellers take an alternative route was the best assumption[28]. This is in line with the detour percentages (between 0% -50%) that were found by Koo and Yim (1998), Kraaijeveld (2008) and Knoop (2009). For the other two accidents that were used for validation, this also gave reasonably good results. Therefore, we use this assumption in the remainder of the thesis. Nevertheless, as was also explained in section 4.3.3, this percentage could be higher or lower in practice. The assessment method that is presented in this chapter is made in such a way that it can easily deal with other percentages when better data about the percentage of drivers that take detours becomes available. It is, for instance, also possible to vary this percentage for the different incident types that are modelled. The detour percentage could, for instance, be set higher for incidents with a large capacity reduction than for incidents with a small capacity reduction.

Finally, the evaluation method that is presented in this chapter does not make use of the incident approximation algorithms of Indy, but instead uses the MIC module in combination with an alternative route choice algorithm. It was shown that the MIC module approximates Indy quite well for the case in which nobody changes routes. Therefore, it is assumed that this is also the case when the alternative route choice algorithms are used. This assumption is based on the fact that the two methods further only differ in the alternative route choice approximation algorithms. Since both give approximations of the same phenomena, it is likely that the outcomes will not differ more, than the inherent variation in the effects of disturbances due to uncertainty in driver's responses. But, this has not yet been proven.

[28] Higher percentages than 15% were not tried. They might give better results.

4.4 Cost-benefit analysis

When the robustness of road networks is improved by taking robustness measures, a cost-benefit analysis should be carried out to weigh the costs against the benefits. In the current practice of cost-benefit analysis, what is most often lacking is a specific evaluation of the robustness/reliability benefits. However, as was recognized by Husdal (2004), these benefits should be included in a cost-benefit analysis (CBA). The OECD/ITF (2009) reported that unreliability of travel times is increasingly recognized as a major network characteristic that should be recognized when considering investment options. The Working Group found that: "*the number of countries that incorporate reliability in the cost-benefit analysis is limited. Only few countries include monetised values of reliability in their appraisal. Reliability is generally not taken into account when evaluating a project and, as a consequence, CBA will not distinguish between two infrastructure projects where benefits (such as the expected time savings) are identical but where the variability of expected travel time differs. Incorporating reliability into cost-benefit analysis requires data parameters on existing travel time reliability, the anticipated reliability level after a policy initiative, and monetary values of reliability disaggregated at the appropriate level of granularity. Even though there are a number of studies internationally looking at the values of reliability, knowledge on how travel time, congestion and the quality of a network affect reliability are missing. Almost all the existing traffic and transport models used for strategic impact assessment provide predictions of average demand flows and average (or expected) travel times. These models do not, however, provide predictions of the change in the standard deviation of travel time; such changes are building blocks to identifying the impact on reliability of an investment. It is therefore necessary to improve currently available traffic forecasting tools. What is being proposed for incorporating reliability is that the temporal journey time improvement should be split into pure journey time improvement and reliability improvement for each user group, location, etc. Revisions to the cost-benefit analysis approach would therefore involve temporal adjustments, user granulation, and monetisation adjustments. It is emphasised that these estimates will not be transferable, but are country-, location-, user- and time-specific.*"

In this section we propose a method for cost-benefit analysis of robustness measures. It is explained how the evaluation method that is presented in this chapter can be used to overcome a part of the model prediction problems as found by the OECD/ITF. The evaluation method that we propose is restricted by the current motorization possibilities. Therefore, this section first explains what these possibilities are in the Netherlands. As can be read below, the monetarised indicator for reliability that is mostly used is the standard deviation of travel times. The methods that are currently used in the Netherlands to compute the standard deviation are explained as well. Finally, this section shows how the evaluation method that is proposed in this chapter fits in this framework and what the advantages and disadvantages are compared to the available methods.

In the Netherlands, decisions on major infrastructure projects are subject to a cost-benefit analysis according to a standardized framework ─ the OEI-framework. Within this framework, a large number of characteristics of the project are included, such as travel time benefits. Before 2005, there was no opportunity to include improvements of travel time reliability in this cost-benefit analysis. This was perceived as a considerable drawback because, in line with the 2004 policy document "Nota Mobiliteit", many projects in the future will focus on improving the reliability of transport. Therefore the Transport Research Centre of the Dutch Ministry of Transport (AVV/DVS) commissioned RAND Europe to organize a meeting with national and international experts to establish a range of provisional values for reliability to use for policy evaluation in the OEI-framework. One of the outcomes from this study was a reliability ratio (RR) for passenger transport by car for all modes of 0.8. This implies that the value of reliability (VoR) ─ the value of one minute standard deviation of travel time ─ is equal to the reliability ratio (RR) times the value of time (VoT). For freight transport, RAND Europe found that not enough information was available to specify a reliability ratio. However, later a RR of 1.2 for freight transport was determined (Kouwenhoven et al., 2005). The choice for using reliability ratios in the appraisal was implicitly also a choice for using the standard deviation of travel times as an indicator for reliability.

The question that remains is how large the standard deviation of travel times is. In the Netherlands, a rule of thumb is often used: the reliability benefits are equal to 25% of the travel time benefits (Besseling et al., 2004). This 25% can be applied only if there was congestion in the reference network. The 25% cannot be applied to travel time gains as a result of shorter routes or higher design speeds of the network. Of course, this rule of thumb is very simple. In order to compute the standard deviation in a more advanced way, two alternative approaches are suggested in the literature:

- *Economic regression approach*: this approach assumes that reliability gains follow the same patterns as time gains in the networks, and is based on empirical analysis. By regressing indicators for standard deviation of travel time on travel times themselves, and on other relevant indicators, such as traffic volumes or speeds, a statistical relationship is derived that can be used for prediction. Examples are found in (Kouwenhoven et al., 2004) and (Peer et al, 2010).The advantage of this approach is that, if such a relationship exists, a great deal of computation time can be saved, since all of the different disturbances do not have to be simulated individually. A disadvantage of this approach is that it cannot consider all project specific circumstances, such as network structure related aspects.
- *Traffic modelling approach*: in this approach the impacts of different disturbances are forecast by means of traffic models. The advantage of this approach is that it has more explanatory power, because it considers all kinds of behavioural responses of drivers (i.e., all of those included in the model), and is, therefore, project specific. This advantage comes at the expense of higher computation times.

These two approaches are explained in more detail below.

Economic regression approach:

Recently, a great deal of econometric research has been conducted into the relationship between travel times (or travel time delays) and variability. In the Netherlands, the model LMS-BT, which is an add-on to the national model system, focuses on the statistical relationship between speeds and reliability. Eliasson (2006), Fosgerau et al. (2008), and Peer et al. (2010) predicted travel time variability based on empirical data. They all showed that travel time variability is well predictable using econometric models. It was found that variability is positively linked to delays. However, the relation between mean delays and variability turned out not to be linear. Instead, the first derivative of travel time variability with respect to mean delay is decreasing in delay. This pattern can be related to the distribution of traffic regimes. A one minute increase in mean delay is associated with a smaller increase in variability if traffic conditions are more congested during that time of the day. Tu et al. (2007) found that variability is hardly related to flow in a free-flow regime, whereas it is positively correlated with flow in the congested regime, and negatively in a hyper-congested regime. For most CBA's, no information on the relative shares of the traffic regimes is available. Therefore, a non-linear model based on mean travel times can be used as an approximation. The fact that these models consider the aggregate relationship between mean delay and travel time variability implies that the effects of all kinds of circumstances that influence the *average* travel times but not the *variability* of travel times (or the other way around) cannot be taken into account.

Traffic modelling approach:
In order to look at specific circumstances, a model is needed. For the Netherlands, there is a model called SMARA (Meeuwissen et al., 2004; Appendix E). In this model, Monte Carlo simulation is used to simulate the effects of variations in demand and supply. This model can analyse the effects of disturbances that vary from local to network-wide disturbances, from disturbances with a small probability to disturbances with a large probability, from supply related to demand related disturbances, and from disturbances with a small effect to disturbances with a large effect. In this way, the complete travel time distribution can be analysed. Since SMARA is a static model, the computation time stays within acceptable boundaries as well (15 minutes up to several hours on a regular PC). Of course, the fact that a static model is used is a disadvantage as well, because network dynamics are not captured and spillback effects are ignored. The underestimation of spillback effects is especially a shortcoming for forecasting the effect of large disturbances. Snelder and Tavasszy (2010b) used this model for the evaluation of the reliability benefits of the measures presented in the Dutch policy paper the "MobiliteitsAanpak" (Dutch Ministry of Transport, Public Works and Water Management, 2008). They showed that the reliability benefits of the presented set of measures for the road network are about of the same size as the travel time benefits of these measures under regular circumstances. This implies that the rule of thumb of 25% could result in a significant underestimation of the benefits. On the other hand, it was shown that the travel time benefits can vary significantly over different projects included in the set of measures, which shows that it is incorrect to apply a general rule of thumb to all projects. Of course, modelling techniques other than Monte Carlo simulation are possible. The method that is presented in this chapter is one of them. It uses dynamic traffic assignment instead of static assignment, which is a big advantage. However, it can consider only local variations in supply.

Choice of method for including the effects of incidents in cost-benefit analysis
In the valuation of robustness benefits, the standard deviation and the rule of thumb are not very useful. The first refers to uncertainty in travel times over a longer period. Since we look at specific incident situations we can't get a complete view on the standard deviation. The second one is too simple and does not closely relate to travel time losses as a result of incidents. Therefore, we propose to use value of time to value the travel time losses caused by incidents. Since there is no generally excepted value of time for unexpected travel time losses, we choose to use an average value of time of €15 per hour. This assumption can be varied in a sensitivity analysis, in order to investigate its effect on the results.

In what follows, we use an example to show the difference between using the method of standard deviations and the method proposed above. In Figure 4.11**,** a fictive travel time distribution is shown with an average travel time of 8.4 minutes, a median travel time of 6.3 minutes, a mode of about 1 minute (if travel time classes of 1 minute size are considered), and a standard deviation of 7.1 minutes. On the vertical axis, the number of trips is shown. In total, 16507 trips were made.

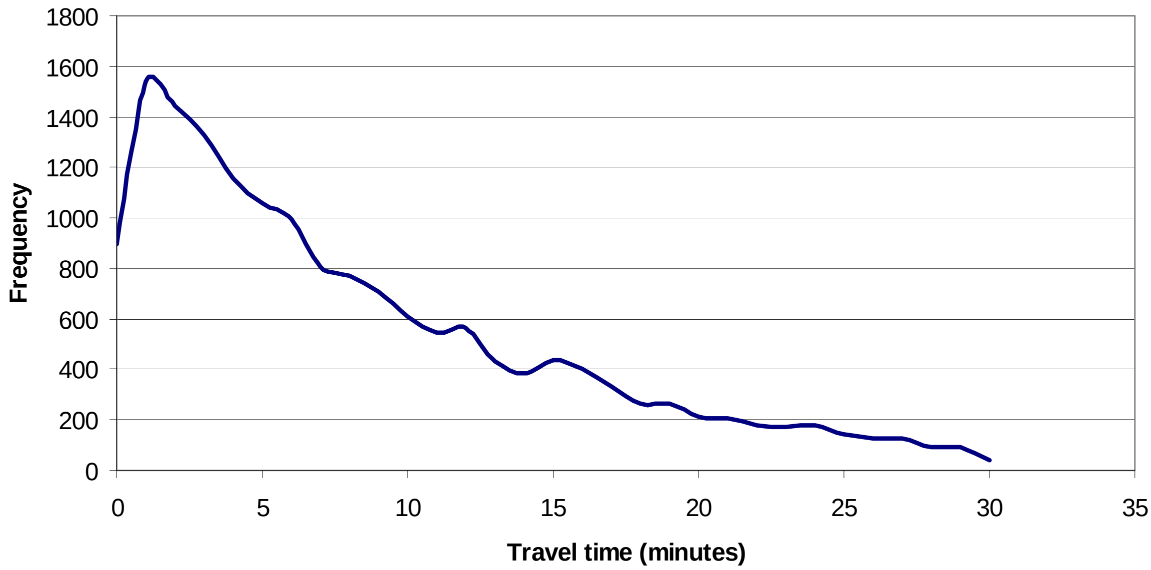


Figure 4.11: Travel time distribution used in the example

The total expected travel time costs are about €34,600, given a value of time of €15 per hour (16507*15*8.4/60). It could also be argued that the total expected travel time costs are about €26,000 (16507*15*6.3/60) if we set the expected travel time equal to the median travel time, and only €4,000 if the expected travel time equals the most frequently occurring travel time class of 1 minute.

The total reliability costs are about €23,400, if the method of standard deviations is used with a reliability ratio of 0.8 for all trips. If we compute the reliability costs as the absolute difference between the realized travel time and the average travel time multiplied by the average value of time and the number of trips, the costs of unexpected delays are equal to €23,800. As can be seen, this is in the same order of magnitude as the reliability costs computed by using the standard deviation. The costs of arriving early are included as well in this computation. If these costs are excluded, the costs of unreliability are €11,900, which is significantly less than the €23,800 estimated above. In this case, travel times are valued equal for arriving late, arriving on time, and arriving early, despite the fact that it is commonly assumed and empirically proven that the value of arriving early is valued lower than the value of time, which is, in turn, valued lower than arriving late.

If the costs of unreliability are computed based on the difference between it and the median travel time or the most frequently occurring travel time, the costs of unreliability are respectively €22,900 and €30,900. The latter case is not unrealistic, since a reference day (which is used in models for analysing the travel times under regular conditions) might very well be a day with circumstances that occur most often. If this is the case, the method of standard deviations underestimates the reliability costs (actually, the costs of unreliability) significantly. Of course, different distributions result in different numbers, and, if another reliability ratio is used, the differences between the two methods will change as well.

Based on this analysis and the above mentioned arguments, in the remainder of this thesis we will use the method in which the difference in travel times in the situation with and without accidents is valued.

Recommendation for taking all disturbances into account in cost-benefit analysis
As explained above, there is currently a rule of thumb that the reliability benefits are equal to 25% of the travel time benefits. The common line of thought at the moment is that vulnerability/robustness benefits need to be added to this. However, this results in the question of how to avoid double counting. In other words, which effects are already included in the 25% rule? In our opinion, the difference between robustness and reliability cannot be made in this way. First of all, the terms robustness and reliability are confusing when they are used like this, since robustness improving measures can lead to reliability benefits. Therefore, in the remainder of this thesis, we refer to the benefits of robustness measures as robustness/reliability benefits. Instead of trying to combine reliability and robustness measures, we propose to use a disturbance based approach. This enables us to distinguish between different values of time for different kinds of disturbances, and it enables us to give special attention to certain types of disturbances. It might, for instance, be necessary to forecast disturbances with a high expected effect (in the tail of distribution function) with more advanced models than the delays of small disturbances. However, this does require that different values of time for different disturbances or effects of disturbances are available, that a distinction in the data can be made between different disturbances, and that the effects of different disturbances can be modelled.

With respect to the available data in the Netherlands, it can be concluded that there are a lot of data available, ranging from traffic counts to weather registration, incident registration, etc. In order to monitor the robustness of a network against different types of disturbances, it is necessary that all those different sources can be linked to each other in the future, and that data are available for the motorways and the secondary network as well. The fact that a National Data Warehouse for Traffic Information (NDW) is being created in the Netherlands is an important step forward in this direction.

With respect to the modelling part: the SMARA approach (or similar methods) for forecasting the effects of roadworks, bad weather conditions, incidents, events, and seasonal variations in demand can already be used to model a very large part of the travel time distribution. Based on the outcomes of the model, many indicators, such as the standard deviation and the vehicle loss hours as a result of disturbances, can be computed. It is also possible to exclude the effects of incidents and to model them in a more advanced way, as described above. A drawback with this is that it requires that two types of models are used: SMARA and a dynamic traffic assignment model (and maybe even a third one: a transport model that considers the demand effects as well). Therefore, in the future the methods should be more integrated.

Two additional remarks need to be made:

- In a cost-benefit analysis, the travel time benefits for the original traffic and the newly generated traffic are computed by using the 'rule of half'. The rule of half estimates the change in consumer surplus for small changes in supply. The question is whether or not improved reliability or decreased vulnerability will result in newly generated traffic and whether or not the rule of half should then be applied. This might be the case; however, little is known about this phenomenon. Because of this, and the computational difficulties in applying the rule of half to the reliability/robustness benefits as computed by the method described in this chapter (they are not computed on the origin-destination level, but only as an aggregate number), we do not apply the rule of half. As long as no additional traffic is generated, this is not a problem. In the case when additional traffic is generated, this results in an underestimation of the reliability/robustness benefits if the robustness of the network improves, because this is implicitly to equal to the assumption that the newly generated traffic did not have unreliability costs before they decided to use the road. If robustness measures result in improved travel times under regular conditions, and in newly generated traffic as a result of that, the additional travel time benefits under regular conditions are included for this new traffic by using the rule of half.

- Improvement in the robustness of the network might have an indirect impact outside the transport market. It could result in different location choices of companies, and people might move. We do not consider the benefits from these kinds of agglomeration forces. However, they might be substantial.

4.5 Summary: framework for assessing the robustness of road networks

In this chapter, an evaluation method was presented that can be used to evaluate the robustness of a road network. This chapter answers the 7th research question: How can robustness be evaluated?

We showed that the dilemma in the model choice for evaluating robustness is we must choose between accuracy (ideally, using a dynamic traffic assignment model with detailed congestion modelling, including spillback effects, multiple types of route choice behaviour during incidents, and accurate intersection modelling) and computation time. In general, the most accurate models take the longest computation time. Using a rule of thumb, for instance, takes hardly any computation time, but is not accurate or detailed. For some applications, a rule of thumb can be good enough to get a quick impression of the robustness of a network. However, to make a well balanced decision about robustness measures in network design, it would be better to look for a method/model that deals with the above mentioned requirements in the best possible way.

We chose to use the dynamic traffic assignment model Indy to compute equilibrium under regular circumstances, because this model has an accurate network loading model that models spillback effects according to the simplified kinematic wave theory of Newell. Furthermore, the model can compute equilibrium route choice and can deal with fixed route choice. Router choice en-route is not possible. However, this is not a problem, since Indy is used only for a basic run in a situation without disturbances, for which an equilibrium assignment is the most appropriate. Indy is a dynamic model, which makes the modelling of time dynamics possible. And, since Indy is able to work with the marginal incident computation model (MIC) (Corthout et al., 2009), it is able to get an estimate of the impact of incidents very quickly. Four incident types were defined, for which the effects are computed for all links. An approximation algorithm is added to the MIC-module to deal with the use of alternative routes. The quality of the method was shown by making a comparison with loop detector data for five incidents. Finally, it was shown how robustness/reliability benefits of robustness measures can be computed with this evaluation method and how they can be included in a cost-benefit analysis.

5 Measures for improving the robustness of a road network

5.1 Introduction

In the previous chapter an evaluation method for robustness was presented as well as a method for including robustness/reliability benefits of robustness measures in a cost-benefit analysis. The questions are: what measures can be taken to improve robustness, and what is their impact on robustness?. This chapter presents a list of measures that can be taken to improve the robustness of a road network. In the sections following the list of measures, we give an indication of the possible effects for the measures that are the most likely to reduce a road network's vulnerability.

5.2 General description of measures

Robustness in the traffic and transport system can be considered according to various categories. A common categorisation divides the transport system into three submarkets (Van de Riet, 1998). Attempts can be made to achieve a more robust situation for each of these three markets: a travel market, a transport market, and a traffic market. Depending on the submarket, different time-scales apply.

- In the *travel market*, the locations of activities are defined. Where will people come to live? Where will they work? Where will companies be located? The demand for activities and the supply of activities in space and time lead to a movement pattern. In the travel market, for example, robustness can be striven for in spatial developments, which will ensure that choices that are made now will provide longer-term benefits. The appropriate time-scale for this covers several years.
- The *transport market* is the market where the demand and the supply of transport possibilities lead to a transport pattern. In other words, this is the market in which the mode choice is made for freight and passenger transport. Questions like, do I go by bike,

by public transport, or by car to my work? In the transport market, different ways of transport complement each other and offer the consumers choice options. Furthermore, to a certain extent, the different modes can function as back-up options for each other. The appropriate time-scale for this market covers several days.

- The *traffic market* is the market in which the demanded transport patterns are confronted with the supply of infrastructure and the traffic management systems linked to this infrastructure (e.g. Dynamic Traffic Management, information systems, traffic light regulation), which results in an infrastructure that is actually used, in the form of a pattern of movements. In the traffic market, a choice has already been made for the mode of transport, and robustness ensures that vehicles or travellers experience few problems with disruptions. This can be achieved, for example, by making alternative routes available.

An alternative to the three-market categorisation of measures is to divide them into strategic, tactical, and operational measures. Another categorization is to characterise them by the element of robustness (prevention, redundancy, compartmentalization, resilience, and flexibility) that they relate to the most. Finally, the measures could be categorised by the type of disturbance they are taken to mitigate.

To the best of our knowledge, there is little literature available about measures that could be taken to improve the robustness of a road network. However, Klem et al. (2004) present examples of measures that could improve the robustness of road networks. In total, 40 measures are described, together with estimates of their effects and costs. These measures are subdivided into measures on lanes, road stretches, roads, networks, and organizational measures. Examples are using VOAC (very open asphalt concrete), improving incident management strategies, and using lower level roads as backups for higher level roads. Below, we present a list of measures that influence the robustness of the road network. The measures are classified in terms of the component of robustness on which they have the most influence: prevention, redundancy, compartmentalization, resilience, and flexibility. If a measure influences more than one component, this is indicated. Additionally, an indication is given as to which market (travel market, transport market, traffic market) they affect and, where relevant, to which type of disturbance they relate, and if they are on the strategic, tactical, or operational level. The measures mentioned in (Klem et al., 2004) are assigned to this classification. Some measures are combined and others are added. The list presented, is not intended to be complete. However, the most relevant measures are likely to be included.

Prevention

1. Construct a road high enough to make sure that traffic can continue to move in times of *floods*. This measure intervenes on the *traffic market* and is of a *strategic* nature.
2. Allow no trees in the close surrounding of the road, so in a *storm* no trees (or parts of trees) can be blown over that will block the road. This measure intervenes on the *traffic market* and is of a *strategic* nature.
3. Use more porous asphalt (ZOAB). Porous asphalt roads cause the capacity to be reduced less during *rainfall*. This measure intervenes on the *traffic market* and is of a *strategic* nature.
4. Spread salt on roads so that glazed *frost* and *snow* lead less quickly to problems. This measure intervenes on the *traffic market* and is of a *tactical* nature.
5. Close the schools in case of *very bad weather conditions*. This measure reduces the trips from and to schools and forces parents to stay at home to stay with their children. This measure is sometimes taken in Canada and in other countries with frequently occurring

bad weather conditions (like *heavy snowfall*). This measure intervenes on the *travel market* and is of a *tactical* nature.

Redundancy

6. *Spread out activities*. In the past, a varied spatial planning policy was implemented that was geared towards spreading or concentrating. Both contrasting strategies can have a positive or a negative effect on robustness. Spreading leads to a balanced distribution of the traffic along the roads. If something happens on one of the roads, another road can provide a route alternative. However, the network has to be geared towards this. If origins and destinations are concentrated, public transport will become more attractive. This measure is not designed to reinforce public transport but to create a robust system by offering an alternative (redundancy). The disadvantage of this measure is that, on the roads, large flows are bundled as well. If a disruption occurs here, the effect can be huge. The question is whether a public transport system can deal with this. The spreading of activities appears in the first instance to have a larger positive effect on the robustness of the traffic and transport systems. To be able to make a statement with any certainty, however, a more elaborate analysis must be carried out in which choice of destination and means of transport are also included. The spreading of activities can, for example, lead to a higher share of car movements, resulting in the roads becoming busier and the network becoming less robust. The effects on trip distances are also unknown. This measure intervenes on the *travel market* and *transport market* and is of a *strategic* nature.
7. *Better alignment, abolition, or liberalization of time windows*. Time windows cause an increase in the number of goods kilometres. As a result of time windows, more trucks must be made available in order to stock the shopping centres in a short period of time. The fact that more trucks are used makes the distribution more robust. If one truck is affected by a disruption, the other trucks can, in certain circumstances, compensate for this. However, the fact that the stocking has to take place in a short space of time makes the situation less robust, because switching to other times is not possible. The liberalization or abolition of time windows, therefore, leads to a more balanced distribution over the whole day. Additionally, the spare capacity (redundancy) that is present in the network outside the original time windows can be made use of, so the flexibility of the transporters increases. This measure intervenes on the *travel market* and is of a *tactical* nature.
8. Many companies have fixed working hours. This means that employees are not flexible with regard to their choice of departure time, so that they are unable to switch to another time. *Making working hours flexible and/or allowing or making telecommuting possible* could increase the reliability of travel times, and could lead to a more balanced distribution of the traffic across the network. A consequence could be that the space on the road that arises in the peak period as a result of employees travelling outside the peak period will be filled by ‘new motorists’. This can result in the spare capacity (redundancy) in the network being reduced, which reduces the robustness. This is not necessarily negative, because the additional demand also results in economic growth. Finally, this measure will increase the flexibility of the employees. This measure intervenes on the *travel market* and is of a *tactical* nature.
9. *Create equivalent routes whereby back-up options have the same capacity as major routes*. Equivalent routes can be created by the “urban/regionalisation” of motorways that are used primarily by urban/regional traffic in certain urban areas. Additionally, equivalent routes can be created by, where necessary, upgrading the underlying road network, restructuring intersections and junctions, and realizing a number of crucial supplementary

connections, predominantly within the urban/regional main structure. This measure intervenes on the *traffic market* and is of a strategic nature.

10. *Introduce a pricing policy that differentiates according to time and place*. A pricing policy is particularly effective in the regular situation (i.e. incident-free). By introducing a pricing policy with a differentiation according to place and time, the peak periods are expected to become longer but less intensive. In this way, the roads will remain busy for a longer time, so less spare capacity is present over time. In the case of disruptions, this can result in a disruption taking longer before its effects are resolved. However, the peak will be a bit less high, so at the busiest time there will be more space to deal with disruptions (redundancy). This applies only if no extra traffic is generated as a result of the levels of the pricing measure. The effect of a pricing policy on robustness is, therefore, difficult to predict. This measure intervenes on the *travel market* and is of a *tactical* nature.
11. *Introduce more robustness in the logistic chain of goods*, for example through:
    - extra stocking of (bottleneck) products,
    - building in reserves in production capacity and vehicle capacity,
    - building in flexibility in processes,
    - dual or triple sourcing: this means that raw materials or services are provided not by one supplier, but by two or three suppliers. These are suppliers of products but also suppliers of logistical services,
    - geographical spreading of suppliers,
    - selecting suppliers that are themselves resilient,
    - decentralization instead of centralization of own locations.

    By building in more robustness in the logistic chain, disruptions on the road network can be dealt with better. In the logistics field, often the term 'resilience' is used. This term is used because in logistics it is of great importance to recover from disruptions as quickly as possible. Redundancy, flexibility, and balance are factors that help achieve this. This measure is placed under redundancy because robustness in the logistic chain is necessary as a back-up option (redundancy) for disruptions on the road network. This measure intervenes primarily on the *travel market* and is of a *strategic/tactical* nature.
12. *Separate through traffic and urban/regional traffic through 'unbundling'*. The advantage of unbundling is that disruptions on the major roads can be dealt with on the parallel roads and vice versa. It is, however, important that thought is given to the way in which the unbundling process occurs. The distance between the places at which the exchange between both roads is possible must be carefully chosen. If the distance is too short, congestion spillback will occur on one road as a result of congestion caused by an incident on another. To allow this measure to work best, flexibility can be necessary in the form of, for example, flexible short cuts. When incidents occur, these flexible short cuts make it possible to switch roads at points where this is not possible under normal circumstances. This measure intervenes on the *traffic market* and is of a *strategic* nature.

Compartmentalization

13. *Introduce buffers to prevent congestion spillback and to regulate the inflow of traffic.* (In buildings, compartmentalization is created by using fireproof doors.) This measure intervenes on the *traffic market* and is of a *strategic/tactical* nature.

Resilience

14. *Introduce incident management*. The introduction of incident management can lead to a quicker resolution of incidents, so that the resilience – and therefore the robustness – of the network increases. This measure intervenes on the *traffic market* and is of a *tactical/operational* nature.

15. *Provide travel information*. By informing companies and motorists better about disruptions, the consequences of such disruptions, and the alternatives available, robustness can be increased, because the network will recover more quickly (resilience). This can be achieved through:
    - improved traffic management. Through a better collaboration between road administrators, better information can be provided and management can be improved;
    - a better connection between in-car systems and road management systems (central offices). This leads to the information from the road administrator reaching the road user better and more quickly.

    This measure intervenes on the *traffic market* and is of a *tactical/operational* nature.

Flexibility

16. *Introduce flexibility in the network*. Flexibility can be introduced through, for example, reversible lanes, short cuts that can be closed off, and flexible intersection design. Flexibility can also be created through constructing road lanes that may only be used in exceptional situations. It is, however, difficult to explain to road users that these road lanes may not be used at times of regular congestion. This measure intervenes on the *traffic market* and is of *strategic/tactical/operational* nature.
17. *Develop Park and Ride (P+R) facilities to link the road network with the public transport network*. P+R facilities can help reduce the effect of incidents. However, the capacity of the public transport network is not big enough to function as a backup option for the road network. Besides redundancy, this measure also increases the choice options of travellers in regular situations, which creates more balance in the transport network. This measure intervenes on the *transport market* and is of a *strategic* nature.
18. *Increase the use of hybrid transport networks/co-modality*. Companies should be able to make use of several networks for their transport. With little loss of value, goods can, for example, be transported via the inland waterways, and the transport of high-value goods can be done via the road network. Furthermore, the stable part of the volumes that have to be transported can be transported via inland waterways on regular basis. The fluctuating part of the volume can then be transported by road. If something happens to one of the two networks, the facilities are available for the transport to be carried out via the other network. This measure intervenes on the *transport market* and is of a *strategic/tactical* nature.

Some of the above-mentioned measures have both advantages and disadvantages. This means that customized measures must be formulated. The effectiveness of some of these measures has already been proven. Immers et al. (2004d) have shown the importance of a cohesive underlying road network. Schrijver et al. (2008) have demonstrated the effectiveness of a combination of these measures. They have proven that the time loss as a result of incidents can be reduced by almost 30% by making a physical distinction according to functions (interregional traffic, urban/regional traffic, and urban traffic), by offering choices, by taking into account the flow- and buffer function of the network, and by unbundling the roads. Finally, Snelder et al. (2010c) explain in detail how the robustness of a transport system can be improved by offering 'robust park and ride facilities'. In the following sections, selections of structure related measures in the following categories are worked out in more detail:

- Creating alternative routes (section 5.3).
- Unbundling traffic flows (section 5.4).
- Creating buffers and spare capacity (section 5.5).

We chose to focus on these categories of measures because they are structure related measures on the traffic market with a strategic nature, which matches best the focus of this thesis. This does not mean that the other measures are irrelevant. It is, for instance, very important to develop good dynamic traffic management strategies (including offering the right information to the right persons at the right time) and incident management strategies, in order to make the best use of the more robust structure of the network.

5.3 The importance of alternative routes

It is clear that having route alternatives is very important for the robustness of a road network. However, having alternatives is not enough; there should also be spare capacity on those alternatives. In this section, this hypothesis is tested. The importance of alternative routes for the robustness of the network is shown by doing a case study for Rotterdam. This section starts by explaining the characteristics of the network that is used for the computations. It is the same network as presented in section 2.2. However, for reasons of clarity, the characteristics are repeated in this section. Thereafter, the method that is used for assessing the importance of alternative routes is described. Finally, the results are presented.

Rotterdam is the second largest city of the Netherlands. In 2007, it had about 590,000 inhabitants. Rotterdam is surrounded by 4 motorways. Together they are the called 'Rotterdamse ruit'. The network of Rotterdam is shown in Figure 5.1.

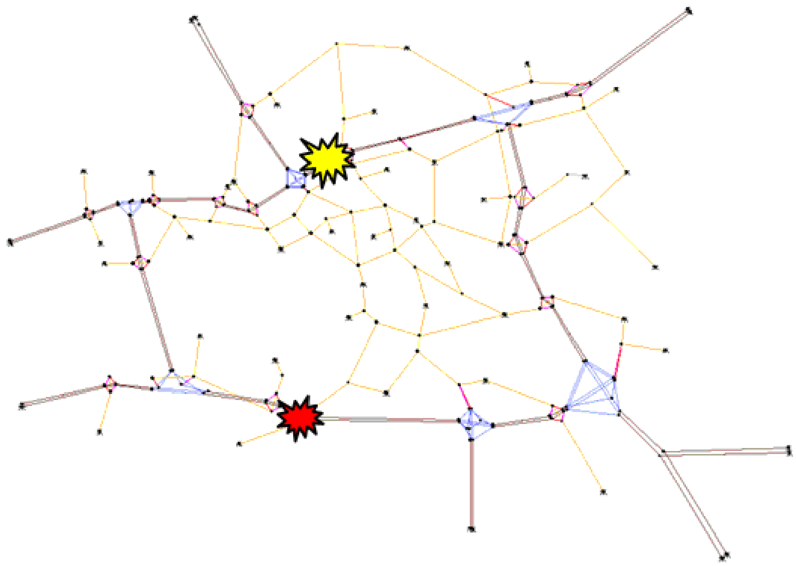


Figure 5.1: Road network of Rotterdam and surroundings

The network that is used in the model contains 51 centroïdes. In total there are 1890 OD-pairs (not all combinations of centroïdes are used), and 7674 paths were generated. Furthermore, the network contains 239 nodes and 570 directed links, of which 468 links are regular links and 102 links are feeder links. The total average OD demand per hour is 80,477 trips. The total demand period used in the case was 3 hours. To obtain the demand for each 10 minutes, the hourly demand was multiplied by the departure fractions that are presented in Table 5.1. The rows in the table contain the departure fractions for each set of 10 minutes in an hour. In the columns, the hours are presented for the demand period from 6.30h -9.30h. In total, for this case, 158,172 trips are made in the morning peak in the network of Rotterdam.

Table 5.1: Departure fractions

	6 - 7 h	7 - 8 h	8 - 9 h	9 – 10 h
x.00 - x.10	-	0.5	1.0	0.4
x.10 - x.20	-	0.7	1.0	0.3
x.20 - x.30	-	0.9	1.0	0.1
x.30 - x.40	0.1	1.0	0.9	-
x.40 - x.50	0.3	1.0	0.7	-
x.50 - x.00	0.4	1.0	0.5	-

In order to examine the importance of alternative routes for the robustness of a road network, 18 simulations were carried out twice. In fact, there are 6 different scenarios for which equilibrium assignments with Indy with and without an incident were carried out. Furthermore, for these 6 different scenarios, two other runs with an incident were simulated: one with fixed routes and one with a new equilibrium. The MIC-module was not used in this analysis because, due to the small number of model runs that had to be made, saving computation time was not needed. As was explained before, doing an Indy simulation with fixed route choice results more or less in the same results as doing an analysis with the MIC-module. However, using the fixed route choice option of Indy has the advantage that, for instance, downstream effects of the incident are not ignored. The approximation methods for alternative routes are not used either, because that would conflict with the method described below for analysing the importance of alternative routes. Finally, two different incidents were analysed (two incidents, 6 scenarios, three runs for each scenario).

The six different scenarios are constructed in the following way. We focused on the through traffic on the Rotterdamse Ruit. This is traffic that does not have an origin or a destination in Rotterdam. In fact, this is the long distance traffic for which the motorways were originally designed. Furthermore, this is the traffic that cities do not want to have on their local networks. In the original network, the through traffic has the opportunity to deviate from the motorways and to take local routes through the city. When an incident occurs, there will be more such traffic. We examined what would happen if this local network were not available as a back-up option for the motorway network. For the OD-relations with on origin and destination on the motorway edges of the network (these are external zones marked with a (green) star in Figure 5.2), the paths were examined. The number of paths varies between 0 and 6. For some OD-relations there is no demand and there are no paths, and for 1 OD-relation there are 6 paths. The other OD-relations have a number of paths in between these values. These paths are ordered according to their lengths on the motorways. In the first scenario, all paths are included. In the second scenario, at most 5 paths are included, and the path with the shortest length on the motorways is deleted. In the third scenario, at most 4 paths are included, and the two paths with the shortest lengths on the motorways are excluded. In the fourth, fifth, and sixth scenarios respectively, at most 3 paths, 2 paths, and 1 path are/is included. This implies that, in the last scenario, the through traffic has no route alternative, and therefore cannot change routes when an incident occurs.

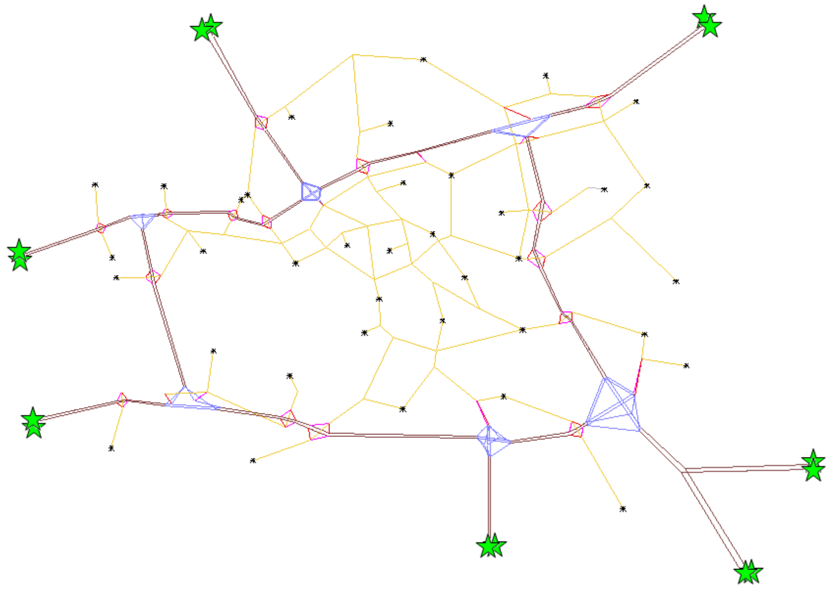


Figure 5.2: External zones, origins, and destinations of through traffic

The first incident that is simulated is on a location that is usually congested during the peak period and, therefore, likely to be vulnerable. The link is indicated with a yellow/gray star in Figure 5.1. It is an incident on the north side of the A20, just before the connection with the A13. The incident is simulated in the direction towards the A13. The incident occurs at 7.45h, and blocks all lanes for 1/2 hour. Thereafter, the emergency services have moved the vehicles that were involved in the incident to the side of the road, and only one lane is blocked. Fifteen minutes later, the capacity is fully restored. In Indy, this incident is simulated by reducing the number of lanes and the capacity. In the first half hour after the incident, the capacity is reduced from 6600 passenger car units (pcu) per hour to 0 pcu/hour, and the number of lanes is reduced from three lanes to no lanes. In the 15 minutes thereafter, the two lanes are in use again, with a total capacity of 3300 pcu/hour. Thereafter, the number of available lanes is back to three, and the capacity is fully restored to 6600 pcu/hour.

The second incident that is simulated is on an off ramp of a location that is usually not congested during the peak period, and therefore is less likely to be vulnerable. The link is indicated with a red/black star in Figure 5.1. It is an incident on the south side of the A15, in the direction towards the A16. The incident occurs at 7.45h, and blocks the off-ramp completely for 1/2 hour. During the next 15 minutes, the capacity is partly restored and, thereafter, fully restored. In Indy, this incident is simulated by reducing the number of lanes and the capacity. In the first half hour after the incident, the capacity is reduced from 3000 pcu/hour to 0 pcu/hour, and the number of lanes is reduced from one lane to no lanes. In the 15 minutes thereafter, the off-ramp is in use again, but with a capacity of only 1500 pcu/hour. Thereafter, the capacity is fully restored to 3000 pcu/hour.

The first indicator for robustness that was examined is the total travel time in the network. In Table 5.2 (for the first incident) and Table 5.3 (for the second incident)**,** the total travel time is shown for the 6 equilibrium runs without an incident, and the increases in travel time compared to these equilibrium runs are presented for the runs with an incident.

Table 5.2: Total travel times in the 18 runs for the incident on the A20

Scenario number	Maximum number of paths per OD relation	Total travel time (TTT) equilibrium without an incident (minutes)	Increase in TTT incident with fixed route choice	Increase in TTT incident with new equilibrium
1	6 paths	22031862	82.7%	3.2%
2	5 paths	22032254	77.1%	3.1%
3	4 paths	22030672	77.5%	3.1%
4	3 paths	22019479	78.7%	3.1%
5	2 paths	22030681	77.1%	3.2%
6	1 paths	22779628	102.9%	4.3%

Table 5.3: Total travel times in the 18 simulation runs for the incident on the A15

Scenario number	Maximum number of paths per OD relation	Total travel time (TTT) equilibrium without an incident (minutes)	Increase in TTT incident with fixed route choice	Increase in TTT incident with new equilibrium
1	6 paths	22031862	5.6%	0.8%
2	5 paths	22032254	5.6%	0.8%
3	4 paths	22030672	5.8%	0.8%
4	3 paths	22019479	5.8%	0.8%
5	2 paths	22030681	5.9%	0.8%
6	1 path	22779628	13.7%	9.3%

From both tables it can clearly be seen that an incident can have large consequences when nobody has a route alternative or is not informed about the incident. This is the situation in which everybody chooses the same route in the situation with the incident as they would choose in the situation without the incident. For the incident on the off-ramp of the A15, in the first four scenarios this impact is about 5.6%; for the fifth scenario, the impact is already 5.9%; and in the last scenario there is an increase to 13.7%. For the incident on the A20, the effects are much more extreme. In the first five scenarios there is an increase of 77.1%-82.7% , and in the last scenario the total travel time more than doubles.

In the other extreme case, where a new equilibrium occurs in the incident situation, the impact is less significant. In the case of an incident on the off-ramp of the A15, in most of the scenarios there is an increase of about 0.8% in the total travel time. In the case when the through traffic has no route alternative, the impact is 9.3%. For the incident on the A20, the delay in the first five scenarios is 3.1%-3.2% of the total travel time. In the last scenario, the delay is 4.3% of the total travel time. Given the fact that the incident does not influence all the traffic (not all the cars are confronted with queues caused by the incident, because some drive in an unaffected part of the network), and also does not influence the traffic during the complete demand period, these percentages may be considered quite high.

Some individual travellers are likely to experience much more delay. This hypothesis was tested by making plots of the average travel time for all OD-relations that have at least 1 path that uses the link on which the incident occurs. The plots are shown in Figure 5.3 and Figure 5.4. The plots are presented only for the first and last scenario, because the other scenarios are quite similar to the first scenario. From these figures, it can be seen that the travel times take extreme values in the fixed route simulations. For the scenario in which all paths are included, some travellers experience a delay of 69 minutes when an incident occurs on the A20. In the scenario with no route alternatives for through traffic, some travellers experience a delay of 81 minutes when the same incident occurs. In the situation without an incident, the same travellers would have had travel times of 25 minutes and 28 minutes, respectively. In the simulations with a new equilibrium, the delay is at most 7 minutes in the scenario in which all paths are included, and at most 9 minutes in the scenario in which no route alternative is available for the through traffic. This does not seem to be much. However, it is still an increase of respectively 29% and 34% in travel time for some travellers.

The incident on the off-ramp results in delays of a maximum of 15 and 34 minutes for the fixed route runs in the first and last scenario respectively. For the equilibrium runs, the delays are 2 and 26 minutes. These differences clearly show that it is important to have route alternatives.

From Figure 5.3 and Figure 5.4, it can also be seen that the effects of incidents remain visible long after all the traces of the incident are removed (8.30h). This is caused by congestion effects. Finally, these figures show that the travel time can, in the worst case, increase by an hour within half an hour. This implies that part of the network loses its function quite rapidly, which is a sign that the network is not robust.

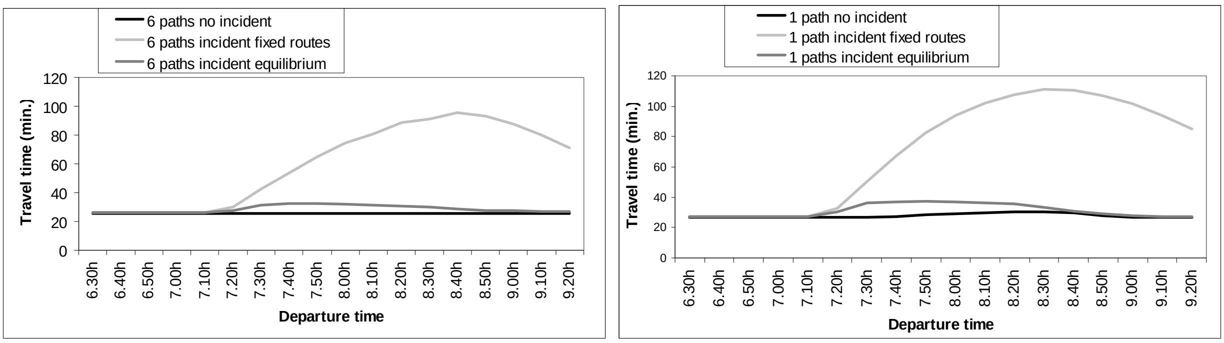


Figure 5.3: Incident on the A20 ─ Average travel times for OD-relations, of which at least one path uses the link on which the incident occurs. (Scenario 1 (left), Scenario 6 (right))

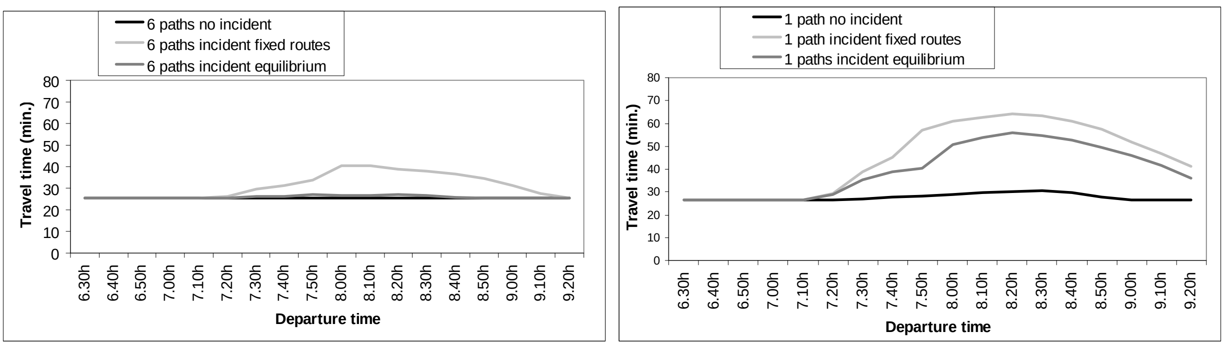


Figure 5.4: Incident on the A15 ─ Average travel times for OD-relations, of which at least one path uses the link on which the incident occurs. (Scenario 1 (left), Scenario 6 (right))

Finally, in Figure 5.5 and Figure 5.6, the situation for the two cases at 8.15h (half an hour after the incident occurred), is shown for Scenario 1 and Scenario 6. From the equilibrium situation without an incident, it can be seen that, at this point in time, there is hardly any congestion. The plots for the fixed route simulations show that congestion builds up quickly. The process continues until long after 8.15h, but it can already be seen that the congestion builds up a bit faster in Scenario 6. Finally, the equilibrium plots for the situation with an incident show that there clearly more congestion in Scenario 6 compared to Scenario 1.

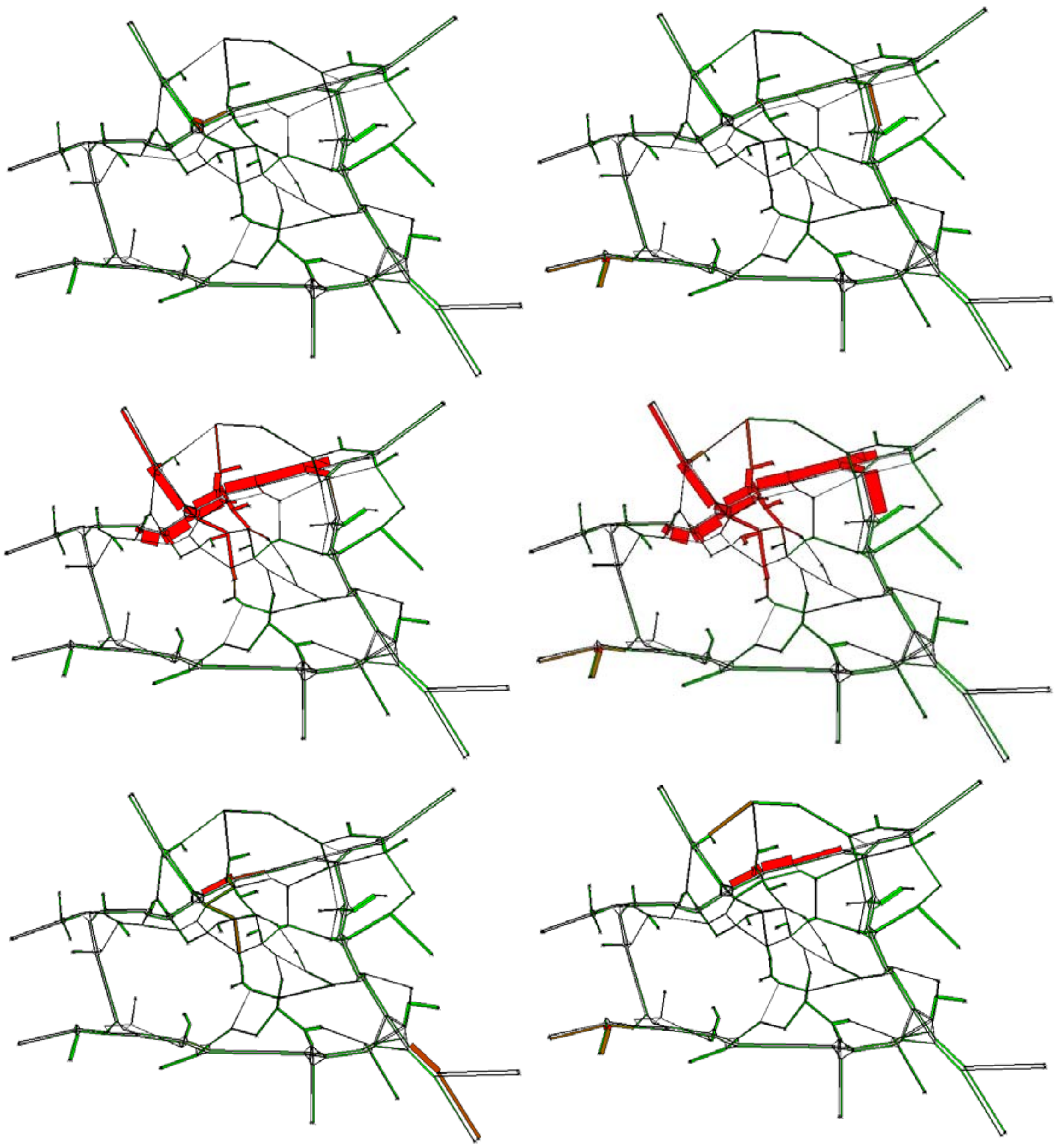


Figure 5.5: Incident on the A20 – Situation on the network at 8.15h in the equilibrium assignment without an incident (top), with an incident with fixed routes (middle), and with an incident with a new equilibrium (bottom) for Scenario 1 (left) and Scenario 6 (right)

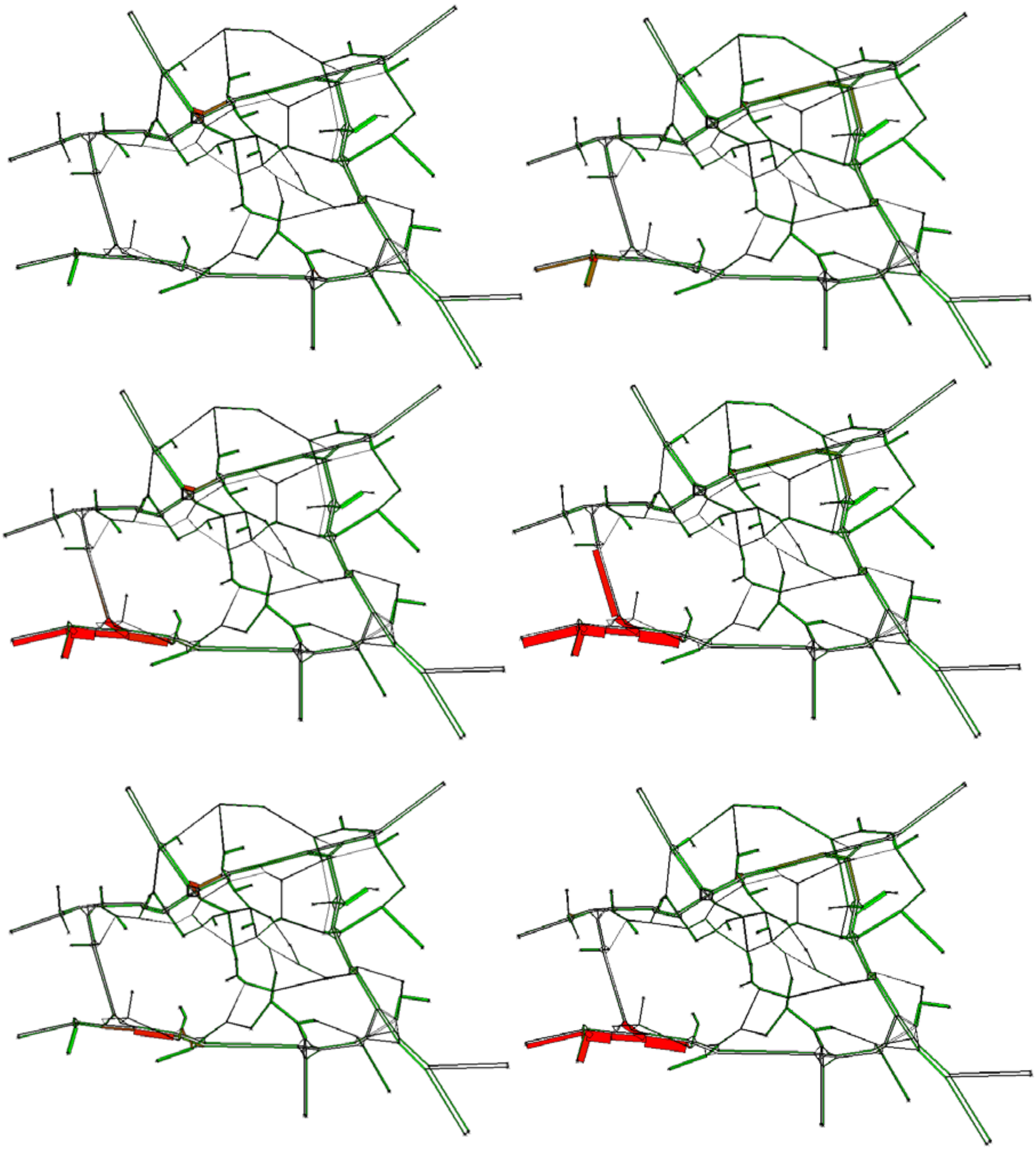


Figure 5.6: Incident on the A15 ─ Situation on the network at 8.15h in the equilibrium assignment without an incident (top), with an incident with fixed routes (middle), and with an incident with a new equilibrium (bottom) for Scenario 1 (left) and Scenario 6 (right)

The case study for the network of Rotterdam suggests that it is plausible that alternative routes are important for the robustness of a road network. The large difference between the situation with fixed route choice and with a new equilibrium after the occurrence of an incident shows that alternative routes can keep the delays that are unavoidably caused by incidents within acceptable bounds. If people are not informed about the incident, or they simply just do not have the opportunity to deviate from their routes, delays of more than one hour can easily occur. This illustrates the importance of timely and reliable en-route traffic information.

Furthermore, six scenarios were examined that varied in the availability of alternative routes for through traffic. In the first scenario, the lower level network can be freely used by through traffic; in the most limited scenario, through traffic can use only one path over the motorway network. This analysis shows the importance of the availability of different paths. From the analysis, it can be concluded that the availability of alternative routes already has a small impact on the travel times in the situation without an incident. This can logically be explained by the equilibrium principle ─if congestion occurs, just enough people take another route in such a way that all used routes have the same travel time. If there is no option for taking another route, everybody will line up in the queue, which results in increased travel times. Furthermore, the results indicate that it is important to have at least one alternative path. If an alternative path is not available, the total delays will be within the region of 9.3%-13.7% of the total travel time in the case of an accident on the off-ramp of the A15. One additional extra path can already reduce the delay to 0.85% - 9.5% of the total travel times. An incident on the A20 results in delays of 4.3% - 102.9% of the total travel time in the situation where no route alternatives are available. One alternative route could reduce these delays to 3.1%-77.1% of the total travel time in the situation where no route alternatives are available. Additional extra paths produce only small improvements. Of course this also depends on the quality (spare capacity and travel time) of the alternative routes; but, in our case, the first alternative route that is added also has the highest quality.

5.4 The importance of unbundling traffic flows

Unbundling traffic flows is defined as the physical separation of traffic flows on a road. Unbundling can, for instance, be used to separate long distance traffic from short distance traffic, or to separate trucks or buses from cars. Figure 5.7 shows an example of an unbundled road ─ the A12 south of Utrecht in the Netherlands, where short distance traffic using the A12 for just a few exits or for traffic from and to Utrecht is separated from traffic using the A12 to bypass the city. In this figure, traffic on the two outermost lanes have the possibility to use the exit to Nieuwegein and Kanaleneiland. The two inner lanes have no exit, so traffic is required to continue driving.

Figure 5.7: Unbundled part of the A12 in the south part of Utrecht in the Netherlands

The concept of unbundling is not new. One of the early reports on unbundling in the Netherlands dates from 1994 (Dutch Ministry of Transport, Public Works and Water Management, 1994). The reasons to unbundle are:

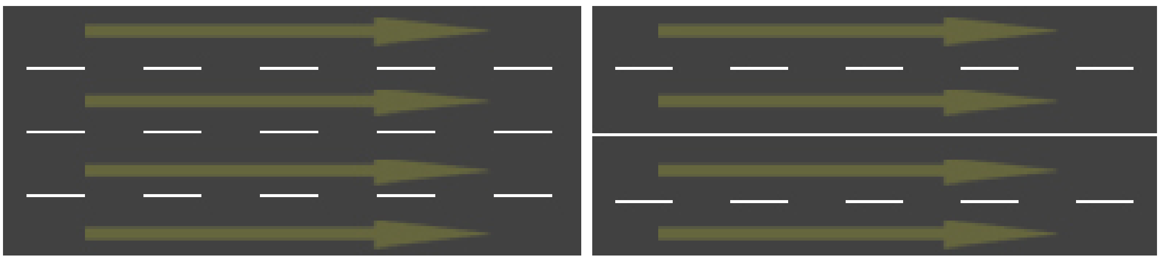
- Offering different quality conditions for different purposes.
- Creating more fluid traffic flows.

- In case of pay lanes, offering an alternative without queues to drivers prepared to pay.

The level of success of the measure depends on the total traffic volume, the flow-capacity ratio, and the share of through traffic (or the share of trucks). Improving robustness is a new argument for unbundling, which has already been adopted (for other reasons) in the policy document *MobiliteitsAanpak* (Dutch Ministry of Transport, Public Works and Water Management, 2008).

Although the precise effects of unbundling on robustness have not yet been determined, it is reasonable to assume that a network can become more robust by unbundling traffic flows, because:

- The chance of a disturbance is usually assumed to depend linearly on the number of vehicle kilometres driven. When a road is unbundled by spreading the available number of lanes equally over two carriageways, the number of expected disturbances per carriageway would then be half of the number of expected disturbances on the same road that is not unbundled (assuming the traffic is equally spread over both carriageways). This applies only if the queue has not passed the connection point of the alternative carriageway. Figure 5.8 illustrates this concept. In Figure 5.8a, the chance of a disturbance on the whole carriageway is taken as p, the number of lanes on which the effect can be measured is q, giving a total effect of p x q. In Figure 5.8b each carriageway is half the number of lanes of the carriageway in Figure 5a, causing the chance of a disturbance per carriageway to be half that of the whole carriageway. If an incident occurs on a carriageway in Figure 5.8b, the number of lanes on which the effect can be measured is also half of those in Figure 5.8a, provided that the queue has not passed the connection point of the alternative carriageway. So, the effect is ½q per carriageway, giving a total effect per carriageway of ½ p x ½q = ¼pxq. The total effect on the whole road in Figure 5.8b will therefore be 2 x ¼ pxq compared to the situation in Figure 5.8a.
- An unbundled road is adjusted to the needs of the user, clustering users that have the same purpose or belong to the same distance class. Therefore, there are no disturbances caused by other users having a different purpose, giving the network a higher level of comfort and a smaller chance of a disturbance. Long distance traffic is, for instance, not disturbed by short distance vehicles changing lanes to get to the off-ramps or to enter the motorway.
- In case of big disturbances on the carriageway for long distance traffic, the parallel carriageway can be used as a backup, as long as there is some spare capacity available on that carriageway. This results in higher robustnesss and, thus, in better traffic flow. The other way around has a smaller effect, since that traffic needs the on- or off-ramp and, therefore, cannot use the carriage way for long distance traffic.



a: Total risk incident p x q b: Total risk incident 2 x ½ p x ½ q= ½ p x q

Figure 5.8: Total risk in the original situation (a) and the unbundled situation (b)

Besides advantages, there are also some disadvantages from unbundling:

- The distribution over the two carriageways is not the same every day of the week, which causes a greater day to day variation than in a situation without unbundling.

- The usage of the capacity on a road without unbundling is equal or higher than the same road where traffic flows have been unbundled. This is because traffic assignment to carriageways is done by a discrete number of lanes, leading to some excess capacity on one carriageway, while traffic on the other carriageway is offered too little capacity. It is, therefore, important to be careful when allocating the number of lanes to the carriageways.
- At the beginning and end of the unbundled sections, extra weaving areas have to be created. This will most likely increase the chance of an incident in these areas. The increase of these chances can be kept relatively small by providing clear information before the unbundled section, and by having long enough weaving areas before and after the unbundled section.
- The spare capacity may be lower due to unbundling, which reduces the potential to absorb the effects of small incidents.
- Building a carriageway that is physically separated from another carriageway requires more space than if no separation is used.

Given this list of potential advantages and disadvantages of unbundling on the robustness of a network and the reliability of travel times, it is clear that an analysis has to be carried out for each specific unbundling project in order to determine whether or not the advantages outweigh the disadvantages. We applied three methods in order to get a better idea in which situations unbundling has a positive effect on the robustness. The first two methods were applied to two fictive example situations of unbundling for the economic evaluation of the *MobiliteitsAanpak*. A detailed description can be found in (Snelder et al., 2009c; 2010b):

- Example 1: separating through traffic from local traffic on a large scale. All motorways between The Hague, Amsterdam, and Utrecht were unbundled in such a way that there are two lanes for the through traffic. The remaining lanes (at least two) are for the local and regional traffic. This implies that no additional lanes were needed, since the capacity of these motorways is already extended to at least four lanes by the 2x4 lanes by other measures proposed in the *MobiliteitsAanpak*. The maximum speed is assumed to be the same on the main road and the parallel road.
- Example 2: separating through traffic from local traffic on a road stretch of about 9 kilometres on the motorway A4 between the splitting with the A44 and the A5.

The model SMARA (Appendix C and Appendix E) was used to evaluate the reliability benefits of these measures. Furthermore, a sensitivity analysis using the vulnerability indicator (see section 3.5) to get a better idea of the effects of unbundling on large disturbances. The analysis with SMARA showed that, for both examples, in case of complete information, the reliability does not change. This is because the network has sufficient spare capacity that can be used as a fallback option in case everybody is completely informed about the disturbance and the availability of the alternative routes. In case of fixed route choice, the reliability improves in Example 1 by 1%, whereas the reliability decreases in Example 2 by 1%. The fact that the impact of splitting through traffic from local traffic differs per project indicates that no general rule of thumb can be found for the effects of splitting through traffic from local traffic. The vulnerability indicator indicates that the network gets more robust. The vulnerability indicator focuses on large disturbances in which roads are, for instance, completely closed. In case the road is split into two separate roads, one road will maintain its function in case the other one is closed. The capacity of the road that is still open might be lower due to rubbernecking effects. Nevertheless, the fact that one road maintains its functions ensures that the effect of the incident is lower than the case in which all lanes would be closed if the road was not split into two roads. Besides that, a part of the traffic that would choose the closed road if it had not been closed can use the road that is still open, if there is spare capacity. However, this is possible only for the through traffic, since the local traffic cannot use the main road, because then they would miss their off ramp. Creating flexible infrastructures can improve the transferability of traffic and, therewith, the robustness.

The third method is based on the evaluation method that is described in the previous chapter. Ketelaars (2010) applied the combination of the dynamic traffic assignment model Indy with the MIC-module and the alternative route choice approximation method in her study, in which she developed an optimization method to determine where which type of unbundling can best be applied. This method is partly described in section 7.5.4. In this section we summarize the most important results with respect to the effects of unbundling on robustness for an example application to the network of Delft in the Netherlands. In general, it was concluded that the most effective type of unbundling depends on the ratio of traffic with specific purposes, which is unique to every network. This also counts for the most effective locations and forms of unbundling.

The vulnerability per link in the Delft network in morning rush hour is shown in Figure 5.9. The vulnerability is expressed by the number of hours lost due to possible incidents per link, reducing the capacity on all links individually by 5% from 7.45 - 8.00 AM, by 20% from 7.45 - 8.30 AM, by 70% from 7.45 - 8.30 AM, and by 15% from 7.45 - 8.30 AM. In this case, the economic cost of the vulnerability of the network for the morning rush hour is €7450.

The optimization process in which the cost of unbundling is weighted against the benefits (only the robustness/reliability benefits are considered; travel time benefits are, for instance, not included), resulted in a situation in which part of the A13 (the motorway that runs on the east side of Delft) is unbundled. This is shown in Figure 5.10**.** The resulting vehicle loss hours are shown as well.

This unbundled network now has an economic cost of €1800 due to the vulnerability of the network. This is a cost reduction of 60% compared to the original situation, and 8% lower than the cost of the vulnerability of the network if one lane would be added to the A13 without unbundling (the original road had three lanes, and one lane is added for unbundling). Considering the cost of modifying the network, which is taken as a 30-year investment of €8 million per lane per kilometre, the economic gain over the 30 years would be 2 million euros by this method of vulnerability cost calculation.

From the above analysis, it can be concluded that unbundling can have a positive effect on robustness. However, the way in which it is done and the locations where it is done need to be carefully chosen.

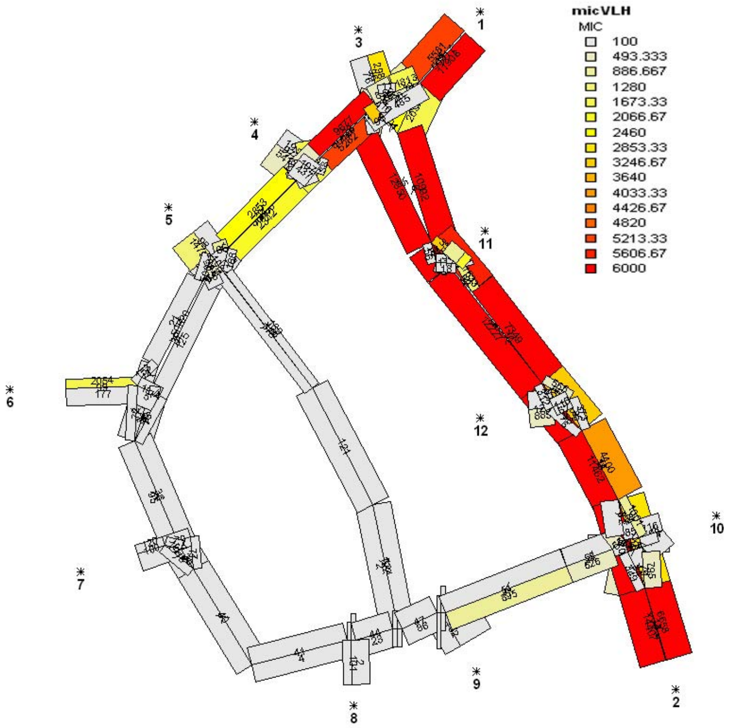


Figure 5.9: Vehicle loss hours in the network of Delft before unbundling

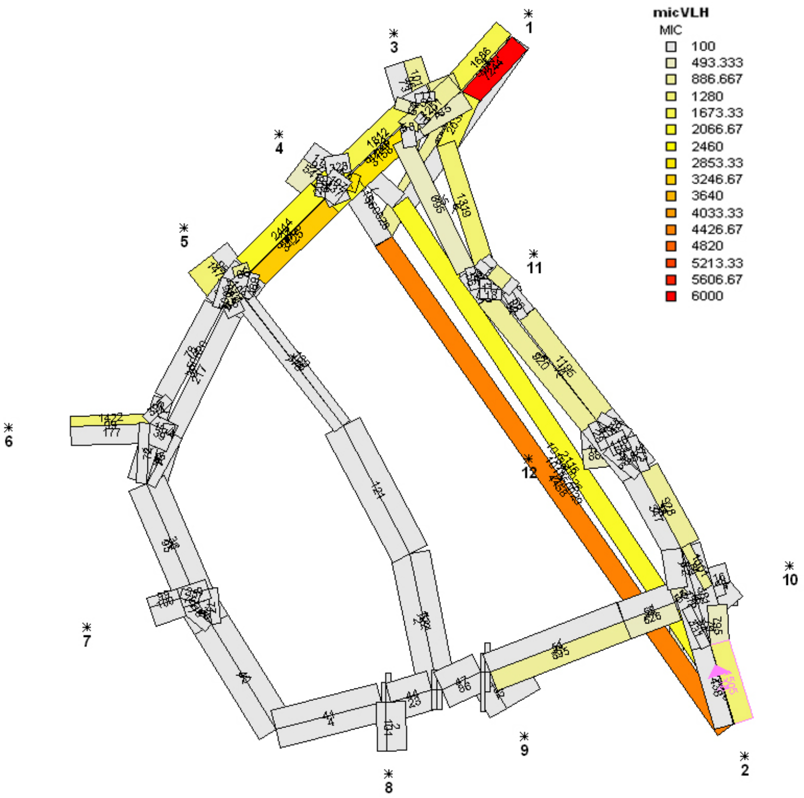


Figure 5.10: Vehicle loss hours in the network of Delft after unbundling

5.5 The importance of buffers and spare capacity

Section 2.2 showed that spillback effects are the main reason why the effects of disturbances can get very large. Due to the spillback effects, local disturbances can lead to queues on a large part of the motorway network. One measure that can be taken is to include buffers. Figure 5.11 shows a buffer strategy that aims at preventing congestion from occurring on motorways and urban roads. In this strategy, three buffer locations are considered:

1. Prevent congestion on the long distance network by creating traffic regulation buffers at on ramps and spillback buffers at exit ramps in the long distance network;
2. Prevent congestion on the urban road network by creating traffic regulation buffers at the urban area 'access points' (and spillback buffers at the 'exits', if needed); this includes keeping urban ring roads as congestion-free as possible;
3. Add spillback buffers upstream from bottlenecks *within* the metropolitan network, if needed.

These buffer locations are in addition to the buffer capacity of the roads themselves.

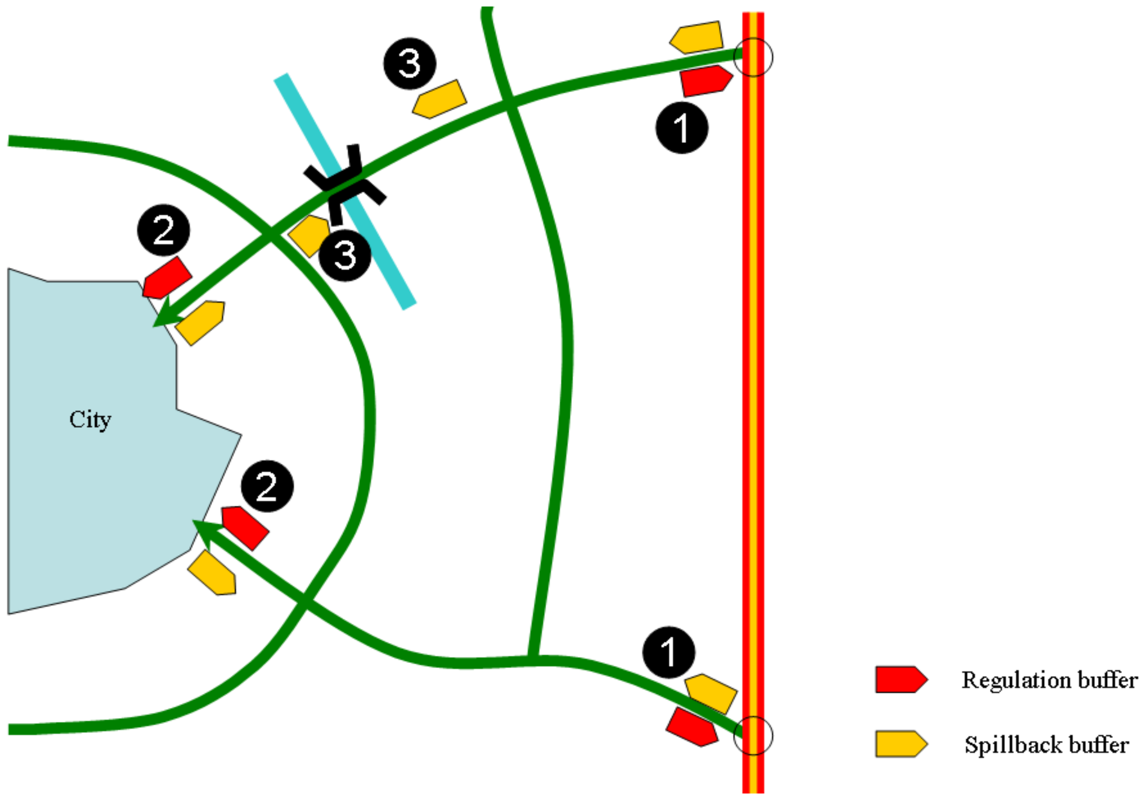


Figure 5.11: Example of buffers in the metropolitan network

Below, based on three examples, an indication is given of the potential effects on the robustness of a network of adding buffers and spare capacity.

Example 1: Spillback effect on a long stretch of road

In this example[29], a fixed bottleneck location is chosen at an intersection of two motorways. A similar approach could be applied to temporal bottlenecks such as accidents, which would result in slightly different results. In this example, the following assumptions are made:

[29] A similar example is given in the course Traffic Flow Theory and Simulation vk4821, Dr. Ir. Serge P. Hoogendoorn, Transportation and Traffic Engineering Section Faculty of Civil Engineering and Geosciences Delft University of Technology.

- The effects are measured on one of the motorways upstream of the bottleneck. It is assumed that this is a long road stretch with no on- and off ramps near the intersection.
- The capacity of the motorway is 6600 vehicles per hour and has 3 lanes.
- Only 4900 vehicles per hour can pass the bottleneck.
- There is a buffer lane of 4 kilometres with a capacity of 2200 pcu/hour.
- The maximum speed is 80 km/hour.
- The demand is 6400 pcu/hour in the peak period and 4000 pcu/hour in the off-peak period.
- The peak period lasts 60 minutes.
- The jam density is 150 pcu/lane km.

This situation is shown in Figure 5.12. Because the demand is higher than the capacity of the bottleneck, a queue builds up in the peak period. The queue only begins to reduce when the demand gets lower. In the situation without a buffer lane, the length of the queue will grow to 15.4 kilometres (based on kinematic shock wave theory). In practice, there would be on ramps and off ramps, and maybe even other motorway intersections, along these 15.4 kilometres that would get blocked by the queue as well. If there were a buffer lane of 4 kilometre, the queue length would be only 9.3 kilometre, which might be just enough not to block other motorways.

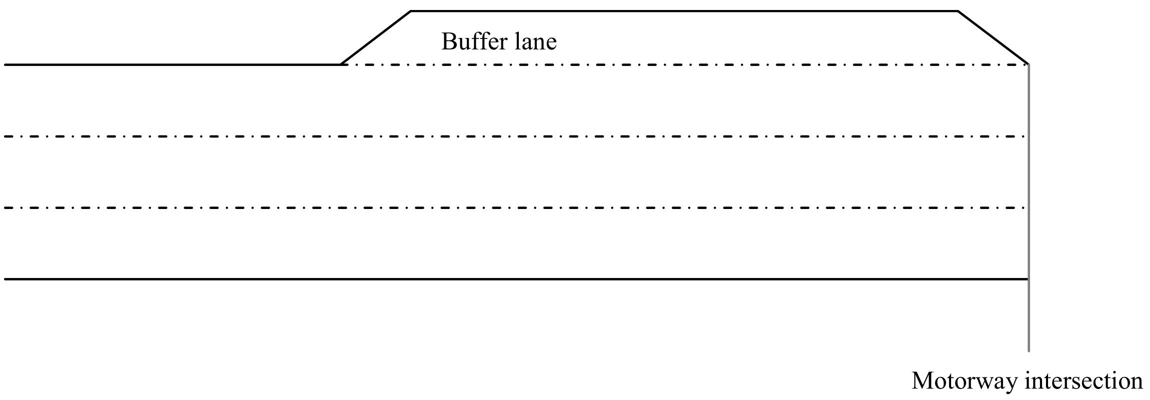


Figure 5.12: Example of a buffer lane

Example 2: Potential effects of a buffer in the network of The Hague

The above buffer strategy, with the three types of buffers, can be applied on a larger scale to a complete network. The potential of buffers was tested by doing two model runs with Indy using the road network of The Hague and its surroundings:

- A model run for the year 2000 without buffers: this situation is simulated by using the link transmission model of Indy in which spillback effects are included.
- A model run for the year 2000 with buffers: this situation is approximated by using the vertical queuing variant of Indy in which spillback effects do not occur (see section 4.3.1). This is a kind of 'maximum buffer strategy' in which buffers are built at all locations where queues build up.

The difference in total travel time between both runs indicates the maximum potential of buffers under regular circumstances. In practice, the benefits will be lower, because it is impossible to create buffers at all locations where queues build up and block other roads.

In Figure 5.13, the simulated network is shown. This network is a part of the NRM-network (regional model system of the Netherlands) for the year 2000. In total, this network has 3308 links, 2412 nodes, and 211 zones. In Table 5.4, the demand for the different time periods is shown. A simple calibration procedure was carried out in such a way that the simulated locations of congestion match the locations where congestion occurs in practice.

Figure 5.13: Network The Hague and surroundings 2000

Table 5.4: Number of trips in 2000

Time	Number of Trips		Time	Number of Trips
6.00-7.00 AM	73146		9.00-10.00 AM	86352
7.00-8.00 AM	131384		10.00-11.00 AM	77209
8.00-9.00 AM	129295		11.00-12.00 Am	73145
Total trips: 570531				

In Table 5.5, the total travel time, travelled distance, and average speed are shown for both model runs. This table shows that buffers can reduce the total travel time by 12.1%, which is an average of about 2 minutes per trip. This reduction is partly explained by a higher average speed in the case with buffers (+2.3 km/hour), and partly by shorter distances travelled. The total distance travelled is 7.6% kilometres lower in the case with buffers. These distance benefits are explained by the fact that people choose shorter alternatives to avoid congestion. The queue length at 8.30 AM, for instance, is 20 km (46%) lower in the case with buffers. In Figure 5.14, the development of queues over time is shown. From this figure it can be concluded that buffers can reduce the queue length during the complete morning peak period. Figure 5.15 shows at which locations the buffers can reduce the queues.

Table 5.5: Network outcomes with and without buffers

Case	Total travel time [hours]	Total distance [km]	Average speed [km/hour]
Without buffers	145929	6469221	44.3
With buffers	128266 (-12.1%)	5976838 (-7.6%)	46.6

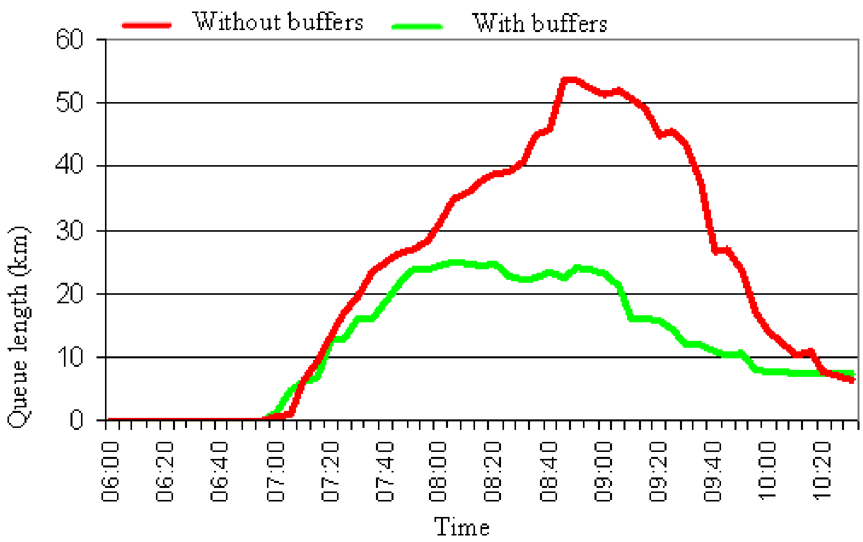


Figure 5.14: Queue length in the case with and without buffers

Figure 5.15: Difference in density in the case with and without buffers

The above analysis showed that buffers have the potential to reduce the average travel time by 12% in the morning peak in the network of The Hague. Of course, the specific impacts will be different in other networks and in other time periods. In more congested networks, the benefits of buffers are likely to be higher. This example is based on a regular situation without incidents. However, it is clear that similar effects will occur in case of incidents. This implies that buffers can make a network more robust.

Example 3: Optimization of spare capacity.
In Example 1, it was shown that adding a buffer lane can reduce the queue lengths, and therewith the travel time delays, significantly. In Immers et al. (2004d), the spare capacity of a road is optimized. These results are briefly summarized below.

A single link between an origin r and destination s is considered. On this network link, the section with the smallest capacity is called a bottleneck. Let *cap* (pcu/hour) be the capacity of this bottleneck. Between r and s we assume a fixed traffic demand D (pcu/hour). Let c_{infr} be the extra cost of the infrastructure with spare capacity, expressed in time units per unit of spare capacity for the period under consideration, and per individual.

It was shown that, if $D \le cap\text{-}dc$ (dc is the capacity reduction), then no spare capacity is needed. If demand exceeds capacity ($D > cap$), then a queue builds up, which never dissipates. These are not interesting cases. Therefore, we consider only the case where $D \le cap$ and $D > cap\text{-}dc$. Immers et al. explained that the optimal spare capacity, rc_{opt} (pcu/hour), could be computed by:

$$rc_{opt} = \frac{\sqrt{p_{dc}}\Delta T_v dc}{\sqrt{c_{\inf r}}\sqrt{2D\Delta T}} + (D - cap) \tag{5.1}$$

The formula highlights the influence of the risk component p_{dc} (probability of an incident) and ΔT_v (disruption period = consequences of the incident). Increasing traffic safety naturally reduces the necessary spare capacity. But, even more important for minimising the length of the disruption period, as shown by the exponents of the different terms, is the resilience of the system. Providing for an adequate incident management service enhances the resilience of the transportation system. As an example, for a motorway link with three lanes, Immers et al. found that the optimal spare capacity is about 800 pcu/hour.

In this example, spillback effects were not considered. Including spillback effects complicates the optimization problem. In the following chapters, the problem of optimizing spare capacity in a dynamic environment with spillback effects is dealt with.

5.6 Summary: measures for improving the robustness of a road network

In this chapter the 8th research question was addressed: Which measures can be taken to improve the robustness of a road network, and what are their effects?

A long list of possible measures for improving robustness were presented. These measures were subdivided into measures that can be taken on the travel market, the transport market, and the traffic market. Furthermore, it was explained to which elements of robustness they related and for which disturbances they are relevant.

For a selection of measures that match best with the focus of this thesis, the possible effects of the measures were worked out in more detail by using examples. The following measures were selected:
- Creating alternative routes
- Unbundling traffic flows
- Creating buffers and spare capacity

It was shown that having one route alternative could reduce the delays by 3.1%-77.1%, depending on the level of information that is given. Additional extra paths produce only small improvements. Of course, the results also depend on the quality (spare capacity and travel time) of the alternative routes.

With respect to unbundling, it was concluded that unbundling can have a positive effect on robustness. However, the way in which the unbundling is done and the locations where it is done need to be chosen carefully, because unbundling can have both advantages and disadvantages for the robustness of a network.

Finally, it was shown that adding buffers and spare capacity to a road network can have large positive effects on the travel times. Buffers have the potential to reduce the average travel time by 12% in the morning peak in the network of The Hague. This impact will, of course, be different in other networks and in other time periods. The more congestion there is, the higher the benefits may become. These conclusions are based on a regular situation without incidents. However, it is clear that similar effects would occur in cases with incidents. This implies that buffers can make a network more robust. Finally, an example was taken from Immers et al. (2004d) that showed that the optimal spare capacity of an example motorway link with three lanes is about 800 pcu/hour. This optimization was done without considering spillback effects. In the following chapters, an optimization method is presented that considers spillback effects in the optimization of spare capacity. In addition, other robustness measures are included in the design method.

6 Formulation of the robust network design problem

6.1 Introduction

In the previous chapters, the concept of robust road networks has been described. Different measures were presented that could improve the robustness of a road network. However, the question where (and when) in the network which measures should be taken has not been answered yet. In the introduction, we explained that designing/improving a network is a difficult task in practice, because many stakeholders are involved whose objectives sometimes contradict each other. Furthermore, there are many interdependencies between the choices of travellers and investment decisions. In the investment decisions, long term and short term problems need to be considered, and all kinds of long term and short term uncertainties need to be taken into account. There are many laws and procedures related to infrastructure investments in order to make sure that good decisions are made. However, these procedures often take a long time. Finally, the political environment plays an important role. In section 6.2, we elaborate on some practical planning dilemmas from a network structure point of view, and we describe how these dilemmas are dealt with in practice. From a theoretical and mathematical point of view, network design is very complicated as well. This is discussed in detail in section 6.3. In the remaining sections of the chapter, we formulate the robust road network design problem by giving a mathematical formulation of the objectives and restrictions that we consider.

6.2 Transport planning in practice

This section describes the planning dilemmas with respect to the structure of a network and the way these dilemmas are dealt with in practice in the Netherlands.

Structural dilemmas
Immers et al. (to be published) present a number of structure related dilemmas related to the design of robust networks:

- *Dilemma 1, the number of systems*: the more subsystems (dedicated networks), like primary and secondary road networks, the better their functions can be geared towards the needs of the traveller. So, offering more subsystems increases the user benefits. On the other hand, reducing the number of subsystems means reducing the investor costs, as this means the capacity offered can be used more efficiently. A practical example of this dilemma is the question of whether short and long distance travel should be combined on the same ring road; this means a high quality road for short distance travel, but disturbance of the long-distance traffic flow caused by the short distance between access points. In general, more subsystems can be offered in more urbanized areas, where the transport demand is higher.
- *Dilemma 2, access point density*: the more access points, the better a road's accessibility. This means that a smaller part of the trip needs to be made on a lower level (and therefore slower) network. On the other hand, the quality of connections (how fast, and how reliable from one access point to another) provided by the subsystem is higher when there are few access points.
- *Dilemma 3, access structure*: apart from defining the ideal structure of the connections between towns, there is the question of where to locate the access points. For instance, a choice can be made for one access point in the middle, or one or more access points at the edges of the built-up area. The first option maximizes the accessibility of the system, but this often leads to 'misuse' of the system by traffic that could use a lower level network. Also, it may affect liveability in the area.
- *Dilemma 4, network density*: once it has been established which cities need to be connected, it still has to be decided whether these cities should be connected by direct links or by way of another city. More links means higher quality connections, because there are fewer detours. But, more links mean higher costs, not only in infrastructure investments, but also in the effects on the environment. What network density will be acceptable depends chiefly on two factors:
    - the amount of traffic: high volumes justify the need for extra infrastructure;
    - the difference in quality between the two subsystems: a greater difference (in design speed) between network levels means that a greater detour is acceptable when using the higher level system.

A general choice in these dilemmas cannot be made, because the choices depend on all kinds of location-specific circumstances. Below, some tools are described that can be used to deal with these dilemmas.

Design tools
Usually, a bottleneck driven approach is followed in deciding where infrastructure adjustments should be made. Locations with the biggest bottlenecks in terms of congestion are considered first. The Netherlands Institute for Spatial Research has published a report (PBL, 2006) in which a more advanced method is described for finding the roads that would lead to the highest economic growth. This method focuses on one objective: economic growth. The method results in an assessment of the existing roads, including the plans for extending the network up to 2020. Besides this approach, there are several other methodologies for deciding where new roads should be constructed. In the Netherlands, Infralab (De Rooij, 2000), IRVS (Wilmink et al., 2002), ARNO and ARKO (Egeter et al., 2007) are examples of these. These methods consider a complete network instead of focusing on bottlenecks, and they make use of expert knowledge. Infralab is a method that was developed by the

Rijkswaterstaat, the implementing body of the Ministry of Transport, Public Works and Water Management. Infralab consists of three sessions (problem formulation, solution, and design) with road users and local residents that aim at finding a solution for a specific traffic problem.

In what follows, the IRVS method is described in more detail (based on Immers et al., to be published). This method is the basis for ARNO and ARKO as well. The difference between ARNO, ARKO, and IRVS is that IRVS focuses on different networks and the interactions among them. ARNO goes a level deeper ─ up to the relation level and details of the road design (number of lanes, road profile, etc.). The ARKO method focuses on intersections as well. These methods are described in this section, because they form the basis for the method that is described in this thesis for designing robust road networks.

The IRVS methodology was developed for the integral design of transport networks of different modes, with a focus on the regional scale. Separate from the present infrastructure, an 'ideal network' is designed as a basis for the analysis. Furthermore, a design is developed together with the stakeholders on the basis of clear, practicable steps. By creating an 'ideal network' separate from the network that is present, a very clear insight is gained into the structure of the network, since it is not obscured by the existing situation, which has emerged historically, and therefore is not always ideal. By confronting this ideal situation with the existing situation, weaknesses in the structure come to light. A second function of the ideal network is that of a long-term horizon within which short-term measures have to fit. By reducing the theoretically highly complex design problem to a number of successive design steps or decisions, this methodology provides insights and can be applied in practical situations. What is important in this respect is that for each step there is commitment from the stakeholders before the next step is taken. The methodology is most effective when it is used in a workshop situation, where the parties themselves participate in the design process. The result of the methodology is that stakeholders gain a clear picture of the crucial dilemmas and decisions. The methodology avoids thinking in terms of end solutions. Instead, the functions of the different parts of the network can be analysed in terms of whether they actually fulfill the functions for which they were designed or for which they are now ascribed. The function of a particular part of the network is, thereby, the leading factor for form and technique. Analysis may result in a whole palette of possible recommendations for each situation, varying from no action, through traffic management function adjustment coupled with modification of the road design and disentangling or expanding existing connections, to the construction of new junctions or new connections. This can be phased in ─ for instance, by first applying traffic management and then, in the longer term, building new junctions or connections.

6.3 The network design problem (NDP) from a mathematical/theoretical point of view

The Network Design Problem (NDP) is a well-known problem in the literature. It has been recognized as one of the most difficult and challenging problems in transport (Yang and Bell, 1998). The problem involves optimal decisions on the expansion of a road network. Usually there is a certain budget restriction. The previous section has explained that, in practice, design methods are used that rely for a large part on experts to deal with the complexity of the NDP. This section illustrates the complexity of the Network Design Problem from a theoretical and mathematical perspective. The objectives, design variables, demand matrices, problem formulations, solution algorithms, and paradoxes that are used in the literature are described.

Example:
If a network is to be designed with 4 nodes, at least 3 links are needed to connect all nodes (the builder optimum), and at most 6 links are needed (user optimum). In total, 20 different networks can be designed with 3 links, of which 4 are not feasible because they do not connect all nodes. With 4 and 5 links, 15 and 6 networks can be designed, respectively. With 6 links, all nodes in the network are completely connected. This shows that, in total, there are 38 different networks that connect all nodes (and 42 networks in total) ─ see Figure 6.1.

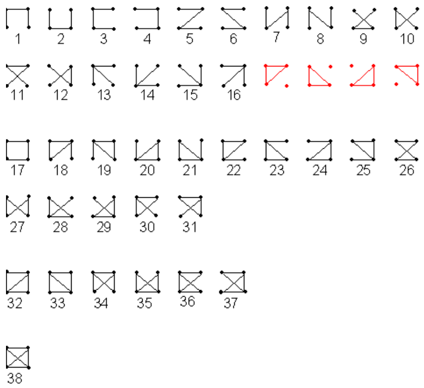
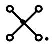


Figure 6.1: Networks with four nodes

Table 6.1: Number of networks that can be created given the number of nodes

Nodes (n)	Number of networks[30]
3	4
4	42
5	848
6	30.827
7	2.069.256

Table 6.1 shows that the number of different networks that can be designed grows very fast with the number of nodes. If intersections are permitted, four nodes can be connected with two links and one intersection as well: . So, permitting intersections results in even more possibilities.

The example above illustrates the complexity of the NDP. In this example, the only decision that is to be made is whether or not a link is constructed. In practice other decisions about, for instance, the number of lanes, the road type, and the maximum speed, have to be made as well. Of course, this increases the complexity of the problem.

Decision variables
Designing road networks raises many questions with respect to the characteristics of the network. These include where new links are to be constructed, the number of lanes and/or the capacity of links, as well as the speed limit, the safety measures, the signal settings, the pricing strategy, and many other design variables.

The approach to road network design that is most often followed in the literature is to start with an existing network and to define links that could be modified. The capacity or the number of links of these modifiable lanes is determined. A choice can be made for zero lanes, which implies that the link is not constructed. The number of lanes is a discrete variable; the

[30] Total number of networks $= \sum_{i=(n-1)}^{\frac{1}{2}n(n-1)} \binom{(1/2)*n*(n-1)}{i}$

capacity is a continuous variable. A mixture between both is possible. This is for example the case if one decision variable is whether or not a new road is or is not constructed, and one determines the capacity of the roads. The choice between these types of variables is important, because they require different solution algorithms. Since solving the continuous version of the NDP is easier than solving the discrete version, the NDP is most often formulated as a continuous or mixed continuous NDP. The discrete approach is, for example, followed by Le Blanc (1975), Poorzahedy and Turnquist (1982), and Xiong and Schneider (1992). The continuous formulation is, for example, used by Suwansirikul et al. (1987) and Friesz et al. (1992). Below the different solution algorithms are discussed.

The design level (e.g. motorways, secondary roads, local roads) is another important factor in specifying the decision variables. For motorways, different decisions need to be made than for urban roads. For urban roads it is possible, for example, to design the network topology or the lane layout at the same time as the signal settings at crossings, as is done by Cantarella et al. (2006). Signal settings for motorways are less relevant.

It should be noted that the decision variables usually relate to links and not to nodes. Also, in the literature, decision variables that are especially related to robustness have not yet been described, as far as is known to the author.

Objective function

In previous sections, we have explained that there are many stakeholders involved in designing road networks. For example, there is the government as network planner and as network manager, there are travellers, there are companies, and there are many others. These stakeholders all have different objectives, which ideally should all be taken into account. However, this is quite difficult. In studies performed so far, often the objective function consists of a single objective; in only few cases are multiple objectives considered. According to Santos (2005), solutions should be efficient, making aggregate accessibility as large as possible or aggregate travel time as small as possible, and solutions should be equitable, robust, and sustainable, and comply with the budget available to improve the network. Except for robustness, Santos expresses all of these aspects in the objective function or through constraints. This is one of the few examples in which multi-objective analysis has been carried out in relation with the Network Design Problem.

Most often, the objective is to minimize the total travel time or the total travel cost subject to a budget restriction on infrastructure costs. Yang and Bell (1998) also mention maximization of the consumer surplus and maximization of spare capacity as possible objectives. Spare capacity is an indicator for capacity reliability. Lo and Tung (2003) maximized spare capacity by maximizing the O-D multiplier. Gao et al. (2005) did the same, but allowed different multipliers for individual O-D pairs. Only in a few studies is an indicator for reliability included in the objective function. Chootinan et al. (2005) present a reliability based network design problem. They introduce an index for capacity-reliability, which measures the probability that all of the traffic links are operating below their respective capacities. In this case, the objective of the network designer (upper level problem) is to determine the optimal capacity enhancements by maximizing the reliability of the network (or minimizing the probability of overloading the network links) under a budget, while taking into account the behaviour of network users. Sumalee et al. (2006) describe the Reliable Network Design Problem. Their objective is to maximize the network total travel time reliability. Li (2009) also includes the reliability of travel times in the objective function. Ukkusuri et al. (2007) also introduce robustness into the objective function. Their aim is to make the network robust against long term uncertainties in demand. According to them, a clear distinction between robust optimization (RO) and stochastic programming (SP) methods is needed in order to differentiate their applicability. SP models account for uncertainty by the minimization of the expected value objective function; RO considers higher moments of the probability distribution in addition to the expected value of the objective function.

Demand

There is a strong interaction between the demand pattern and the quality of the infrastructure. One the one hand, the infrastructure design should match the demand pattern. The demand for travel requires a certain level of infrastructure. On the other hand, the infrastructure and the quality of the infrastructures is also an important determinant in the location choice of residents and companies. Furthermore, the mode choice and departure time choice of travellers depends on the quality of the infrastructure and the services offered. For example, if there is no congestion on the road network, more travellers will prefer the car over public transport than in a situation where congestion exists, and, if a traveller is aware that the road network is more congested than normal, he or she might depart earlier. This supply-demand interaction should be taken into account when designing road networks.

A second demand issue is the extent to which future demand should be taken into account. Infrastructure has a long economic lifespan. Furthermore, due to many factors, there is often a long time between the designing of a road and the completion of its construction. Both the economic lifespan and the long construction time require that future demand be taken into account, as is done by Ukkusuri et al. (2007). It is possible to make a network robust against long term uncertainties in demand.

The interaction between future supply and future demand is not often considered in the literature. Most often, a fixed origin-destination matrix for one time period is assumed. This simplified approach is probably used because including the demand-supply interaction is difficult ─ it requires an OD-estimation method to be run every time that the network is adjusted. This significantly increases the computation time, and convergence is not guaranteed. Nonetheless, several attempts have been made to include the supply-demand interaction (elastic demand) in the NDP. Already in 1980, Boyce and Janson (1980) formulated the combined distribution and user-equilibrium assignment NDP and the combined distribution and system-optimal assignment NDP in a 10 node network. Asakura (1999) mentions that cancelled or suspended travel demand must be considered, but does not propose a way of doing this. To the best of our knowledge, there are no examples of elastic demand in the design of large networks.

The reason for not taking into account the time from design until construction, and not mentioning the economic lifespan of infrastructure, is probably that whenever a fixed OD-matrix is assumed, the year from which this demand is derived is not relevant for explaining the functioning of the solution algorithm of the NDP. There are some exceptions in which the time aspect is considered. For example, Groothedde (2005) presented a method for hub network design in which the network development path is explicitly taken into account.

Since reliability has been recognized as an important aspect, it could be claimed that the reliability of travel times should also be taken into account in trip distribution. Modelling reliability requires a dynamic assignment with dynamic OD-matrices. Accidents could, for example, cause a change in departure time of some travellers, or perhaps even a change in mode choice or destination choice. Few examples can been found in the literature about these kinds of approaches. One example is Li (2009). She included the reliability of travel times as a factor in route choice and departure time choice, but not in other choices, such as mode and destination.

Problem formulation
In different articles and books, overviews are presented of problem formulations and methods that solve the NDP (Yang and Bell, 1998; Magnanti and Wong, 1984; Steenbrink, 1974). These overviews clearly show that there are many possible methods. Although the methods differ in many ways, they also have a number of similarities. For example, the NDP is usually formulated as a non-convex and non-differentiable bi-level problem. The top level addresses the question where new links should be constructed or where the capacity of existing roads should be extended, given the transport flows. The objective function on this level considers the total costs of road investments and the travel costs. The lower level problem is the assignment problem. In the assignment problem, the traffic is assigned to the network, where the objective function considers the individual travel times. This bi-level problem can be seen as a Stackelberg game[31], in which the network designer (on the top level) is the leader and the travellers (at the lower level) are the followers.

Typically, the NDP can be formulated as the following bi-level program:

$$\textbf{(UP) Minimize } Z_1(\mathbf{f}, \mathbf{l}^{\mathbf{new}})$$
$$s.t. \quad G_1(\mathbf{f}, \mathbf{l}^{\mathbf{new}}) <= 0$$
$$\textbf{where } f = f(\mathbf{l}^{\mathbf{new}}) \textbf{ is implicitely defined by :}$$

$$\textbf{(LP) Minimize } Z_2(\mathbf{f}, \mathbf{l}^{\mathbf{new}})$$
$$s.t. \quad G_2(\mathbf{f}, \mathbf{l}^{\mathbf{new}}) <= 0$$

In this formulation, Z_1 is the objective function of the upper level problem (UP) and l^{new} is the decision vector of the upper level problem. G_1 is the constraint set of UP. In the lower level problem (LP), Z_2, f, and G_2 are respectively the objective function, the decision vector, and the constraint set.

The NDP is not always formulated as a bi-level program. Meng et al. (2001) transferred the bi-level program of the continuous network design problem into a single level, continuously differentiable, but still nonconvex, optimization problem.

Solution algorithm
Solving the robust network design problem for road network is complex. Since designing other networks is complex as well, Tero et al. (2010) tried to learn from the way biological networks are designed. Biological networks have been honed by many cycles of evolutionary selection pressure, and are likely to yield reasonable solutions to such combinatorial optimization problems. Furthermore, they develop without centralized control and may represent a readily scalable solution for growing networks in general. It was shown that the

[31] It is a Stackelberg game if the road authorities want to design the network in such a way that it matches the demand. However, if road authorities design their network in line with other objectives, such as influencing travel behaviour, it is no longer a Stackelberg game.

slime mold Physarum polycephalum forms networks with comparable efficiency, fault tolerance, and cost to those of real-world infrastructure networks (in their case, the Tokyo rail system). The core mechanisms needed for adaptive network formation can be captured in a biologically inspired mathematical model that may be a useful to guide network construction in other domains. These recent findings are very interesting and promising. However, the method is not (yet) applicable to the robust road network design problem, since it requires the inclusion of many constraints and exogenous factors that cannot (yet) be captured by their method.

It has been known for many years that the complexity of the network design problem makes it extremely difficult to find a global optimal solution (the discrete version of the problem is proven to be a NP-complete problem (Johnson et al., 1978)). Therefore, many attempts have been made to find algorithms/heuristics by which the problem can be solved in the best possible way and in an acceptable computation time.

LeBlanc (1975) used a Branch and Bound algorithm, and Poorzahedy and Turnquist (1982) used a Branch and Backtrack algorithm to solve an approximation of the discrete version NDP. Steenbrink (1974) proposed an iterative decomposition algorithm in which the user optimal flows were approximated by system optimal flows. More recently, Gao et al. (2005) used the support function concept in a generalized Benders decomposition with some simplifications with respect to the demand matrix (a fixed demand was assumed). Also, genetic algorithms are often used, since they are capable of dealing with problems that do not possess nice mathematical properties, such as continuity, differentiability, uni-modality, and convexity. Santos et al. (2005) compared three different approaches: an add+interchange algorithm (AIA), a basic genetic algorithm (BGA), and an enhanced genetic algorithm (EGA). The EGA improves the BGA in several aspects. It comprises an intervention procedure, an interchange procedure, and an add procedure. It was found that AIA always outperforms BGA, and that EGA outperforms AIA. However, depending on the network size, EGA is 2 to 8 times slower than AIA. Xiong and Schneider (1992) used a cumulative genetic algorithm to solve a discrete version of the NDP. A large majority of the recent papers about NDP use a genetic algorithm to solve the problem.

The continuous network design problem (CNDP) is solved by Pearman (1979) using a simple approximation method. Friesz et al. (1992) use simulated annealing, and Abdulaal and LeBlanc (1979) use the methods of Hooke and Jeeves and Powell. The method of Fibonacci, Golden section and Bolzano search are applied by Suwansirikul et al. (1987). It is shown that these algorithms are faster than the algorithm of Hooke and Jeeves in case of convex investment functions. Chiou (2005) proposed four variants of gradient-based methods to solve the general CNDP. He compared these variants with some of the above mentioned and other methods. In all the test cases, a gradient-based method produced similar results to those of the Simulated Annealing algorithm (which can be regarded as a globally optimal solution to the CNDP), with much less computational effort.

The lower level problem can be solved using different types of assignments. Choices have to be made between static and dynamic assignments, deterministic and stochastic assignments, and single user and multi-user class assignments. Also, a decision has to be made as to whether or not to use an equilibrium assignment. In previous research on network design, different assignment methods are used for solving the lower level problem of the NDP. For example, Friesz et al. (1992) and LeBlanc (1975) use the Frank-Wolfe algorithm to carry out a deterministic user equilibrium assignment. Xiong and Schneider (1992) use a neural network approach to carry out this same assignment. The stochastic user equilibrium assignment is used by Lo and Tung (2001). Davis (1994) uses this kind of assignment in combination with, respectively, the continuous and discrete versions of the NDP. Chootinan et al. (2005) use a variant of the method of successive averages, using the exponential average to represent the learning process of network users on a daily basis, which results in daily variation of the traffic-flow pattern and Monte Carlo stochastic loading. Sometimes very simple assignment procedures are used. For example, Santos (2005) uses an all-or-nothing assignment. No examples have been found in literature of the NDP in combination with a multi-user class assignment; examples of the network design problem in combination with a dynamic assignment are also rare, but exist. For example, Waller and Ziliaskopoulos (2000) present a CNDP in which a system optimal stochastic dynamic traffic assignment is used. Reliability has been considered in some assignment algorithms. For example, Lo and Tung (2003) introduced a probabilistic user equilibrium assignment.

Since travellers feel that reliability is important, it is likely that they also take this aspect into account in their route choice. Lo and Tung (2003) postulate that drivers would select routes to lower their travel time variability, just as they would select routes to lower their travel time. Therefore, reliability should be included in assignment algorithms, as is, for instance, done by Li (2009) and Lo and Tung (2003). This is called a Probabilistic User Equilibrium (PUE), which can be seen as an extension to the deterministic user equilibrium (DUE). The PUE conditions require that the travel time distributions of any used routes have the same mean, which is equal to the minimum mean origin-destination travel time. Moreover the travel time distributions of any used routes have variabilities that are within specified bounds.

Finally, Meng et al. (2001) transferred the bi-level programming formulation into a single level optimization problem. This resulted in a continuously differentiable, but still nonconvex, optimization formulation of the CNDP that is solved by a locally convergent augmented LaGrange method. This method was also compared with MINOS, Hooke and Jeeves, EDO, and Simulated Annealing, and it proved to be at least not inferior to these algorithms.

Paradoxes

In road network design, some seemingly strange phenomena might occur. This is illustrated by the Braess-paradox and the Pigou-Knight-Downs paradox. These paradoxes demonstrate that adding a new road segment or enhancing the capacity of an existing link in a congested network without considering the response of network users may actually increase network-wide congestion or user travel costs. Many authors have discussed these paradoxes (e.g., Braess (2005)). Pas and Principio (1997) and Penchina (1997) showed that the occurrence of the Braess-paradox depends on the link congestion function parameters and the demand for travel. Yang and Bell (1998) showed that the addition of a new road segment to a road network may actually reduce the potential capacity of the network. They also showed that this paradox can be avoided by introducing network spare capacity into the network design problem. Yin and Ieda (2002) give an example of the reliability paradox, which is a reliability version of Braess's paradox. It shows that increasing the capacity or decreasing the variability for some links may actually lead to a less reliable network. These paradoxes, illustrate the difficulty of designing road networks.

6.4 Formulation of the robust road network design problem

In this section the robust road network design problem is presented as a bi-level problem in which the lower level problem involves two decisions: route choice and trip/demand choice (destination and mode choice). This is the formulation that is used in the remainder of this thesis. In the top level, the actual network design is made. The network design depends on the flow, which is determined in the lower level by means of a stochastic dynamic traffic assignment. This assignment not only depends on the capacities (and road types and speeds) that are determined in the top level, but also on the demand. The demand depends on the travel times, which are an outcome of the dynamic assignment and, therefore, also depends on the top level, in which the capacities (and road types and speeds) are determined.

(Top level)

Maximize $Z_1(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D})$

$s.t.\quad G_1(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) <= 0$

where $f = f(\mathbf{l}^{\mathbf{new}}, \mathbf{D})$ is implicitely defined by the lower level (route choice) and

$D = D(\mathbf{f}, \mathbf{l}^{\mathbf{new}})$ is implicitely defined by the lower level (trip choice)

(Lower level : route choice)

Minimize $Z_2(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D})$

$s.t.\quad G_2(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) <= 0$

(Lower level : trip choice)

Minimize $Z_3(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D})$

$s.t.\quad G_3(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) <= 0$

In the above formulation, Z_1 is the objective function of the top level problem and l^{new} is the decision vector of the top level problem. G_1 is the constraint set of the top level problem. In the lower level problem, Z_2, f, and G_2 are respectively the objective function, the decision vector, and the constraint set of the route choice. Finally, Z_3, D, and G_3 are respectively the objective function, the decision vector, and the constraint set of the trip choice. Trip choice refers to the destination and mode choice of travellers.

In the following section, we discuss the choices that we made with respect to the objective function, decision variables, and constraints in more detail for each level. The algorithm that was used for solving this problem is described in chapter 7.

6.5 Top level: network design

In this section, the top level network design problem is formulated.

Decision variables

We chose to use a discrete formulation of the network design problem. This implies that the number of lanes that have to be added to each link, $\mathbf{l}^{\mathbf{new}}$, is used as a decision variable on the top level. This variable can also be used for new links. A potential new link at first has 0 lanes. When $\mathbf{l}^{\mathbf{new}}$ is larger than 0, a new link is constructed.

We chose to use this discrete decision variable, and therefore a discrete formulation of the network design problem, because this comes closest to practice. In practice, a decision is made about the number of lanes of a road. This is, for instance, either one or two lanes, and not 1.2 lanes. The choice for a discrete decision variable has consequences for the solution algorithm that is used, as is explained in the next chapter.

By using this discrete decision variable, the following robustness measures can be modelled:
- *Adding spare capacity to an existing link*. Spare capacity is added by adding additional lanes.
- *Adding spare capacity to alternative routes*. Spare capacity is added by adding additional lanes to existing alternative routes, or new additional routes can be created.
- *Adding buffers*. Buffers can be modelled by including very short links at locations that are logical for buffers. In the previous chapter, we explained that buffers can be added to prevent spillback and to regulate the inflow. Since regulating inflow is more a dynamic traffic management measure than an infrastructure measure, only spillback buffers are considered in our formulation.

In the next chapter, we show that not only the number of lanes, but also the maximum speeds and the road type can be changed in the network design process. These link characteristics are, however, not explicitly included in the mathematical formulation, because they are optimized by expert judgement and not by means of a mathematical model.

Objective function
Below, the objective function Z_1 is shown. The benefits of the generalized travel costs under regular conditions (= total travel time cost benefits *TTCB* + total distance related cost benefits *TDCB*) and the reliability/robustness benefits (*TCVB)* minus the infrastructure costs, including maintenance costs (*TCI)* are maximized.

$$\max_{l^{new}} \quad Z_1(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) = TTCB(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) + TDCB(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) + TCVB(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) - TCI(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) \tag{6.1}$$

The total travel time costs *TTC* [euro] are computed by formula (6.2). The total travel time is a summation of the equilibrium travel time $\bar{\tau}_{jk}^{rs}$ times a value of time α^{vot} for the departure flow f_{jk}^{rs} from each origin *r* to each destination *s* for each path *j* and each departure time interval *k*. Since the departure flow is expressed as a flow per hour, it has to be multiplied by the size of the departure time interval η. Both the flow $f_{jk}^{rs}(\mathbf{l}^{\mathbf{new}}, \mathrm{D})$ and the travel times $\bar{\tau}_{jk}^{rs}(\mathbf{l}^{\mathbf{new}}, \mathrm{D})$ are results from the lower level problem. They depend on the number of new lanes $\mathbf{l}^{\mathbf{new}}$ and the demand **D**. In order to keep the notation simple, these dependencies are not shown explicitly in formula 6.2, and are not shown in formulas used in the remainder of this chapter.

$$TTC(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) = \sum_{r,s,j,k} \eta f_{jk}^{rs} \bar{\tau}_{jk}^{rs} \alpha^{vot} \tag{6.2}$$

The benefits of the total travel time costs (TTCB) are computed by using the rule of half on the OD-level, in which the travel time costs $\bar{\tau}1_{jk}^{rs}$ in the adjusted network are compared to the travel time cost in the original network $\bar{\tau}0_{jk}^{rs}$ by taking into account the changes in demand (formula 6.3) .

$$TTCB(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) = \sum_{r,s,j,k} \tfrac{1}{2}\eta\,(f1_{jk}^{rs} - f0_{jk}^{rs})(\bar{\tau}0_{jk}^{rs} - \bar{\tau}1_{jk}^{rs})\alpha^{vot} \tag{6.3}$$

The total distance related costs *TDC* (euro) are computed by formula (6.4). The total distance related costs are a summation of the distance travelled over each road type w from each origin r to each destination s for each path j and each departure time interval k multiplied by the average driving costs per kilometre α_w^{vod} and the average costs of external effects per kilometre α_w^{ext}[32]. Both the flow f_{jk}^{rs} and the distances travelled are results from the lower level problem. They depend on the number of new lanes $\mathbf{l}^{\mathbf{new}}$ and the demand **D**.

$$TDC(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) = \sum_{r,s,j,k,w} \eta\, f_{jk}^{rs}\, \overline{dist}_{wjk}^{rs}(\alpha_w^{vod} + \alpha_w^{ext}) = \sum_{r,s,j,k,w} \eta\, f_{jk}^{rs} E_{jk}^{rs} \tag{6.4}$$

The benefits of the total distance related costs (TTCB) are computed by using the rule of half on the OD-level (formula 6.5).

$$TDCB(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) = \sum_{r,s,j,k} \tfrac{1}{2}\eta\,(f1_{jk}^{rs} - f0_{jk}^{rs})(E0_{jk}^{rs} - E1_{jk}^{rs}) \tag{6.5}$$

The total costs of vulnerability *TCV* (formula 6.6.) are a summation of the travel time loss $(\tau_{jk}^{rs}(i_a) - \bar{\tau}_{jk}^{rs})$ caused by all incident types i that happen with a probability p_i on link a multiplied by the number of vehicles that experience that delay $\eta\, f_{jk}^{rs}$. The travel time losses are computed for each path j from origin r to destination s. The probability of an incident depends on the number of vehicle kilometres driven. Therefore, it has to be multiplied by the number of vehicle kilometres driven on the link $\omega * v_{at} * L_a$ in order to compute the expected number of incidents. Finally, the total travel time loss as a result of incidents has to be multiplied by an average value of robustness α^{voro}. The reliability/robustness benefits (*TCVB)* are computed by taking the difference between the total costs of vulnerability in the original network and the costs of vulnerability in the adjusted network.

$$TCV(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) = \sum_{i,a,t} p_{ia}\,\omega v_{at} L_a \sum_{r,s,j,k} \left(\eta\, f_{jk}^{rs}\left(\tau_{jk}^{rs}(i_a) - \bar{\tau}_{jk}^{rs}\right)\right)\alpha^{voro} \tag{6.6}$$

The total investment and maintenance costs *TIC* (formula 6.7) can be computed by multiplying the new number of lanes $\mathbf{l}^{\mathbf{new}}$ by the variable costs per kilometre of that specific link c_var_a. If the link is new ($\delta_a^{new} = 1$), additional costs c_fixed_a are added. By making the infrastructure costs link specific, the location of the link can be considered. In this way, the fact that it is more expensive to add lanes to a road (or construct new roads) over water than it is to add lanes to other roads can, for instance, be considered.

$$TIC(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) = \sum_{a \in A} \left(l_a^{new} c\_\mathrm{var}_a + \delta_a^{new} c\_fixed_a\right) \tag{6.7}$$

For each of the cost/benefit components in the optimization function, the net present value is computed by using the following formula: NetPresentValue_cost = cost*(1-(1+ *dr*)$^{-n}$)/*dr,* in which n is the number of years for which there are costs or benefits (set to infinity) and *dr* is the discount rate.

[32] The external costs can be computed in more detail if needed ─ for instance, by making them dependent on the level of congestion and the location.

Constraints
Several constraints can be added to the network design problem. One often added constraint is a budget constraint. In formula (6.8), a budget constraint is shown. We prefer not to add this budget constraint, because we would like to find the optimal investment strategy for which the benefits are higher than the costs, regardless of the cost. Of course, in practice the budget might be a restriction. Therefore, in the next chapter we present a solution algorithm in which the budget constraint (6.8) can easily be added.

$$TIC(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) \leq \text{Budget} \tag{6.8}$$

Other restrictions that can be added are restrictions that relate to the fact that a network structure has to be logical. The number of lanes on a road structure cannot change too frequently. In fact, it is usually preferable that the number of lanes remain constant between two large intersections. Another constraint could be that a road have the same number of lanes in each direction. These kinds of structure-related constraints are not included in our formulation. Instead, the network structure is chosen in such a way that 'strange' network configurations have a low chance of occurrence. For instance, if we want the number of lanes between two intersections to be constant, the network is modelled by only one link instead of multiple links. Furthermore, in the solution algorithm there is an opportunity to correct for all kinds of 'strange' network configurations. These corrections can be made without pre-specified constraints because what might seem to be a logical constraint for one location does not have to be a logical constraint for another location.

Finally, the fact that the flows, travel times, vulnerability costs, and distance-related costs are results from the lower level problem can be seen as constraints as well.

6.6 Lower level: route choice - dynamic user equilibrium

In this section, we formulate the route choice problem on the lower level network design problem. The traffic flows are determined given the number of lanes and the number of new lanes that result from the top level network design problem, and given the demand that results from the lower level network design problem (trip choice), which is presented in section 6.7.

We choose to use a dynamic formulation of the traffic assignment problem instead of the static formulation, because traffic flow dynamics are required when robustness issues are considered. This section is based on (Bliemer, 2005).

Decision variables
The total number of vehicles travelling from origin r to destination s that want to depart at time instant k is given by D_k^{rs}. One or more routes are available between this origin-destination (OD) pair (r,s) given by $\mathbf{J}^{\mathbf{rs}}$. The question is, how many vehicles will take each of the routes $j \in \mathbf{J}^{\mathbf{rs}}$ at each departure time interval k. The number of vehicles taking route j then defines the route flow rates denoted by f_{jk}^{rs}. This is the decision variable. Through link loading principles, the path flows determine the link loads and travel times.

Objective function and constraints
In practice, there are a lot of different drivers, all of whom have different route choice behaviours. We distinguish three different driver types v:

- Drivers taking a fixed route (driver type I)
- Drivers taking the perceived cheapest/fastest route (driver type II)
- Drivers taking the actual cheapest/fastest route (driver type III)

The following constraint holds: $f_{jk}^{rs} = \sum_{v \in V} f_{vjk}^{rs}$. The percentage of drivers belonging to each class are an input to the model.

In order to keep the formulation of the top level problem comprehensive, this distinction is not included in that formulation. When flows in the top level problem are considered, a summation has to be made over all driver types.

We assume that drivers make their route choice decisions based on the generalized costs of the available routes. To this end, route travel costs c_{vjk}^{rs} are defined, denoting the cost for vehicles using route j to travel from origin r to destination s departing at time instant k. Generally, the route travel costs have the following components:

$$c_{vjk}^{rs} = \alpha_v^{vot} \tau_{vjk}^{rs} + E_{vjk}^{rs} \tag{6.9}$$

where α_v^{vot} denotes the value of time for driver class v, τ_{vjk}^{rs} is the actual travel time of route j for driver class v, and E_{vjk}^{rs} denotes all other cost components of the cost function. These other costs can include flow dependent costs, time dependent costs, and constant costs. For example, dynamic tolls that are flow dependent, time-varying tolls that are time dependent (based on a fixed schedule), and constant tolls can be taken into account. It is important to point out that E_{vjk}^{rs}, and therefore c_{vjk}^{rs}, need not have an additive cost structure. This means that these route costs cannot be computed by adding consecutive link costs together — i.e., they directly depend on the origin and/or destination and on the route taken.

Below, we first consider drivers that take a fixed route; we then consider drivers having (im)perfect route information.

Drivers taking a fixed route (driver type I)

Drivers of type I take a fixed route and, therefore, do not consider the route costs, but simply take their fixed route. Since we do not model individual drivers but assume aggregate flows, route proportions y_{j1}^{rs}, denoting the (departure time independent) proportion of vehicles taking route j from origin r to destination s, have to be known. Furthermore, the proportion of drivers of each type, h_d, has to be known. Given this, the flow rates for driver type I on each route f_{jk1}^{rs} is computed by:

$$f_{jk1}^{rs} = h_1 y_{j1}^{rs} D_k^{rs} \tag{6.10}$$

Drivers taking the perceived cheapest/fastest route (driver type II)

Drivers of type II are assumed to take the perceived cheapest (or fastest) route, evaluating the generalized route costs c_{vjk}^{rs}. The route proportions y_{jk2}^{rs} for each departure time instant k are in this case not input, but depend on the perceptions the drivers have on the route costs. To simulate imperfect information, an error component ε_{jk}^{rs} is added to the actual route costs, yielding perceived route costs $\tilde{c}_{jk}^{rs}$:

$$\tilde{c}_{vjk}^{rs} = c_{vjk}^{rs} + \varepsilon_{vjk}^{rs} \tag{6.11}$$

We model the behaviour of the drivers as aiming to minimize their perceived travel costs. Different assumptions on the form of the error components ε_{jk}^{rs} lead to different approaches for computing the route proportions y_{jk2}^{rs} . We assume that ε_{vjk}^{rs} are identically and independently Gumbel (Extreme Value type I) distributed, resulting in a simple multinomial logit (MNL) model to be used to compute these proportions:

$$y_{jk2}^{rs} = \frac{\exp(-\mu c_{jk}^{rs})}{\sum_{j^1} \exp(-\mu c_{j^1k}^{rs})} \tag{6.12}$$

where $\mu \geq 0$ is the scale parameter of the MNL model, which can be seen as a route dispersion parameter. If μ is small, then drivers have an inaccurate perception of the route travel costs, so there is a wide dispersion over the available routes. On the other hand, if μ is large, then they have good knowledge of the route travel costs, so most travellers will choose the actually cheapest/shortest routes and will not end up widely dispersed among the available alternative routes. Implicit in the assumption that all error terms are independent is that overlap between routes may not correctly be taken into account. As a result, routes (and therefore links on those routes) that have a large overlap receive too much flow, and the other routes do not get enough. This distortion may be corrected by simply using an extended version of the MNL model, e.g. C-logit or path-size logit, or by more advanced choice models, e.g. nested logit, paired combinatorial logit, etc. (see Hoogendoorn-Lanser et al., 2004).

Using the route choice proportions y_{jk2}^{rs} and the known proportion h_2 of drivers of type II, we can calculate the route flow rates as follows:

$$f_{jk2}^{rs} = h_2 y_{jk2}^{rs} D_k^{rs} \tag{6.13}$$

The route flow rates f_{jk2}^{rs} are determined by the route travel costs c_{jk}^{rs}. However, these travel costs depend on the route travel times τ_{jk}^{rs}, which in turn depend on the dynamic loading of the route flow rates f_{jk2}^{rs}. This compartmentalization, which essentially yields a fixed point problem, makes the problem difficult to solve. Following the equilibrium law of Wardrop (1952), when no driver has the incentive to switch routes anymore (i.e., all travellers choose their perceived cheapest route), then the system is said to be in user equilibrium. This dynamic user equilibrium (DUE) is usually called a stochastic DUE, indicating that a stochastic perception error is used. It can be shown that the route choice problem in which all drivers choose their perceived cheapest route can be written as the following variational inequality (VI) problem in which the equilibrium route flow rates $\bar{f}_{jk2}^{rs}$ are to be determined (see also He, 1997):

$$\int_{k \in K} \sum_{r,s} \sum_{j \in J_m^{rs}} \left(\bar{f}_{jk2}^{rs} - h_2 y_{jk2}^{rs} D_k^{rs}\right) \frac{\partial c_{jk}^{rs}}{\partial f_{jk2}^{rs}} \left(f_{jk2}^{rs} - \bar{f}_{jk2}^{rs}\right) \geq 0, \forall \quad f_{jk2}^{rs} \in \Omega \tag{6.14}$$

where Ω is the set of feasible route flow rates defined by

$$\Omega = \{f_{jk2}^{rs} \geq 0 : \sum_{j} f_{jk2}^{rs} = h_2 y_{jk2}^{rs} D_k^{rs}\} \tag{6.15}$$

The set Ω expresses the feasible flow rates bounded by two constraints ─ the non-negativities of the flow rates and the flow conservation constraints. These latter constraints require the driver type II travel demand to be satisfied; i.e., all vehicles should be assigned to a route. Solving the VI problem for the equilibrium flow rates $\bar{f}_{jk2}^{rs}$ solves the route choice problem for drivers taking the perceived cheapest/fastest route.

Drivers taking the actual cheapest/fastest route (driver type III)
Drivers of type III take the actually cheapest (or fastest) route, and are assumed to have perfect knowledge of the route costs. This driver type is essentially a special case of the previous driver type, in which the driver takes the perceived cheapest route. If the error component is zero (i.e., $\varepsilon_{jk}^{rs} = 0$), the drivers are said to have perfect knowledge. In order to compute the route proportions y_{jk3}^{rs}, (6.12) is still valid, although we are now considering the limiting case in which $\mu = \infty$. The driver behaviour is now modelled such that all drivers try to minimize their own actual route travel costs. The result will be a deterministic DUE in which no driver has the incentive to unilaterally change routes. Using the dynamic extension of Wardrop's first principle, this means that in equilibrium all used routes have equal and minimal route costs. Mathematically:

$$\bar{f}_{jk3}^{rs} \geq 0 \Rightarrow c_{jk}^{rs} = \pi_k^{rs} \tag{6.16}$$

where f_{jk3}^{rs} are the route flow rates of driver type III, and π_k^{rs} is the minimum route travel cost for users travelling from *r* to *s* at time instant *k*:

$$\pi_k^{rs} = \min_{j \in J^{rs}} \{c_{jk}^{rs}\} \tag{6.17}$$

Equation (6.16) states that a positive route flow rate for a certain vehicle class (i.e, the route is used by that vehicle class) implies that this route has minimum route cost for that vehicle class. Hence, the route flow proportions y_{jk3}^{rs} have to be determined in such a way that the equilibrium condition in Equation (6.16) holds.

As before, the equilibrium route flow rates $\bar{f}_{jk3}^{rs}$ depend on the route costs c_{vjk}^{rs} , while these route costs depend on the route flow rates through the dynamic network loading model, yielding a fixed point problem. The VI problem formulation given in Equation (6.14) can now be simplified to (see, for example, Bliemer and Bovy, 2003):

$$\int_{k \in K} \sum_{r,s} \sum_{j \in J_m^{rs}} \bar{c}_{pk}^{rs} \left(f_{jk3}^{rs} - \bar{f}_{jk3}^{rs} \right) \mathrm{d}k \geq 0, \forall f_{jk3}^{rs} \in \Omega \tag{6.18}$$

where $\bar{c}_{jk}^{rs}$ denote the route travel costs corresponding to route flow rates $\bar{f}_{jk3}^{rs}$. Solving the VI problem in Equation (6.18) for the equilibrium flow rates $\bar{f}_{jk3}^{rs}$ solves the route choice problem for drivers taking the actual cheapest/fastest route.

The flows, travel times, and distance related costs depend, through network loading principles, on the number of lanes of each link, which are a result of the top level problem, and on the level of demand, which is a result of the lower level problem with respect to trip choice (see section 6.7).

6.7 Lower level: OD demand – trip choice

In this section, the trip choice on the lower lever network design problem is formulated. In this level, the demand for car (and truck) trips is determined. This demand depends on the trip production and attraction of each zone, and on the mode choice.

Decision variables
We use one basic OD-matrix with a departure time profile. The OD-matrix that is used is representative for all periods. This means that the demand D_{km}^{rs} for each period k is a fixed but time dependent percentage of the demand for the representative period k^*. In the remainder of this section, the index k^* is left out the formulation of the problem. This means that $D_{k^*m}^{rs}$ is the decision variable of the lower level network design problem with respect to trip choice.

Objective function and constraints
We chose to use a gravity model formulation in which trip distribution and modal split happen simultaneously, because utility maximizing travellers make their choice of destination considering both the utility of the activities at the destination and the disutility of the travel to the destination. Of course, other formulations of demand models can also be used. Bovy et al. (2006) give an overview of several different approaches. In general, a choice can be made to model distribution and modal-split sequentially instead of simultaneously. Furthermore, a choice can be made to use a growth factor model instead of a gravity model, or a combination of both. Finally, many different distribution functions can be used. In practice, the choice among all these options should depend on travel patterns in the area of consideration. We chose to use a log-normal distribution function, since this is a formulation that is often used in practice.

The simultaneous trip distribution- modal split model is formulated by means of the following gravity model formulation:

$$D_m^{rs} = \varsigma_r \varphi_s Z_r A_s F_{rsm} \qquad \forall r \in \mathrm{R}, \forall s \in \mathrm{S}, \forall m \in \mathrm{M} \tag{6.19}$$

This implies that the number of trips D_m^{rs} from origin r to destination s by mode m is a function of the production of trips Z_r in zone r, the attraction of trips A_r in zone $s,$ and a mode specific accessibility/distribution function F_{rsm} ,together with parameters ζ_r and φ_s. These parameters can be seen as balancing factors (balancing the attractions and productions) combined with a measure of the average trip intensity in the area under consideration.

The production and attraction of trips can depend on the number of inhabitants, the number of jobs, the area, etc. We chose to make them solely dependent on the number of inhabitants and the number of jobs.

The production of trips Z_r in zone r is, therefore, determined by the following production function:

$$\begin{aligned} &Z_{rh} = \psi_h^z * residents + \upsilon_h^z * jobs \qquad \forall h \in \mathrm{H}, \forall r \in \mathrm{R} \\ &Z_r = \sum_h Z_{rh} \qquad \forall r \in \mathrm{R} \end{aligned} \tag{6.20}$$

In this function ψ_h^z and υ_h^z are parameters that determine the number of trips that are produced in zone r per resident and job for each trip purpose h.

The attraction of trips A_s in zone s is determined by the following attraction function:

$$A_{sh} = \psi_h^a * residents + \omega_h^a * jobs \qquad \forall h \in \mathrm{H}, \forall s \in \mathrm{S}$$
$$A_r = \sum_h A_{sh} \qquad \forall s \in \mathrm{S} \tag{6.21}$$

In this function ψ_h^a and ω_h^a are parameters that determine the number of trips that are attracted to zone s per resident and job for each trip purpose h.

Furthermore, we chose to use the following log-normal distribution function

$$F^{mo}(c_{mpk}^{rs}) = \beta^m \exp^{(\gamma^m \ln^2(c_{mpk}^{rs}+1))} \tag{6.22}$$

The generalized costs $c_{mk^*}^{rs}$ are the equilibrium costs for the lower level problem with respect to route choice for the mode car (and truck). They are a weighted average over all paths p in time period k^* (weighted by the path departure flows). The generalized costs for the other modes are assumed to be fixed.

A double constrained gravity model is used. This implies that the number of trips that start in zone r should be equal to the number trips produced in zone r, and that the number of trips that end in zone s should be equal to the number of trips that are attracted to zone s. This leads to the following constraints:

$$\sum_s \sum_m D_m^{rs} = Z_r \tag{6.23}$$

$$\sum_r \sum_m D_m^{rs} = A_r \tag{6.24}$$

Finally, the demand is determined given the number of lanes and the number of new lanes that result from the top level network design problem, and given the travel costs that results from the lower level network design problem with respect to route choice, as described in the previous section.

6.8 Summary: formulation of the robust network design problem

In this chapter, the robust road network design problem was explained in detail. We showed which dilemmas are important when a network is designed. Furthermore, we showed the objectives, decision variables, and constraints that are important. The robust road network design problem was formulated as a bi-level problem in which, at the top level, a network is designed that is robust against incidents, and in which, at the lower level, the route choice and trip choice are determined, given the network structure. By formulating this robust road network design problem, the 9th research question: "How can robustness be integrated in a network design method?" is partly answered. In the next chapter, we continue answering this question by presenting a solution algorithm that actually solves the robust road network design problem.

The formulation we present differs from formulations of the Network Design Problem in the literature, because it combines destination choice, mode choice, and dynamic route choice in the lower-level problem, and it includes short-term variations in supply caused by incidents in the Network Design Problem. The next section describes a solution algorithm for the formulated Robust Network Design Problem.

7 Design method – solution algorithm

7.1 Introduction

In the previous chapter, a mathematical model was presented that consists of three levels. It is generally known that the Network Design Problem is one of the most complicated problems in transportation. The literature presents many formulations and solution algorithms to solve this nonlinear, nonconvex mathematical program, which is difficult to solve optimally (Abdulaal and LeBlanc, 1979; Chiou, 2005; Davis, 1994; Friesz et al., 1993; LeBlanc, 1975). If the network and demand are simultaneously optimized for both regular and irregular circumstances, the problem becomes even more complex. Solving such a problem analytically on a large scale network is impossible. It is also undesirable. A model can never capture all decision variables that play a role in practice. Therefore, expert knowledge of researchers and stakeholders is useful in designing a network. Experts have extensive knowledge about the possibilities, impossibilities, and difficulties with improving 'their' network. The help of experts not only can improve the network design, it also increases the level of acceptance of the robust design by the different stakeholders. Methods like Infralab, IRVS, ARNO, and ARKO make use of expert knowledge in network design (see section 6.2).

In this chapter, an architecture for designing robust road networks is presented that combines the best of both worlds. Models are used to support the experts in the design process. In addition, some general design rules are specified as well, which can be of extra help and which make it possible to design based on capacities, and also to design road types and maximum speeds. In this way the following measures for improving the robustness of a road network can be included (see chapter 5):

- Adding spare capacity on a road.
- Creating new alternative roads or adding spare capacity on alternative roads.
- Unbundling of roads.
- Creating buffers.

In the next section, the framework of the solution algorithm is presented. The framework consists of three steps, which are described in more detail in sections 7.3, 7.4, and 7.5. In section 7.6 the functioning of the algorithm is demonstrated for a small test network. (In chapter 8, the algorithm will be applied to a large network.) The quality of the algorithm is analysed in section 7.7.

7.2 Framework of the solution algorithm

In this section the design method is presented. The method is summarized in Figure 7.1. Each step of the method is briefly explained below. In the following sections these steps are discussed more thoroughly.

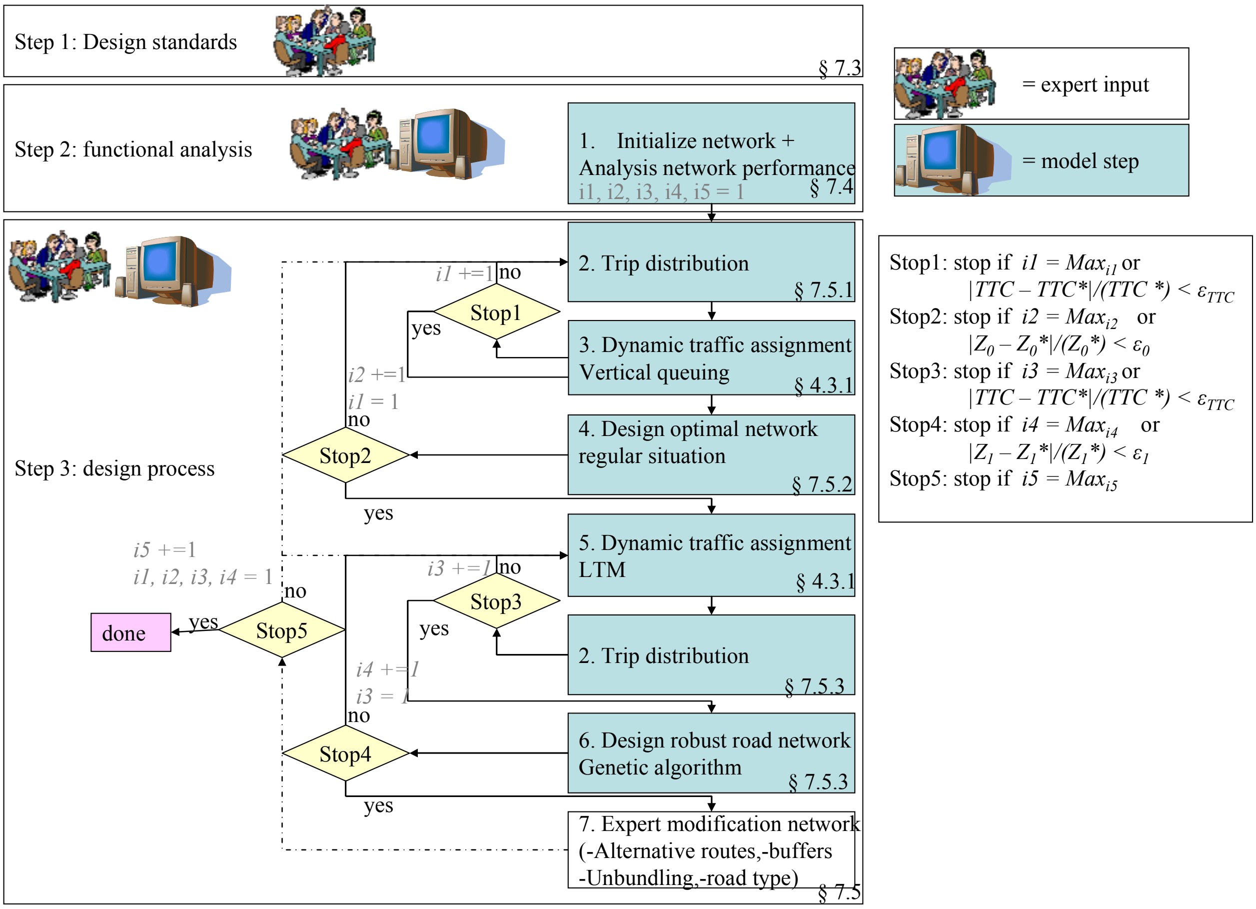


Figure 7.1: Design method

Step 1: *Design standards*. Designing a road network starts with specifying the design standards. The objective of designing/improving the network, the design variables, and the restrictions are specified. This can be done in half a day by experts in a workshop setting.

Step 2: *Functional analysis*. A functional analysis is an analysis in which the network performance of the reference network is evaluated. The reference network is the network that is used as a starting point for the network design. It could be the current network, but it could also be a planned future network. The choice for the reference network is made in the first step.

Step 3: *Design process*. In this step the robust road network is designed. The design standards and the results of the functional analysis are used as a starting point. A model can be used to get an idea of where adding capacity would be the most useful. At first the 'optimal' number of extra lanes for the regular situation is determined. Thereafter the 'optimal' number of extra lanes of spare capacity for improving the robustness is indicated. After this modelling exercise, a group of experts (these could be stakeholders and/or traffic engineers) have the opportunity to adjust the suggested network. They can indicate which alternative routes should be used, where capacity should be added, and how much capacity should be added there. Furthermore, they can add all kinds of other measures to improve the robustness of the road network. Of course, this will result in an expert dependent suboptimal design. However, since there are many different near optimal solutions, this is not necessarily a problem. Furthermore, when the stakeholders are involved who are responsible for the implementation of the design, their commitment to the design is important.

After adjusting the network, the performance of the network needs to be evaluated again. The quality of the robust design is tested with the same model as is used in the functional analysis. This evaluation also indicates what the remaining problems in the network are. A choice could be made to go back to the first design steps and change the network design slightly to eliminate or reduce these problems. This makes designing an iterative process.

There are several other feedback loops in the figure, which are intended to achieve an equilibrium between demand, travel times, and network design. The setup of the architecture is such that all building blocks can be used independently of each other, and the number of times that each building block is called can also be specified by the user. In fact, some building blocks can even be left out (for instance, the loop with the trip distribution model), which of course does imply that the user (or group of experts) has to make some basic assumptions on the outcomes of these skipped building blocks. In the following sections the building blocks are explained in more detail.

7.3 Step 1: Design standards

An objective is always required when a network is designed. Traditionally this objective includes:

- minimization of travel times under regular circumstances,
- minimization of number of vehicle kilometres,
- minimization of emissions,
- maximization of safety.

We added to this list the maximization of the robustness of the road network, by minimizing the number of vehicle loss hours due to small and severe incidents. The exact specification of the objective function is presented in chapter 6. However, the method is designed in such a way that it can deal with different objectives as well.

The design variables are also important ─ which changes can and which changes cannot be made to the network. In the modelling part, we optimize the capacity. To be more precise, we optimize the number of lanes of the existing network first in the regular situation and then with disturbances. However, the method allows for all kinds of other design variables in the phase in which the experts modify the network. They can add alternative routes (new routes), modify the road types, adjust the maximum speeds, reduce the number of lanes of certain roads if that is desired, separate through traffic from local traffic (unbundling), add buffers, add dynamic traffic management measures, and/or include flexibilities in the network.

In designing networks, there are many restrictions. One obvious restriction is the budget. Although the model might indicate that investing in regular extra capacity and spare capacity has a benefit/cost ratio that is larger than one and is thus profitable for society, it might very well be that the investment costs are too high. Therefore, it is possible to include a budget constraint. Other possible restrictions could be the number of lanes for each road that can be added or the location where new roads can be constructed (for instance, given certain spatial limitations).

Assumptions about the behaviour of drivers also have to be made. For instance, what percentage of travellers will take an alternative route during incidents? In fact, the modellers or the group of experts will have to make a choice about all of the parameters in the model. Appendix G describes the parameters.

In addition, a choice has to be made on which building blocks to use and the order in which they are called. This requires the modeller's knowledge ─ to indicate what is possible and what is not possible. The building blocks that are used have to be in line with the objectives of the network design. A choice could for instance be not to adjust the demand or not to optimize the capacity for the regular situation.

7.4 Step 2: Functional analysis

After the design standards are specified, a functional analysis is carried out by means of a dynamic traffic assignment with Indy. The network loading is done with the LTM module. In this way spillback effects are included. Furthermore, the vulnerability of the network is analysed with the MIC-module. This initial analysis is also the initial step in the modelling approach. It assumes that an initial demand matrix is known. However, if this is not the case, an OD-matrix can first be generated with a demand model combined with a dynamic traffic assignment. This is explained in more detail in section 7.5.1.

The initial assignment gives insight into the characteristics and the problems in the network. It shows:

- The speed ratios (speed as percentage of the maximum speeds) and densities per time step. These can be used to find the locations and severities of the bottlenecks in the network.
- The percentages of through traffic, traffic that has an origin or destination in the area on which the network design focuses, and local traffic.
- The vulnerability of the different links in the network.

Based on this functional analysis, the group of experts is asked to indicate where possible new (alternative) routes are to be located. Later in the process a decision is made on which of these potential new alternative routes are actually constructed.

7.5 Step 3: Design process

In section 3.3, we described the five design elements: prevention, redundancy, compartmentalization, resilience, and flexibility. In order to include these elements in the network design, we need to elaborate on them further. In order to do so, we first introduce four general design principles (Schrijver et al., 2008):

1. Healthy balance between supply and demand,
2. Building in options and flexibility,
3. Designing with flows *and* buffering in mind,
4. Form follows function.

Below, each principle is explained and an indication is given about which elements of robustness they support. These principles are intended to guide the design, especially during the phase when the experts are involved. Model runs should show whether or not the robustness of the network really improves after applying these principles.

1. Healthy balance between supply and demand

In a robust road network, in a normal situation, a 'healthy' balance exists between supply and demand. The acceptance of a certain degree of (local and temporary) congestion is then tolerable and even desirable: by accepting a limited amount of congestion, the highest peaks in the traffic demand are 'spread' across a slightly longer period, so that the costly infrastructure does not need to deal with this 'peak within the peak'. In this way, the network continues to function as it should, and a certain scarcity of traffic space is still created that keeps the mobility-generating effect of infrastructure within limits. In fact, it is not different from the queue at the checkout counter in the supermarket: if it is (extremely) busy in the shop, short queues are generally accepted, but if queues form at times when it is not so busy, this is a sign that structurally there are either too few checkout counters or too few counters open.

Just as the management team of the supermarket weighs the quality for the customer (which includes the option to go shopping at the busiest time of the day) against the costs of having extra checkout counters, the weighing of quality of accessibility against the costs of producing and maintaining the infrastructure is a policy-driven argument for the provider of the infrastructure, but an argument that can be based on solid cost-benefit analysis. The direct and indirect costs of more infrastructure can then be weighed against the direct and indirect costs of the absence of this infrastructure for consumers and producers. This weighing of factors concerns not only improved accessibility, but also the consequences for living comfort, use of space, and economic development.

In terms of increasing the robustness of the network, a 'healthy balance' means the following: the direct and indirect costs of adding capacity, including spare capacity, to the network to increase the <u>redundancy</u> must be in equilibrium with the direct and indirect benefits of less travel time loss as a result of incidents. The robustness will be increased only if the capacity (including spare capacity) does not result in a mobility-generating effect. If desired, this can be avoided by choosing the design speed of the infrastructure carefully, or by introducing road pricing. If the average speed of travelling from *A* to *B* under normal circumstances does not increase, no extra mobility is generated.

2. Building in options and flexibility

A robust road network offers several possibilities for getting from *A* to *B*. Alternative routes can run across the same network and across the underlying or even overlying network. They can also use other modes (public transport and bicycles for passenger traffic; inland waterways and railways for goods traffic). From the perspective of the road administrator, providing options is desirable. It leads to a levelling off of peaks, because road users can choose another route when faced with bad circulation on a certain route, even though this may be of lower quality (driving comfort), or involve travelling a longer distance. In addition to providing options, these alternative routes also offer back-up options (redundancy) in case a road section or intersection becomes temporarily or permanently inaccessible, whether expected or not. In such cases it is extremely important to provide the road user with the necessary information, in combination with the implementation of dynamic traffic management, which must ensure that the traffic flow over the alternative route – which is more intensive than normal – takes place effectively. In order to switch between the route alternatives, flexibility in the network is required. Park and Ride facilities can, for example, expand the options in the network by enabling a flexible switch from the road to rail.

Example:
If the traffic flows are evenly distributed across the network and the capacity is, therefore, also evenly spread across the network, disruptions can more easily be dealt with. One route can then act as a backup option for another (redundancy). Figure 7.2 shows an example of an unbalanced and a balanced distribution of traffic lanes across alternative routes. In the top situation (red arrows), there are two routes available for driving from south to north. One route has four lanes and the other route has two lanes. If something happens on the route with four lanes, there are only two lanes available on the other route (minus the capacity already being used) to take up the traffic from the other route. The middle situation (blue arrows) shows a more balanced distribution across both routes. Both routes now have three lanes, which means that they can function more easily as backup options for each other. The bottom situation (green arrows) is even more balanced still. A third route has been created here, so that all three routes have two traffic lanes. If an incident were to take place on one of the routes, there would always be four traffic lanes (minus the capacity already being used) available on the other two routes to take up the traffic from the affected route. It is important to note here that the routes are well interlinked.

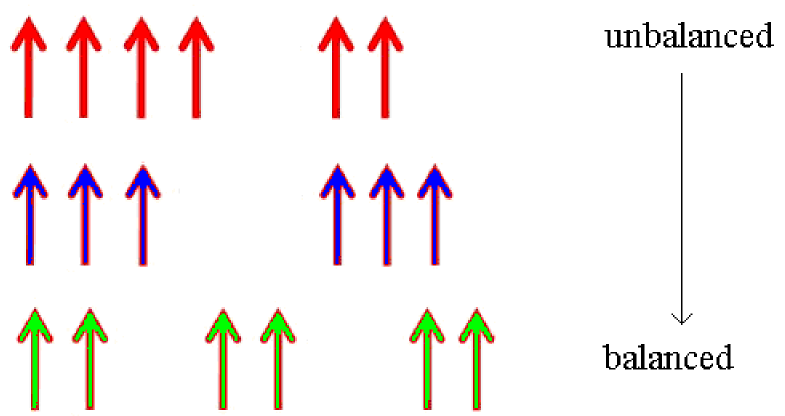


Figure 7.2: Balanced distribution of traffic lanes across routes

3. Designing with flows and buffering in mind
Because a certain amount of congestion is allowed for in a robust road network, when designing roads, intersections, and junctions, the buffer function of road sections must be fully taken into account, as well as the flow function. Each road section has its own 'natural' buffer capacity. Where this buffer capacity is insufficient, extra buffers can prevent congestion-forming occurring on other routes; they also serve to regulate the traffic in such a way that certain roads remain as congestion free as possible, such as through roads or the urban road network. Therefore, buffers result in a better compartmentalization of a network.

Since buffers also enable a network to recover more quickly from incidents, they make the network more resilient.

4. Form follows function

Motorway networks were originally designed for long-distance traffic. In practice, large portions of many motorways are used for short distances, because there is often no other route of sufficient quality available. Long distance trips make use of the same network. As a result, the two network functions interfere with each other. The form is therefore not optimally in line with the function, which results in an inefficient use of the traffic space. In a robust road network, form follows function. So, where necessary, physical barriers should be introduced between road networks for long distances and those for short distances. This makes the whole network, especially in urban areas with much traffic, more efficient (e.g. in terms of use of space), and enables the various functions to not work against each other (compartmentalization). As a result, traffic congestion on the network for through traffic does not directly lead to congestion for short-distance traffic, and vice versa. Additionally, by creating different independently functioning networks, the redundancy and resilience of the whole road network is increased, because both networks can function as backup options for each other. The road network therefore can recover more quickly from disruptions. Flexibility in the network may be required if a switch has to be made from the long-distance network to the short-distance network, and vice versa. The network for long-distance traffic needs fewer access roads and exits than the network for short-distance traffic. This means that some part of the long-distance traffic may no longer be able to switch to the network for short-distance traffic in the case of disruptions. A flexible short cut, such as a disaster crossing of a barrier (CaDo), can provide a solution. If an incident takes place on the network for short-distance traffic, flexible short cuts to the network for long-distance traffic are also desirable, since the network for long-distance traffic would otherwise not be able to offer a backup option for the short-distance traffic. Finally, the *form follows function* principle also leads to increased balance, because not all long- and short-distance traffic comes together on one motorway.

Example Rotterdam - Delft screen line

The form follows function principle is illustrated using the Rotterdam - Delft screen line (shown in Figure 7.3) as an example.

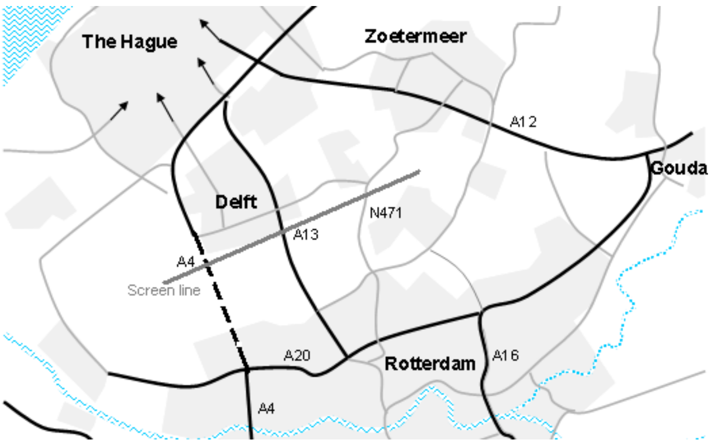


Figure 7.3: Screen line Rotterdam-Delft

In order to be able to process the traffic demand between Delft and Rotterdam expected in 2020, an average of seven lanes will be needed per direction. This number includes the limited extra growth in demand due to the improved network; however, this growth may be slowed down as the result of a road pricing policy. It can be computed that one extra lane is needed for spare capacity (see section 5.5). In Figure 7.4 the required number of lanes per road type is shown.

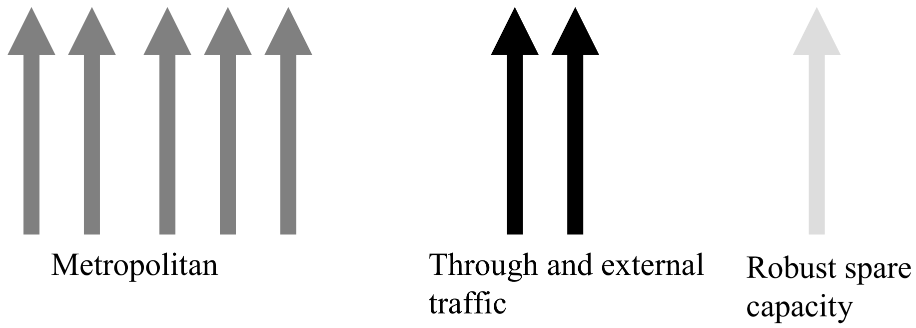


Figure 7.4: Rotterdam-Delft number of lanes needed in 2020

According to the Government's policy plans, in total 6 lanes are planned for 2020 for this section (see Figure 7.5): the current A13 (3 lanes) and N471 (1 lane), plus the A4 Midden-Delfland, which has yet to be built (2 lanes). This means that, based on theoretical capacity, there is one lane too few, and based on 'robust capacity' there are two lanes too few. Moreover, the distribution of the lanes over road types does not match the traffic demand: five motorway lanes are being provided, while only two such lanes are needed for through traffic; it would suffice to size the other lanes as metropolitan arterial roads.

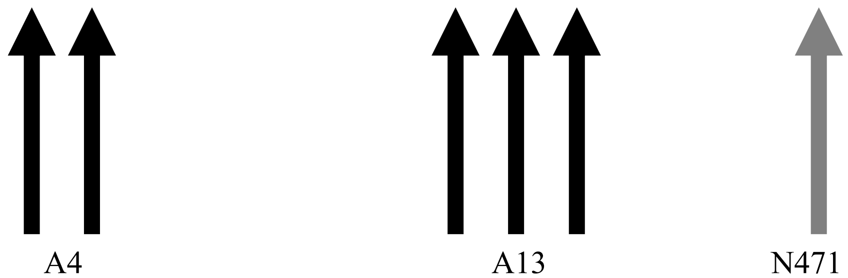


Figure 7.5: Rotterdam-Delft planned lanes for 2020

A robust network would have six lanes for metropolitan traffic that is equally distributed across the three 80 km/h routes ─ the 'N4', the 'N13', and the N471. Due to the equal distribution of capacity, a reasonable processing of traffic is still possible should an incident on one of the three routes lead to stoppages. This means that the N471 needs to be expanded to 2 x 2 lanes, but that the A13 can have one lane less in each direction, while the remaining 2 x 2 lanes could also be a little narrower. The space gained could be used for the realization of a south tangent-like public transport lane beside the N13 (OV), leading to a more robust public transport network in this corridor (see Figure 7.6).

In addition to these six lanes (in total) for metropolitan traffic, two lanes per direction for through traffic are still needed. We choose to locate these next to the A4, since that complies best with current policy. It means that the A4/N4 would get a total of four lanes in each direction – two for metropolitan traffic (80 km/h) and two for through traffic (120 km/h). That is two more than the number projected in the current plans.

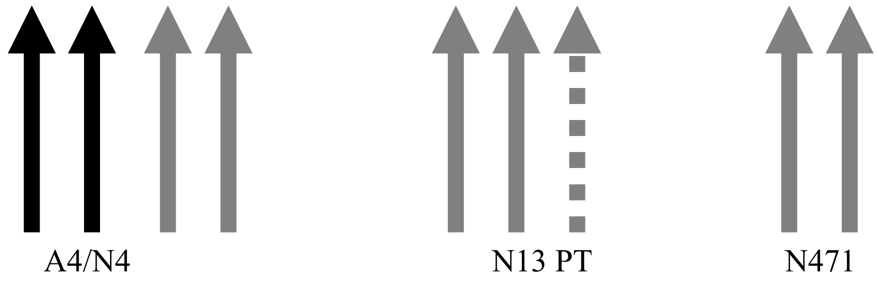


Figure 7.6: Rotterdam-Delft number of lanes in robust network in 2020

The four principles presented at the beginning of this section are the basis for the design process, as described in the following sections.

7.5.1 Create OD-matrix

The design method presented in Figure 7.1 contains two loops in which the OD-matrix (= the demand) is determined. Figure 7.7 repeats these loops.

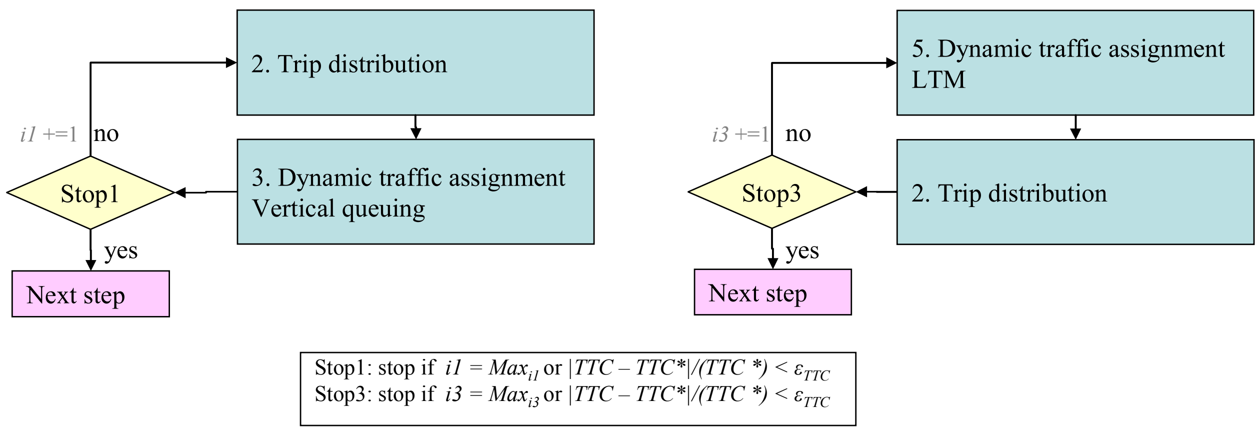


Figure 7.7: Create OD-matrix

One of the effects of adding extra capacity is that, in general, the travel times are decreased at first, which makes it more attractive to use the routes on which the travel times improve. This is not only true for drivers that already used that route, but it also becomes attractive for other travellers.

To be more precise the following effects could occur:

- *Location choice effects*: in the long term, people and companies might adjust their location choice based on the available infrastructure and the travel times.
- *Generation effects*: due to the decreased travel times, new trips are made. Generation effects are usually very small.
- *Distribution effects*: due to changes in travel times, some destinations might become more attractive than others. That is why the destination choice might change. Distribution effects are usually much larger than generation effects.
- *Modal split effects*: people might switch the mode that they use for a trip — e.g., from public transport to the car, if travel times by car decrease.

- *Departure time effects*: if travel times in the peak period decrease, travellers that used to travel just before or after the peak to avoid congestion might switch back to the peak period. Effects like this are called departure time effects.
- *Route choice effects*: changes in travel times can result in different route choices.

If these secondary effects are ignored, the travel time benefits might be overestimated. That is, adding capacity might seem to solve the congestion at first sight, but, due the secondary effects, congestion remains. However, this does not imply that adding capacity has no benefits, since it does improve mobility ─ more travellers are travelling according to their preferences.

We used a four step approach to creating an OD-matrix that captures generation, distribution, modal split, and route choice effects. Location choice effects and departure time effects are not included in the model. However, there is an option to adjust departure times and the locations (starting points of trips) if the group of experts thinks that this is necessary. This will have to be done based on the assumed reactions of travellers, spatial general equilibrium models, or elasticity models that are not (yet) included in the architecture.

The input for the demand model consists of a number of households and jobs per zone (if needed, extra zonal/demographic variables can be added). The number of trip productions and attractions are computed based on these variables. The productions and attractions are a function of the number of jobs and households per trip purpose. These functions were specified in chapter 6.

After trip generation, a gravity model is used to determine the distribution of trips and the modal split simultaneously by means of a distribution function. This function was specified in chapter 6.

Usually, a static assignment model is used after the trip generation, distribution, and modal split to compute the travel times on the road. Furthermore, an assignment algorithm is used to determine the travel times in the other transport modes. Instead of using a static model, we chose to use the dynamic assignment model Indy to compute the travel times on the road. Since this is a dynamic model, it needs time-dependent OD-matrices as an input. This can either be different matrices per time step, or a single matrix that is adjusted for different periods by multiplying the matrix by fixed departure time fractions. These fractions indicate the percentage of the demand that departs in each time interval. The fact that different time steps are used in the assignment means that either the trip generation, distribution, and modal split should also be done separately for these periods, or that one time period should be chosen as being the norm. The other OD-matrices then would depend (for instance, through the fixed departure times) on the normative matrix. We chose to use the last option, because this requires less data and less computation time. However, the algorithm can easily be adjusted for the first approach as well. Switching to the first approach might be necessary if the pattern of demand (trip distribution) differs significantly in different periods.

After the assignment model is run, the travel costs change. Therefore, the demand needs to be recalculated by running the distribution model again. This process continues until a maximum number of iterations $Max_{i1, i3}$ is reached or the total travel time costs do not change more than a certain ε. Convergence is guaranteed by using a method of successive averages on OD-trips.

The design process becomes a lot easier if the demand is assumed to be fixed. As explained at the beginning of this section, this might lead to serious underestimations of secondary effects. Nonetheless, there are two arguments that can be used for keeping the matrix fixed anyway:

- If road pricing is introduced, the volumes can be regulated. In this way, the level of demand can be kept more or less constant by increasing the price of travel. So far, only travel times have been mentioned as a factor in the decisions of travellers. But, in fact, generalized costs are also important. The generalized costs include travel times multiplied by a value of time, the out of pocket costs (for instance, fuel), and the distance related costs for road pricing[33]. Therefore, introducing road pricing might increase the generalized costs, which would compensate for the reduction in generalized costs resulting from reduced travel times. This still does not automatically imply that the demand will stay stable, especially not on the OD-level. It might, for instance, happen that some OD-relations benefit more from the reductions in travel times and other OD-relations experience higher increases in the generalized costs as a result of road pricing. The effects of road pricing heavily depend upon the type of road pricing and the intended goals of introducing road pricing.
- If the primary goal of the network design process is to make the network more robust, it could be argued that the travel times under regular situations do not have to be improved. However, this is a decision that has to be made by policy makers, since it gives a preference to reliable travel times over lower travel times and, therewith, increased mobility. If this decision is taken (in the first step of this design process), adjusting the maximum speeds (making them lower on selected links) could prevent the travel times under regular situations from decreasing, which means that the demand would not increase.

The above discussion illustrates that it is very difficult to make an accurate forecast of demand, especially for a forecast year far from the year in which the design is made. Therefore, it is advisable to work with different scenarios in which different matrices are used. Of course, this will result in different network designs. But it will give insights into which investment decisions are robust for different demand scenarios, and which investment decisions are useful only in certain scenarios. Finally, in the literature, some attempts have been made to work with stochastic demand to design a network that is robust for changes in the future demand (Ukkusuri et al., 2007). However, as far as known to the author, these approaches have been applied only on small theoretical networks.

7.5.2 Design optimal network for the regular situation

In this section, the building block for optimizing capacity under regular conditions is presented. Although the focus of this thesis is on robust road network design, which is more or less synonymous with irregular situations, the regular situation is very important. As is explained in the first design rule, there should be a healthy balance between supply and demand in both the regular and irregular situations. For example, if there is a lot of congestion in the regular situation, adding 'spare' capacity to improve robustness will actually not be spare capacity, because in practice this capacity will be fully used in the regular situation. Therefore, before making the network more robust for irregular situations, attention should be paid to the regular situation.

[33] There are many other costs that have in impact on trip decisions – e.g., car ownership cost. But the reliability of travel times (or actually the costs of unreliability) could also influence the decisions. We intended to explain the general principles; our intention was not to give a complete list of costs.

A building block is included in the architecture that ‘optimizes’ the road capacity under regular conditions. This building block is largely based upon the model ‘ROADNet’ (Snelder, 2003), which can be used to design a road network from scratch. In (Snelder et al., 2007), this algorithm is explained in detail and is applied to redesign the Dutch road network. The results show what the road network should look like if we could start all over again, ignoring the existing road network. The model iteratively optimizes the capacity by means of a golden-section algorithm, and does an ‘all-or-nothing’ assignment to compute the travel times. This process continues until the algorithm converges (the total investment costs and travel time costs stay more or less stable).

For this architecture, the model has been adjusted in several ways:

- The model optimizes the capacities of existing roads. New roads cannot be added.
- A different pricing mechanism is chosen. In the previous version of ROADNet, a pricing mechanism was used that prices a road differently in different periods according to the number of users (exchange between congestion cost and cost for road infrastructure) (Meeuwissen, 2003). This mechanism was needed to be able to design a network from scratch. For the current purpose, a method is chosen that allows including a price per kilometre.
- It is now a discrete model that optimizes the number of lanes, instead of continuous model that optimizes the capacity. This change has been made in order to more consistent with the design model that is used to optimize the robustness of the network.
- The link travel times are now computed based on the adapted Smulders function (Smulders, 1988; De Romph, 1994), instead of on a piecewise linear travel time function.
- Instead of an ‘all-or-nothing assignment’, a dynamic user equilibrium assignment with vertical queues is used. This type of assignment is chosen since it captures the dynamics of time, which is necessary to be consistent with the optimization of robustness, in which time effects are very important. Secondly, vertical queues are chosen, because the location of the vertical queues exactly match with the bottlenecks in the network. These are the first places to be adjusted. The fact that spillback effects are ignored in the evaluation of the different network designs is, of course, still a simplification. Under specific circumstances this could lead to an underestimation of the extra capacity that is needed under regular circumstances. However, using a model with spillback effects could result in investments in illogical places and would make the design process a lot more complex and time consuming.
- In ROADNet, a fixed OD-matrix was used. This matrix reflected the number of trips that would ideally be made by car if the road network were ‘perfect’: no detours and no congestion. Since the aim was to design an ‘ideal’ network, this matrix was suitable for this purpose. The outcomes of the model showed that even in an ‘ideal’ network there is some congestion. This is because it is economically not efficient to design a network to handle the short peak periods. This would, therefore, have an impact on the demand matrix. But, it was shown that these effects were very small. For the purpose of improving an existing network, this approach is not suitable. Therefore, we included the possibility to iterate with a demand model, as is shown in Figure 7.8.

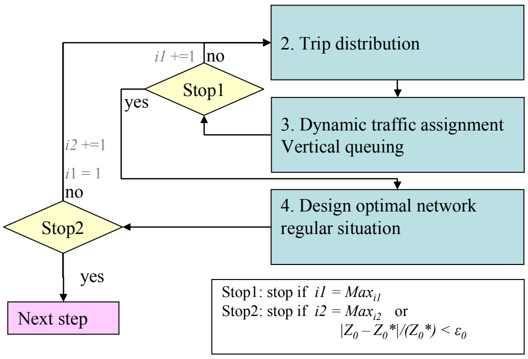


Figure 7.8: Design optimal network regular situation

In summary, the algorithm works as follows:
Below, the objective function Z_0 is shown. This is a simplified version of the objective function Z_1 presented in chapter 6, because the total costs of unreliability/vulnerability *TCV* are excluded.

$$\max_{l^{new}} \quad Z_{10}(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) = TTCB(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) + TDCB(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) - TCI(\mathbf{f}, \mathbf{l}^{\mathbf{new}}, \mathbf{D}) \qquad (7.1)$$

In the optimization of the network design under regular conditions, this objective function is optimized:

- First, the demand is determined by iterating between the trip distribution model and the dynamic traffic assignment with vertical queuing, as described above.
- Thereafter, for each link, the optimal number of lanes that is to be added is determined. If a lane is added, the travel times change, and possibly the OD-matrix. Ideally, the demand model and the assignment model should be run after each change in the network. However, since this takes too much time, the demand and the link flows are assumed to be fixed during the optimization of the number of lanes. The travel times are recomputed by using the adapted Smulders function (Smulders, 1988; De Romph, 1994), which is also used to calculate the travel times in the vertical queuing version of the dynamic traffic assignment model Indy. The adapted Smulders function has an uncongested part and a congested part, as is shown in Figure 7.9. The optimum number of lanes is found by a very simple optimization process: add a lane until the objective function no longer improves. This simple procedure works because the objective function per link is unimodal. (A function is unimodal on the interval [*a*,*c*] if the function is strictly increasing on the interval [*a*, *b*] in which *b* is the optimum, and the function is strictly decreasing on the interval [*b*,*c*].)
- After the number of extra lanes for each link is optimized, the trip distribution and assignment model are run again. This is necessary, because the optimum found by optimizing each link is not necessarily, or even most likely not, a network-wide global optimum.
- This process continues until a maximum number of iterations Max_{i2} is reached or the objective function Z_0 does not change more than a certain ε. It is not proven that this process converges. Furthermore, this process is not guaranteed to find the global optimum. The quality of the algorithm is examined in section 7.7. For the example network, the algorithm converged and the global optimum was found.

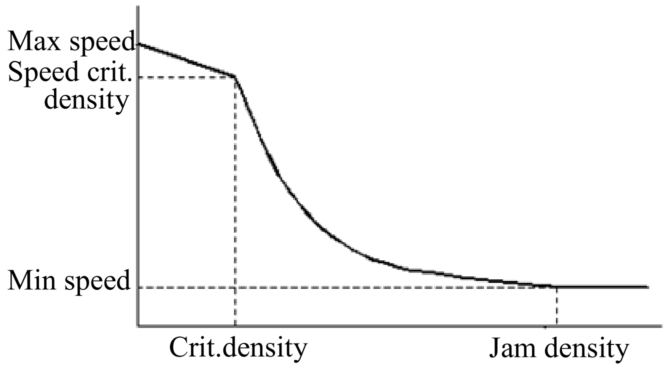


Figure 7.9: Adapted Smulders speed-density function

7.5.3 Design a robust road network

The first step in the design process for designing a robust road network is to compute the 'optimal' spare capacity in the network. This step starts with the results of the optimization of the network under regular conditions. (If this step was skipped, it starts with the original network.) At first, new routes are not considered. A discrete optimization algorithm is used that optimizes the number of additional lanes that are required as spare capacity. The objective functions were specified in the previous chapter, and can be changed if needed.

As was explained before, just like adding capacity for regular circumstances, adding spare capacity can result in many changes in choices by the traveller: location choice, number of trips made, destination choice, mode choice, departure time choice, and route choice. Again, these choices should ideally be dealt with simultaneously, but from a computational and practical point of view that is impossible. If, for every possible network, only the route choice were to be evaluated by means of a dynamic traffic assignment model with spillback effects, the computation times would very quickly add up to many years on a regular PC. Since waiting for that could be longer than the life time of the infrastructure, it is clear that a simplified algorithm is needed.

Figure 7.10 shows the design dilemma of optimizing the spare capacity of the network in the case in which the travellers stay with their initial choices. (This implies that they do not change location, destination, mode, departure time, or route choice.) The traffic flows from origin 1 to destinations 2 and 3. Disturbances on links 4 and 5 lead to spillback effects on links 3, 2, and 1. Disturbances on link 3 lead to spillback effects on links 2 and 1, and disturbances on link 2 lead to spillback effects on link 1. These spillback effects occur only in the case in which the demand exceeds the capacity. Furthermore, in those situations, delays occur on the disturbed link as well. The question now is where spare capacity can best be added. Spare capacity on link 1 prevents spillback to region 1 and can be useful for incidents on all links. Spare capacity on link 2 is useless for incidents that occur on link 1. Spare capacity on link 3 is useless for incidents that occur on links 1 and 2. Spare capacity on links 4 and 5 is useless for incidents that occur on links 1, 2 and 3. However, spare capacity on these links can prevent or slow down spillback to links 1, 2, and 3, which allows travellers to destinations other than destinations related to the disturbed link to complete their trips without additional delays.

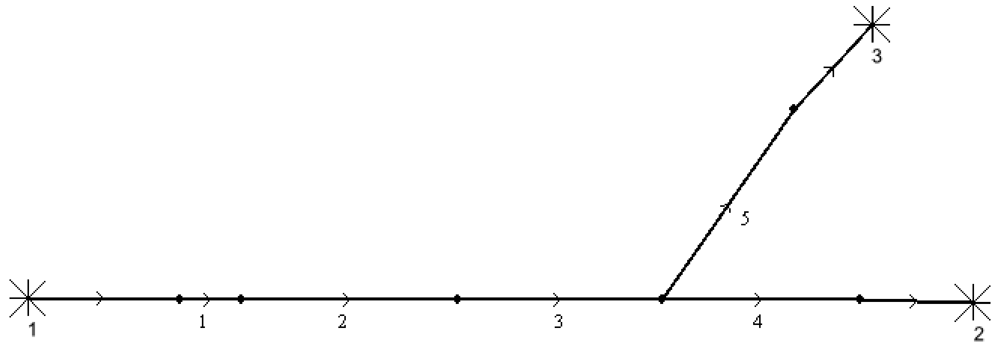


Figure 7.10: Example network

We chose to use a genetic algorithm to optimize the spare capacity (number of lanes) as was done by Li (2009), but using a different objective function. We made this choice because the design problem that we are solving is a discrete, nonconvex, and nonlinear problem. A genetic algorithm does not require continuity, differentiability, and unimodality of the evaluated functions. The genetic algorithm used to generate and evaluate the design strategies is illustrated in Figure 7.11.

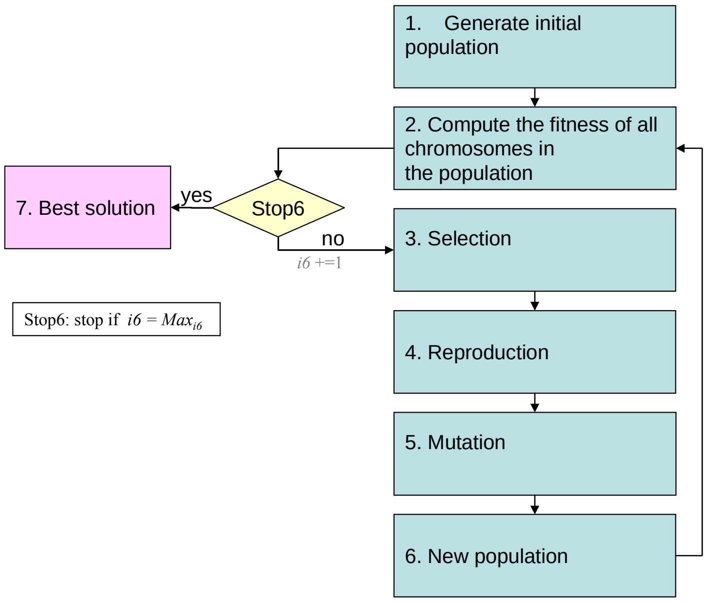


Figure 7.11: Genetic algorithm

The genetic algorithm starts with generating an initial population of chromosomes. A chromosome consists of a string of integer values that contains the number of additional lanes (design variable) for each existing link. The chromosomes are chosen by randomly selecting the number of additional lanes for each link. The maximum number of extra lanes for each link is determined by computing the total costs of the vehicle loss hours that occur downstream of the link as a result of all simulated incidents (chance*effect) that affect the link. These costs are divided by the costs of adding an extra lane to the link. The result is the maximum number of lanes that can be added to the link. There is no point in adding more lanes, because the costs will then exceed the benefits. However, it could be optimal to add fewer lanes, because, as is shown in the example above, spare capacity at other locations can reduce the impact of incidents as well. The maximum number of lanes that can be added can also be restricted by the user. Furthermore, two other solutions are added to the initial population. The first is a solution in which no extra lanes are added (0 lanes for all links). The second is a solution in which the maximum number of lanes is added to all links.

For all chromosomes in the initial population, the fitness is computed by evaluating the objective function. It is assumed (a simplification) that the flows under regular conditions do not change as a result of the extra lanes, which implies that Indy does not have to be run to compute the fitness of each chromosome. Instead, the fitness of a chromosome is determined directly with the marginal incident computation model (Corthout et al., 2009), which is an attached module to Indy, in combination with the alternative route approximation (see chapter 4). Since this module makes it possible to simulate the effects of many incidents on different locations within a very short time, it can be used in a genetic algorithm in which many evaluations have to be made.

The fact that alternative routes are not considered in the module is a problem when we look at robustness, because alternative routes are important for the robustness of a network (as shown in section 5.3). The module is likely to be improved in the future in such a way that alternative routes can be considered, for instance by changing the base cumulative flows. In the meantime, we use the simplified approach that is presented in section 4.3, which does not capture the dynamics of traffic completely, but does give an indication of the effects that occur when alternative routes are considered. The level of detail can in fact be considered to be consistent with the level of detail of the assumptions that have to be made anyway about the number of travellers that choose an alternative route.

After the fitness of all chromosomes has been computed, the genetic algorithm goes through a number of steps in which the population at the beginning of each step is replaced by another population. The chromosomes at each new generation are produced by means of genetic operators (denoted as reproduction and mutation). A detailed explanation of the working of genetic algorithms can be found in Goldberg (1989) and Deb (2002). The genetic algorithm stops when it reaches a predefined number of generations.

The solution space is very large. We can add some constraints to the problem to reduce the solution space, and thereby reduce the computation time. One of those constraints is a budget constraint. Other possible constraints are the number of extra lanes per link (given the number of lanes on upstream and downstream links). However, these constraints have not yet been implemented. We did add an evaluation method at the beginning of the genetic algorithm. It determines the maximum number of lanes that it is beneficial to add given the costs of vehicle loss hours on the link and upstream of the link[34] as a result of incidents on all links, and given the costs of adding a link. In this way, if, for instance, a link already has a lot of spare capacity, which prevents incidents to cause vehicle loss hours upstream of the link, the maximum number of extra lanes is 0, which implies that the link can be left out of the optimization.

[34]Upstream of the link implies that if an incident occurred on the link or downstream of the link that caused vehicle loss hours on the link, the vehicle loss hours caused by that same incident upstream of the link are also included. This is done because adding spare capacity to the specific link could prevent the queue from spilling back to those upstream links and, therefore, could prevent vehicle loss hours from occurring upstream of the link.

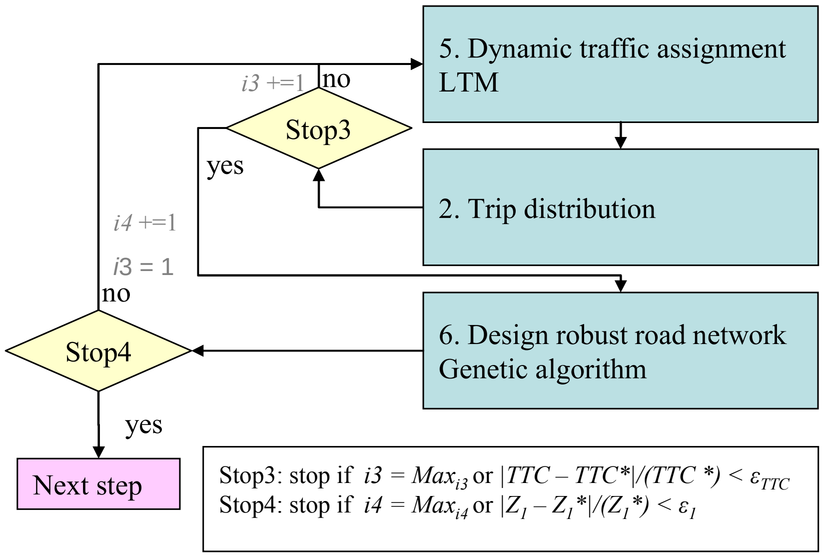


Figure 7.12: Design optimal network for the situation with incidents

Adding capacity affects the travel times under regular conditions. This means that, if capacity is added, the density decreases and the speed increases. The resulting lower travel times under regular conditions could result in higher travel time losses in case of major incidents in which, for instance, a road is fully blocked. This paradox is likely not to occur very often. In any case, the changes in travel times under regular conditions affect the number of vehicle loss hours that result from disruptions. (A vehicle loss hour is defined as the travel time loss compared to the regular situation.) In addition, route choice effects could occur. These two effects are captured by using an iterative approach, as is shown in Figure 7.12. Distribution and model-split effects are also considered by including the demand modelling building block.

7.5.4 Expert modification

The result of the network design for regular situations and for irregular situations (robustness) is an advice on extra lanes per links. However, as is explained before, adding new alternative routes has not yet been considered in the previous steps. Therefore, this has to be considered in this step. The extra lanes could also be built on alternative routes. This would improve the balance of the network, it might improve the redundancy of the network, and it also might improve the travel times under regular conditions.

The example in the previous section already illustrated how difficult it is to find the optimal spare capacity for all links if spillback effects are considered. Besides adding spare capacity on the routes that pass the incident location, spare capacity can also be added on new or existing route alternatives. In Figure 7.13 the same network is shown with two route alternatives (links 6 and 7). In this case, spare capacity can also be added on those links. If for instance an incident occurs on links 1, 2, 3, or 5, travellers can also use link 6 to reach their destination. Link 7 is a route alternative that can be used in case of incidents on links 1, 2, and 3 as well. This makes designing robust road networks even more complex. It shows that incidents (or other disturbances) on links have to be considered, the spillback effects of the incidents to all other links have to be considered, and all the links in the network have to be considered as possible locations for adding spare capacity. In other words, there is a many to many relation between the incident location, the locations on which the effects of incidents are noticeable, and the locations where spare capacity can be added. (Of course, the structure of the network excludes certain options.)

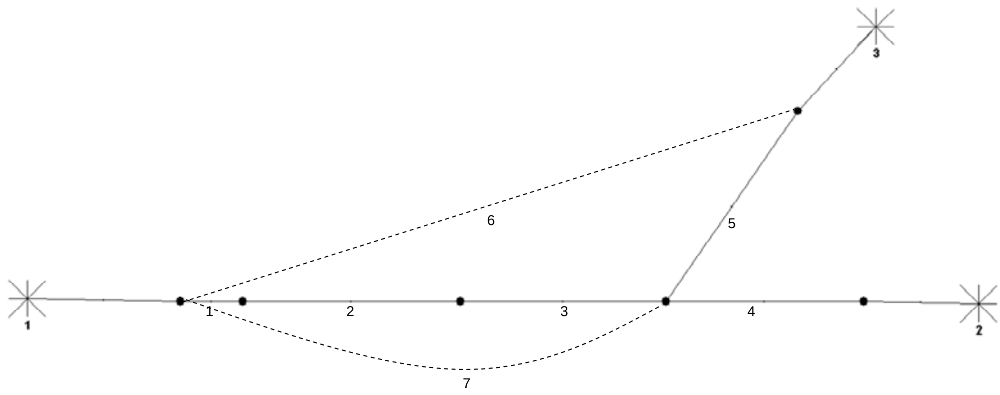


Figure 7.13: Example network with route alternatives

The alternative routes are considered in the genetic algorithm in a simplified way in evaluating the fitness of the chromosomes. However, (new) alternative routes are not considered as possible locations for adding spare capacity. Therefore, we add an extra step to the design process in which experts decide where to add capacity.

The experts decide where new alternative routes have to be added. Furthermore, in this step they can also alter the maximum speeds and road types of the design, add buffers, and unbundle roads. The design principle presented at the beginning of this section could be of help with this. Furthermore, some additional features were implemented that could be helpful:

- A procedure that shows the percentage of local traffic/through traffic. For the network of Delft, this results, for instance, in the plot presented in Figure 7.14.

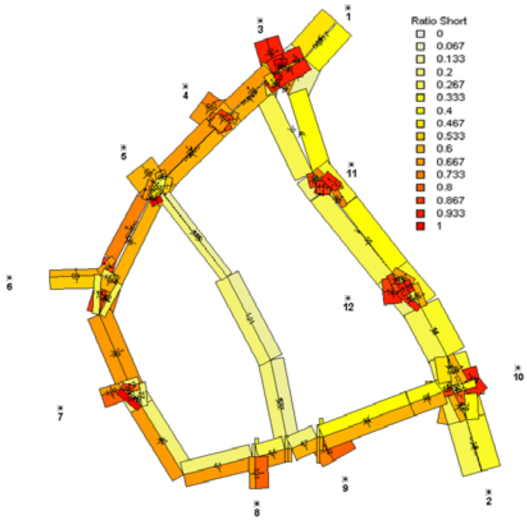


Figure 7.14: Ratio of short distance traffic in the Delft network (Ketelaars, 2010)

- An optimization module that determines which form of unbundling can best be applied. This algorithm is described by Ketelaars (2010). It results, for instance, in the advice below for the network of Delft:

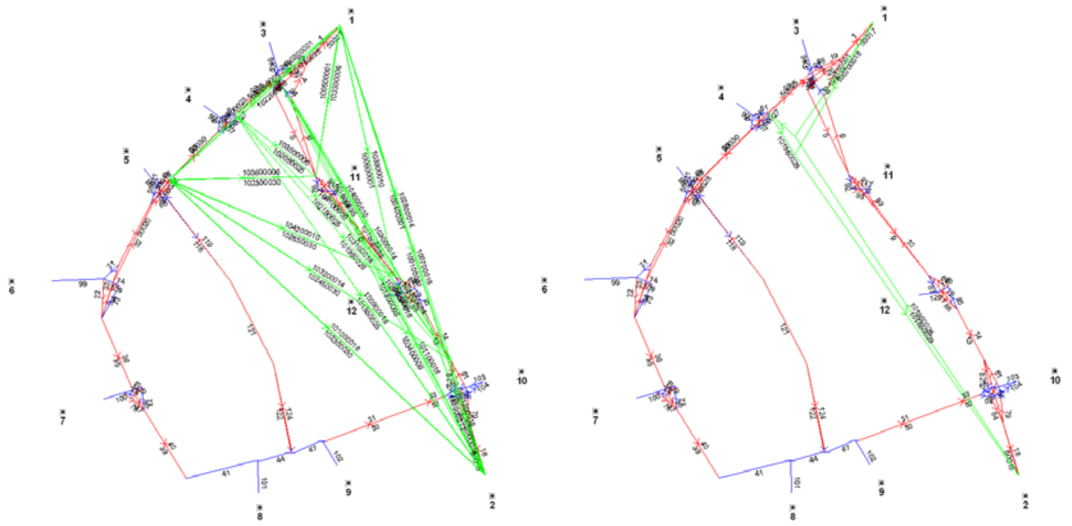


Figure 7.15: Advice on unbundling (left: large set of suggestions; right: chosen option) [35]

After the expert modification, a choice can be made to go back to one of the previous steps in the design process. This can be an evaluation of the new network under regular and irregular conditions. A decision can also be made to optimize the spare capacity again under regular and irregular conditions.

7.6 Example test network

In this section, the functioning of the algorithm is shown on a small test network. In Figure 7.16 a simple test network is shown. An example for a large realistic network is presented in Chapter 8.

Step 1: design standards
We chose to optimize the network according to the objective function that is specified in the previous chapter, without computing the net present value of the costs (this is not needed for this example). Furthermore, we chose to go through the full optimization process as specified in the previous section, which implies that changes in demand are considered and that the capacity is optimized for both regular and irregular conditions.

Furthermore, we chose to make at most 10 iterations between the distribution model and the dynamic traffic assignment model each time this combination of models is used. The maximum number of iterations between the optimization model for regular circumstances and the combination of the demand model and the dynamic traffic assignment model is 5. The maximum number of iterations between the optimization model for irregular circumstances and the combination of the demand model and the dynamic traffic assignment model is also 5. The loops stop when this maximum number of iterations is reached or when the model outcomes change less than 1%.

We chose to include the following design variables/measures: the spare capacity is optimized by the model and the modellers can change the structure of the network at the end of the process by adding new alternative routes. After this adjustment, the network is not optimized again. It is only evaluated. We included a small link at the beginning of the network (link 4),

[35] Source: Ketelaars, 2010

which has the potential to become a buffer, because the maximum number of lanes that can be added to this link is large.

We did not include a budget constraint.

Finally, the population size of the genetic algorithm was set to 50 and the number of generations was set to 20.

Step 2:Functional analysis
In the functional analysis, the network is analysed in detail. Before this is done, we present a more detailed description of the network and zones. There are 3 zones. The number of jobs and residents per zone are shown in Table 7.1.

Table 7.1: Zonal data test network

Zone	Jobs	Residents
1	0	85000
2	27500	0
3	27500	0

We assumed that there are two trip purposes:

- From home to work:
  - The number of trips produced for this purpose for each zone are computed as: 0.09*residents.
  - The number of trips attracted for this purpose for each zone are computed as: 0.20*jobs.

- From home to a place other than work:
  - The number of trips produced for this purpose for each zone are computed as: 0.10*residents.
  - The number of trips attracted for this purpose for each zone are computed as: 0.30*jobs.

The parameters that are used in these production and attraction functions are arbitrarily chosen. Of course, in realistic networks, these parameters have to be estimated. Furthermore, we assumed that there are two modes: car and public transport. The travel times for public transport are assumed to be fixed: from zone 1 to zone 2, 30.5 minutes; from zone 1 to zone 3, 32.2 minutes. For comparison: the free flow travel times by car are 15.1 minutes and 15.9 minutes, respectively. The travel times for the car are selected for a reference time step 90 minutes after the start of the simulation. In an equilibrium state, the travel times for the car are: from zone 1 to zone 2, 22.7 minutes; from zone 1 to zone 3, 23.6 minutes. These travel times are the equilibrium result of iterations between the demand models and Indy. In the equilibrium state (reached after 5 iterations) at the reference time step, 46% of all the trips are made by public transport and 54% of the trips are made by car. The number of trips between origin 1 and destinations 2 and 3 are shown in Table 7.2.

Table 7.2: Demand matrix reference time: trips per hour summed over both purposes

	Zone 2	Zone 3
Car: from zone 1	4262	4428
PT: from zone 1	3813	3647

The demand period is 3 hours, which is divided into time slices of 10 minutes. The following departure fractions per 10-minute time slice are used: 0.1, 0.3, 0.4, 0.5, 0.7, 0.9, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 0.9, 0.7, 0.5, 0.4, 0.3, and 0.1.

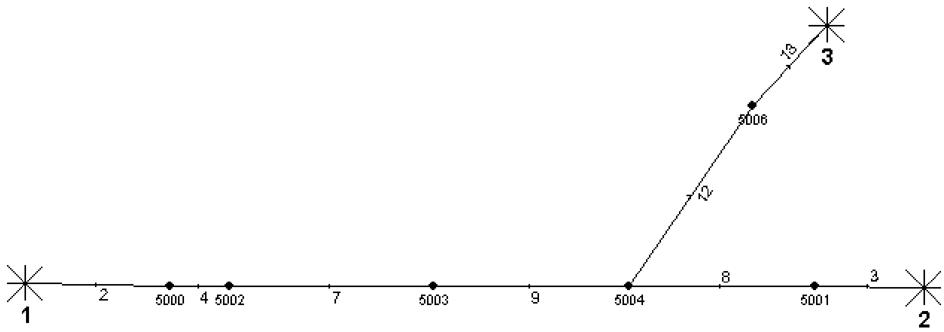


Figure 7.16: Test network

Table 7.3 shows the network characteristics. It can be seen that the feeder links/connectors have 10 lanes. The capacity of these links is not adjusted in the optimization process. In order to be sure that the feeder links are not bottlenecks, their capacities are set to 10 lanes.

Table 7.3: Network characteristics

Link	Length [km]	Lanes	Capacity [pcu/h]	Free flow speed [km/h]	Fixed infrastructure costs [€/week]	Variable infrastructure costs [€/pcu/km/week]
2	5.00	10	24161	120	-	-
3	5.00	10	24161	120	-	-
4	1.84	4	9664	120	5298.8	1.70
7	6.36	4	9664	120	5298.8	1.70
8	5.80	2	4832	120	5298.8	1.70
9	6.01	3	7248	120	5298.8	1.70
12	7.34	2	4832	120	5298.8	1.70
13	5.00	10	24161	120	-	-

In this initial equilibrium state, congestion occurs on link 7, because the capacity of link 9 is smaller than the capacity of link 7.

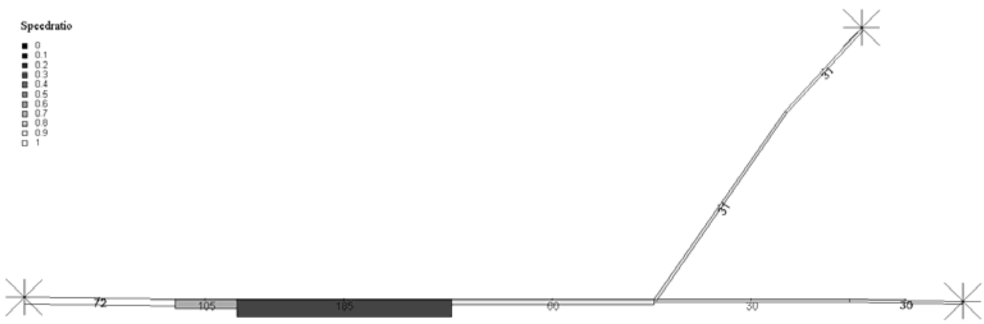


Figure 7.17: Density (numbers on link in pcu/km) and speed ratio (colour) at 90 minutes in the initial network

Figure 7.18: Expected cost (euros) of vehicle loss hours caused by incidents on each link in the initial network

Step 3: design process
After the initial assignment, the optimization process starts. First, the network is optimized for regular circumstances. This implies that disturbances are not considered. The optimization algorithm finds the following optimal solution: links 4 and 7 get two extra lanes, link 9 gets three extra lanes, and links 8 and 12 get one extra lane. These extra lanes are enough to remove the bottleneck. As a result of the improved travel times, the demand for car trips increases by 46%.

The figure below shows the development over the 'optimization iterations' of the infrastructure cost, the travel time benefits, and the cost-benefit on the y-axis on the left. On the y-axis on the right, the figure shows the number of trips per hour. On the x-axis, the solution for each 'optimization iteration' is shown between brackets, and the number of lanes for each link are shown in a string. For instance, [00050] means that links 4, 7, 8, 9, and 12 get 0, 0, 0, 5, and 0 extra lanes, respectively.

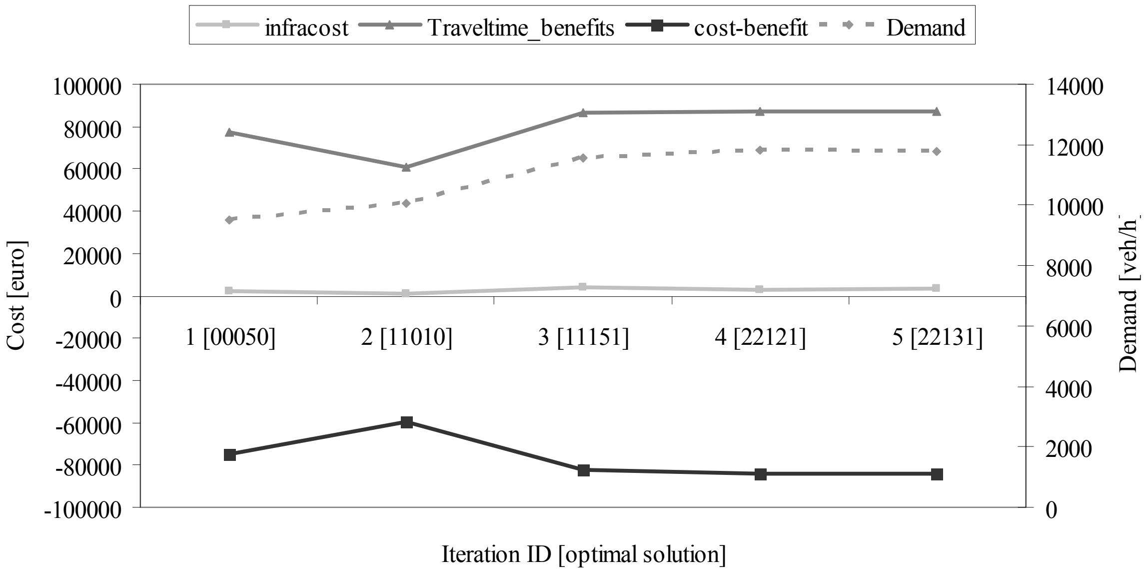


Figure 7.19: Optimization under regular circumstances

In the first iteration, 5 lanes are added to link 9. This was the bottleneck. Since the bottleneck was removed, the travel times decrease and the demand increases. However, the 5 additional lanes appeared to be too much. In the second iteration, the algorithm finds a solution in which one lane is added to links 4, 7, and 9. What the algorithm did not know during the optimization process was that, by doing this, the bottleneck reappears. Therefore, in subsequent iterations, 5 lanes are added to bottleneck link 9. Again, this appeared too much, so the number of extra lanes is reduced to 2. Also, links 4 and 7 get an extra lane as a result of the increased demand. Again this creates a bottleneck, so link 9 gets an extra lane. After this step, convergence is reached and the calculations end.

The remaining question is how much extra spare capacity has to be added to the network in order to make the network robust for disturbances. In order to answer this question, the genetic algorithm is run, followed by the dynamic traffic assignment and trip distribution models. All models are run iteratively until convergence is reached.

The first time the genetic algorithm is run, it finds the solution [11010]. The second time the algorithm is run, it finds the same solution [11010], so the algorithm is stopped, because convergence is reached. Between the iterations, the demand model and Indy are run iteratively. It is expected that the demand does not increase much anymore, because the level of demand depends on the travel times under regular conditions. The network was already optimized for regular conditions. In fact, it appeared that demand did not change at all in between the iterations.

Figure 7.20 shows the total number of lanes for each link after optimization. Figure 7.21 shows the density and speed ratio at the reference time step. In this network there is no congestion anymore. Figure 7.22 shows the expected costs related to the vehicle loss hours caused by incidents on each links. From this figure, it can be seen that link 12 is the most vulnerable link. This implies that incidents on that link cause the highest number of vehicle loss hours.

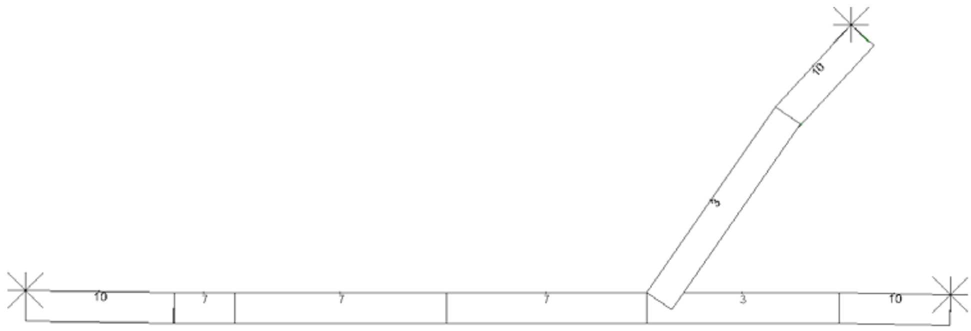


Figure 7.20: Total number of lanes

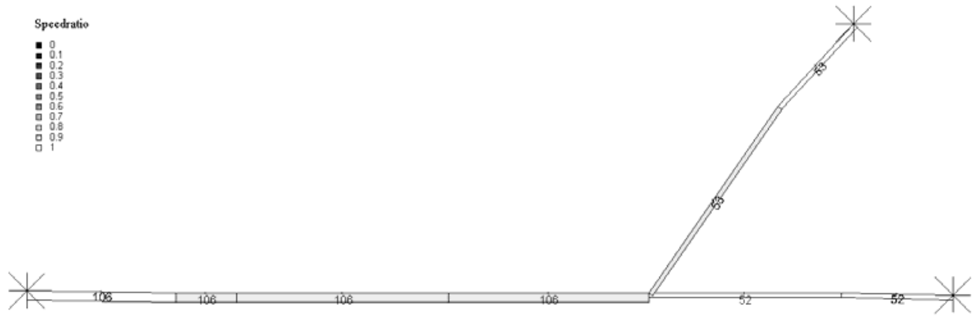


Figure 7.21: Density (numbers on link in pcu/km) and speed ratio (colour) at 90 minutes at the end of the algorithm

Figure 7.22: Expected cost (euro) of vehicle loss hours caused by incidents on each link in the optimized network

The optimal spare capacity, expressed by the optimal number of extra lanes, is shown in Figure 7.23. This spare capacity can be added to the links, as is shown in the figure. However, it can also be used to create a parallel road structure or to create completely new route alternatives. Expert knowledge can be used to define one or more alternatives based on the number of lanes that can be added. In Figure 7.24, an example is shown in which the local traffic (traffic travelling to zone 3 on top of the figure) is separated from the through traffic (traffic travelling to zone 2 at the bottom of the figure). The road for the local traffic gets 3 lanes, and the road for the through traffic gets 4 lanes.

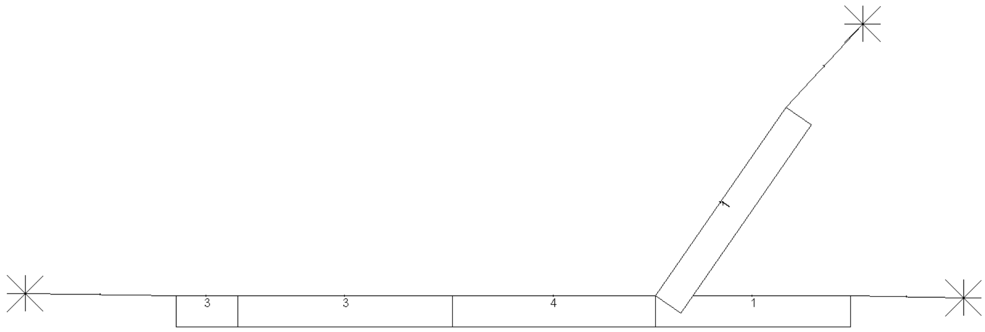


Figure 7.23: Number of extra lanes

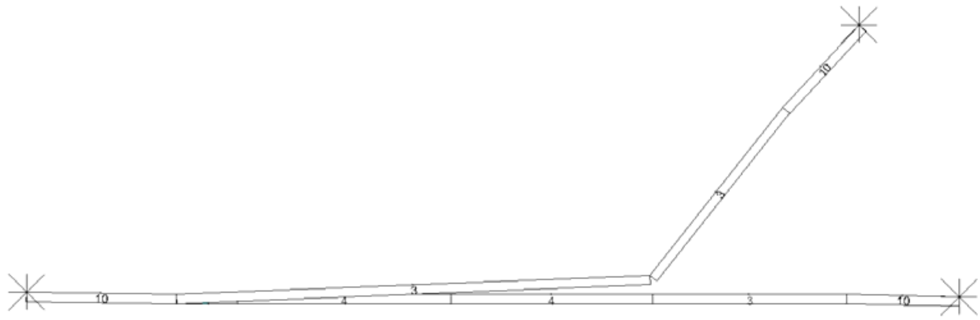


Figure 7.24: Number of lanes in variant structure

Figure 7.25 shows the performance of this network under regular conditions. It can be seen that the speed ratios are high, which matches the fact that the densities (numbers on the links) are low. This implies that, in this network, there is no congestion under regular conditions.

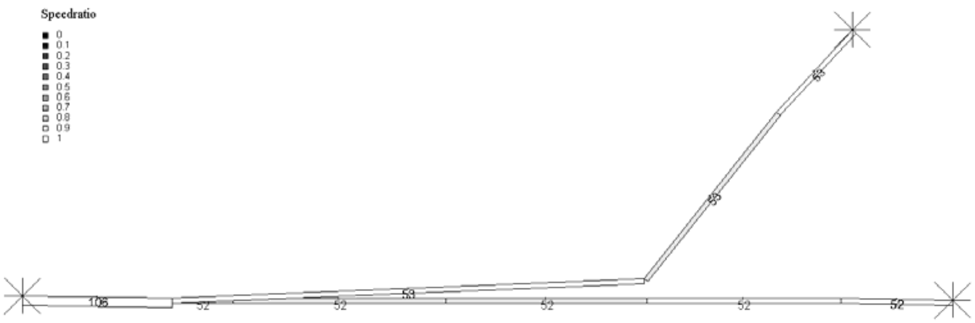


Figure 7.25: Density (numbers on link in pcu/km) and speed ratio (colour) at 90 minutes in the variant

The costs of vehicle loss hours caused by incidents are shown in Figure 7.26. The figure shows that the road for local traffic is more vulnerable than the road for through traffic. This is caused by the fact that there is less spare capacity available for this road.

Figure 7.26: Vehicle loss hours caused by an accident on each link

In Table 7.4, the network characteristics are summarized for the different optimization steps. All indicators are expressed for the simulation time period *T*. The costs in the last two columns are the costs of the extra spare capacity ─ i.e., the costs for the network improvement under regular circumstances are not included. The same is true for the benefits. The table shows that the investment costs are lower than the benefits. It also shows that the vehicle loss hours caused by incidents are significantly decreased ─ from 7480 to 5452 (= -27%) ─ which implies that the network is more robust. Despite the fact that the demand is 46% higher, the total vehicle loss hours caused by incidents are 27% lower.

Table 7.4: Network indicators

	Initial network	After reg optimization	After sparecap optimization	After network adjustment
Investment costs in *T* [euro]	0	3,496	1,044	1,044
Total demand in *T* [pcu]	17,106	24,975	24,975	24,975
Total travel time in *T* [hours]	6,087	6,458	6,458	6,458
Total distance travelled in *T* [km]	526,817	768,987	768,987	768,987
Average speed [km/hour]	87	119	119	119
Total travel cost under regular circumstances in *T* [euro]	512,755	712,057	712,057	712,057
Total travel cost under irregular circumstances in *T* [euro]	7,480	8,312	5,896	5,452
Total benefits under regular circumstances in *T* [euro]	0	29,751	0	0
Cost-benefits in *T* [euro]	0	-25,423	-1,372	-1,816

7.7 Quality of the algorithm

For the test network, it was possible to compute all the possible combinations of adding extra lanes to all links (with a maximum of 4 extra lanes per link) within acceptable computation time. This was done for the situation with elastic demand and fixed demand.

A comparison between the results from this "full computation method" and the results of the solution algorithm presented in the previous example provides insights into the quality of the solution algorithm. The solution algorithm is not guaranteed to find the global optimum, because it optimizes the capacity for regular conditions and for irregular conditions separately. It can occur that the benefits of adding an extra lane under regular conditions do not outweigh the cost of adding an extra lane, but that the benefits of adding, for instance, 0.6 extra lanes would outweigh the cost. However, since it is not possible to add 0.6 lanes, the lane is not added. When the spare capacity is optimized for irregular conditions, a similar problem could occur. Optimizing the capacity for both the regular and irregular situations at the same time might, however, justify the addition of one extra lane. This is a problem that might occur in larger networks, but it did not occur in this network. Furthermore, a (meta)-heuristic approach is used, which is by definition not guaranteed to find the global optimal solution.

The "full computation method" showed that the global optimal solution with elastic demand is [33141]. This is the same solution that was found by the solution algorithm, as was shown in the previous section. This shows that the algorithm is capable of finding the global solution. Of course, this does not prove that the algorithm will always find the global solution. But it does show that the algorithm works properly, and that it is at least capable of finding the global optimum in this small network. In case of fixed demand, the global optimal solution is [11021]. This means that, in case of fixed demand, the number of lanes that need to be added in order to create a robust road network is much lower, which demonstrates that latent demand plays a crucial role in road network design.

The full "computation method" also provides other insights into the network design problem. For instance, it shows that the top-10 solutions are: (1) [33141], (2) [33142], (3) [43141], (4) [43142], (5) [44141], (6) [44142], (7) [33131], (8) [43131], (9) [34141], (10) [34142]. Looking at this top 10, it can be seen that, in all elements of the top 10, link 8 has 1 extra lane and link 9, which is the bottleneck link, has 4 extra lanes (except for element 8 of the top 10). For all other links, the number of lanes varies either between 3 and 4 or between 1 and 2. An analysis like this, shows how likely it is that a small increase in the demand would require an additional lane. It can also show which infrastructure adjustments can best be made first. In other words, it can be of help in prioritizing the infrastructure adjustments. Insights like this can also be obtained by looking at the near optimal solutions that are found by the solution algorithm. For instance, it can be checked what the best solution was in each generation of the genetic algorithm.

Sensitivity analysis
The full computation allows for an easy sensitivity analysis of the algorithm for a change in infrastructure costs and the value of time. Table 7.5 shows the optimal solution that is found when the variable costs of infrastructure decrease or increase. In the optimization algorithm, the variable infrastructure costs are set to €1.70 per capacity unit (pcu) per kilometre per week. The table shows that if these costs increase to €5 per pcu/km/week, or decrease to €0.75 per pcu/km/week, the optimal solution changes. Of course, this is something that needs to be taken seriously, because it is quite difficult to make an accurate estimate of the infrastructure costs beforehand.

A more accurate sensitivity analysis shows that, in case of elastic demand, the optimal solution remains the same if the variable infrastructure costs stay in the range of €1.63-2.65 per pcu/km/week. This implies that, even if the infrastructure costs are underestimated by a maximum of 55.6% or overestimated by a maximum of 4.5%, the optimal solution remains unchanged. The margin of 55.6% underestimation is large enough to make it relatively certain that the investment will pay off. The smaller margin of 4.5% is less important because, an overestimation of cost will only lead to underinvestment. It is easier to increase the number of lanes, maybe a few years later, than to correct for too much capacity. In case of fixed demand, the range in which the optimal solution does not change is €1.00-1.82 per pcu/km/week, which means that the accurate estimation of the variable infrastructure costs are more critical than in case of elastic demand.

Table 7.5: Sensitivity analysis variable infrastructure costs without fixed costs

	Elastic demand		Fixed demand	
Variable infrastructure cost [€/pcu/km/week]	**Optcost**	**Solution**	**Optcost**	**Solution**
0.5	-187260	[44143]	-312250	[33042]
0.75	-186396	[44142]	-311759	[22031]
1.70	-183409	[33141]	-310481	[11021]
5	-175855	[32131]	-308757	[00010]
10	-165196	[32131]	-307460	[00010]
25	-134760	[22031]	-303571	[00010]
50	-99135	[21021]	-297088	[00010]
75	-69378	[11021]	-293115	[00000]
100	-43675	[00010]	-293115	[00000]
1000	-872	[00000]	-293115	[00000]

We did not include fixed costs in our analysis because the roads already exist. However, it could also be argued that these fixed costs should be included because large start-up costs are made when a road is to be extended. The solutions that are found if fixed costs of 5,298.80 €/km are considered are shown in Table 7.6. This changes the optimal solution to [32131] in the case of elastic demand, and to [00010] in the case of fixed demand. In the case of elastic demand, the bottleneck is removed and the capacity of the network is increased by two lanes (spread of two links after the split in the network). Furthermore, an extra lane is added at the first link, which can be seen as a buffer. In the case of fixed demand, only the bottleneck is removed.

Table 7.6: Sensitivity analysis variable infrastructure costs with fixed costs

	Elastic demand		Fixed demand	
Variable infrastructure cost [€/pcu/km/week]	**Optcost**	**Solution**	**Optcost**	**Solution**
0.5	-180784	[33131]	-309578	[11020]
0.75	-180240	[32131]	-309413	[11010]
1.70	-178208	[32131]	-309043	[00010]
5	-171180	[32131]	-308188	[00010]
10	-160521	[32131]	-306891	[00010]
25	-130808	[22031]	-303002	[00010]
50	-96497	[11021]	-296519	[00010]
75	-66770	[11021]	-293115	[00000]
100	-43107	[00010]	-293115	[00000]
1000	-872	[00000]	-293115	[00000]

A similar sensitivity analysis can be done by changing the value of time. The value of time for both the regular and irregular situations was set to €15/hour in the optimization process. Table 7.7 shows what happens if this value is varied for both situations between €5/hour and

€50/hour. Of course, the value of time for the regular and irregular situations could be varied independently. It is for example not unthinkable that the value of time for unexpected time losses is higher than for expected time losses. This table shows that the number of extra lanes goes up when the value of time increases, which is in line with the expectations. A more detailed analysis shows that the range in which the solution does not change is 9.7-15.7 €/hour in the case of elastic demand and 14.0 – 25.7 €/hour in the case of fixed demand. This shows that it is important to make an accurate estimate of the value of time.

Table 7.7: Sensitivity analysis the value of time

Value of time [€/hour]	Elastic demand		Fixed demand	
	Optcost	Solution	Optcost	Solution
5	-58541	[32131]	-102909	[00010]
10	-120760	[33141]	-206466	[11020]
15	-183409	[33141]	-310481	[11021]
25	-309149	[44142]	-518818	[11021]
50	-624069	[44143]	-1040733	[33042]

Convergence

The solution algorithm is an iterative procedure in which a dynamic traffic assignment is carried out, the trip distribution model is run, and the network is optimized. This raises the question of whether or not the algorithm converges and, if so, how fast it converges. Figure 7.27, which was also shown in the previous section, shows that the algorithm that is used for optimizing the capacity under regular situations converges after four iterations.

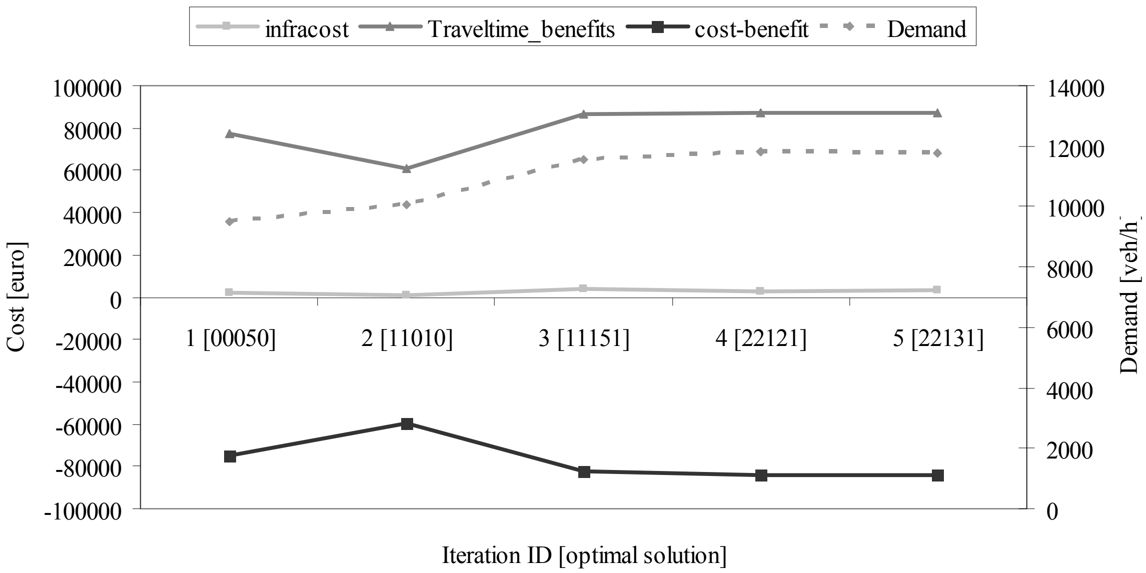
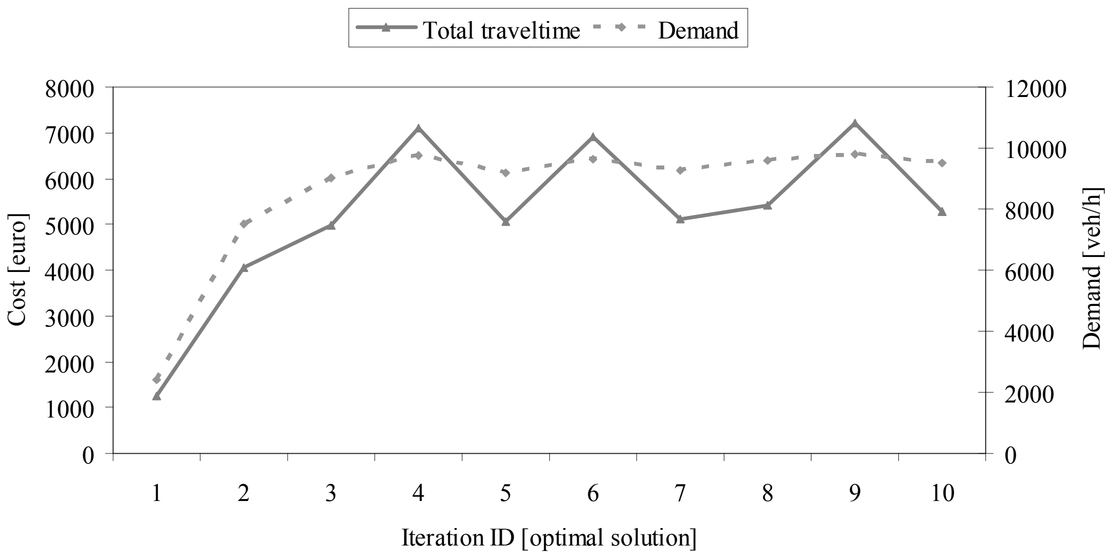
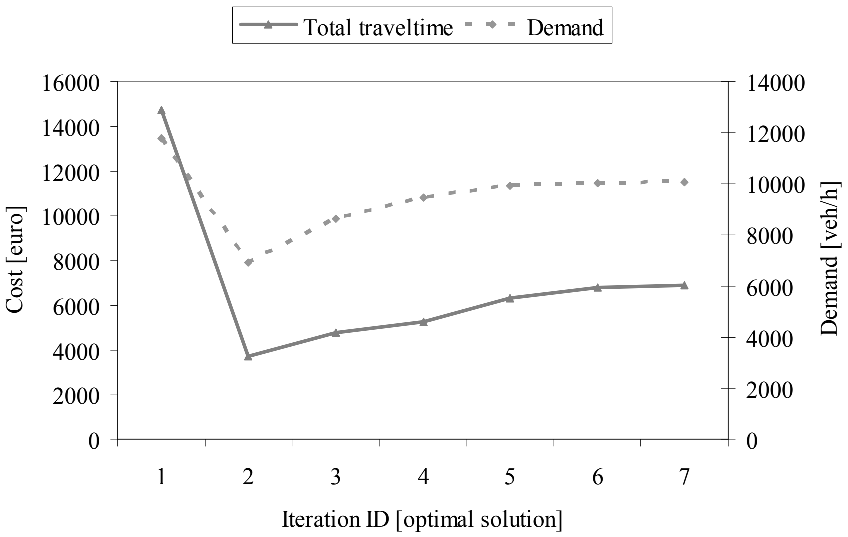


Figure 7.27: Optimization under regular circumstances

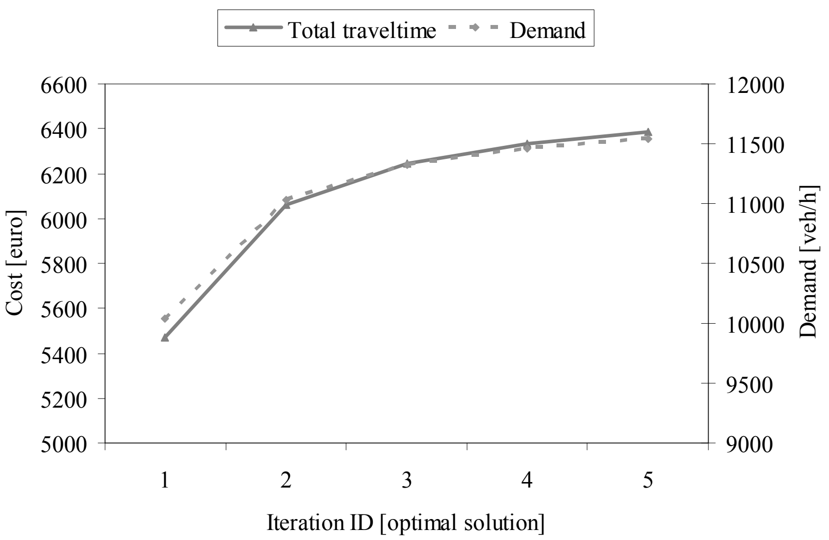
The dynamic traffic assignment model and the trip distribution model are run in between the iterations of the optimization process under regular circumstances. For the first three iterations, the convergence is shown in Figure 7.28.



1)



2)



3)

Figure 7.28: Convergences demand-assignment model in the first three iterations 1), 2) and 3) of the optimization model under regular conditions

In the first iteration there is a lot of flip-flop behaviour. Thanks to the method of successive averages, which was added to this "inner loop", the total level of demand converges quickly. However, the travel times react more slowly to small changes in demand. In this case, 10 iterations of the demand-assignment model were not enough to achieve full convergence. However, this is not a big problem, since the total optimization process did not stop there. The fact that full convergence was not yet reached only caused the optimization process to continue with a sub-optimal solution. In the three following iterations of the optimization procedure, the demand-assignment model combination converged in 7, 5, and 2 iterations, respectively. The number of iterations that are needed to converge decreases with the number of iterations of the optimization process, because the network changes become less significant the further the process gets.

Thereafter, the optimization algorithm under irregular circumstances is run, which produces exactly the same results in the second iteration as in the first iteration. This fast convergence is explained by the fact that the demand does not change much when the spare capacity for irregular circumstances is optimized, because this does not change the travel times under regular conditions (much).

Computation time
In total, the demand model was run 35 times, the dynamic traffic assignment model was run 38 times, and the MIC-model was run 3150 times. The total computation time was 1 hour and 20 minutes on a Intel(R) Core(TM)2 Duo CPU T7500 @ 2.2 GHz 2.00 GB of RAM laptop.

7.8 Summary: solution algorithm

In this chapter, a solution algorithm is presented that combines the expert knowledge of stakeholders with advanced modelling techniques. This makes it possible to design a large-scale robust road network with many design variables (= robustness measures) that is likely to be supported by all the stakeholders that were involved in the design process.

The method consists of three steps (with feedback loops), which are briefly described below:
- Step 1: *Design standards*. Designing a road network starts with a specification of the design standards by experts. The objective of designing/improving the network, the design variables, and the restrictions are specified.
- Step 2: *Functional analysis*. A functional analysis is an analysis in which the network performance of the reference network is evaluated. This is done based on expert knowledge of the different roads in the network and based on an evaluation model that evaluates the robustness of all the links in the network. The evaluation model first determines the travel times and congestion under regular conditions and analytically determines the cost of the travel time losses for four different incident types that occur with certain probabilities on each link. An approximation algorithm for alternative route choice during incidents is added to that evaluation method.
- Step 3: *Design process*. In this step the robust road network is designed. The design standards and the results of the functional analysis are used as a starting point. A model can be used to get an idea of where adding capacity would be most useful. At first the 'optimal' number of extra lanes for the regular situation is determined. Thereafter the 'optimal' number of extra lanes of spare capacity for improving robustness is indicated. After this modelling exercise, a group of experts (these could be stakeholders and/or traffic engineers) have the opportunity to adjust the suggested network. They can indicate

which alternative routes should be used, where capacity should be added, and how much capacity should be added there. Furthermore, they can add all kinds of other measures to improve the robustness of the road network. Some design principles were proposed that help the experts in their task.

After adjusting the network, the performance of the network needs to be evaluated again with the same evaluation model as is used in the functional analysis. Furthermore, this evaluation indicates what the remaining problems in the network are. A choice could be made to go back to the first design steps and change the network design slightly to eliminate or reduce these problems. This makes designing an iterative process.

The fact that experts get the opportunity to participate in the optimization process can be considered as a big advantage, since this makes it possible to enrich the model inputs and outcomes with knowledge of stakeholders that can never all be included in a model. Furthermore, this makes it possible to include many kinds of robustness measures in the optimization process. It is very difficult to include this many design variables in a full model approach. The fact that extensive use of models is made is an advantage, since this gives useful input about the vulnerability of different elements of the network for the part in which the experts are involved. The optimization results from the models gives suggestions about how much spare capacity is to be added on which locations. Finally, this gives quantitative support to the decisions that are made.

The solution algorithm allows optimizing the spare capacity that is needed in order to make a road network robust against incidents by considering traffic dynamics. This adds to the existing literature in which the spare capacity is optimized for static traffic flows under regular conditions without considering spillback effects. Furthermore, the proposed algorithm contributes to the literature about the network design problem by allowing multiple design variables (measures) to be considered for making the network robust against incidents. Finally, the proposed solution algorithm shows which simplifications are needed to solve the (robust) road network design problem for large-scale real-world networks.

Together with the formulation of the robust road network design problem in chapter 6, this chapter answers the 9th research question: "How can robustness be integrated in a network design method?" The solution algorithm was designed in such a way that it can be applied to large-scale networks. In order to make the algorithm applicable to large networks, some simplifications had to be made in the optimization algorithm:

- The capacity under regular conditions and irregular conditions is optimized separately. This might result in adding too few lanes, because there are cases in which the travel time benefits under regular conditions and the reliability benefits are individually not high enough to justify an extra lane, whereas the combined benefits might be sufficient to justify an extra lane.
- The regular capacity is optimized using a link-based heuristic. During this optimization process the network is evaluated only by means of dynamic traffic assignment with vertical queues in between the different iterations.
- The spare capacity is optimized with a network-based genetic algorithm. During this optimization process, the network is evaluated only by means of dynamic traffic assignment with a link transmission model in between the different iterations.
- The marginal incident computation model combined with alternative routes is a simplified model for analysing the effects of incidents. For instance, it does not consider downstream effects, and it does not consider effects on alternative routes.

An extensive test of the model on a small network shows that, despite the first three simplifications, the model is able to find the global optimum for this network. Thanks to the simplifications, the number of times that a full model run with a dynamic traffic assignment needs to be done stays small, which makes an application of the model to large networks possible.

8 Application to the network of Amsterdam and surroundings

8.1 Introduction

In the previous chapters, a design method was presented by which networks can be designed that are robust against disturbances. In this chapter the method is applied to Amsterdam and surroundings in order to show that the method can be applied to large networks, and in order to show how that can be done and which issues need to be addressed.

Recently in the Netherlands, visions on a robust road network design have been made for the areas The Hague- Rotterdam (Schrijver et al., 2008), Utrecht (Schrijver et al., 2009), and the 'Stadsregio Amsterdam' (Egeter and Snelder, 2010). The design steps that are used in order to make those designs resemble in many ways the design steps that were presented in this thesis. A functional analysis was carried out and a robust network design was made first in workshop settings and worked out back in the office by using, among others, the design principles shown in the previous chapter. The difference with our work is that models were not used in the design phase in these visions. In the evaluation phase, models were used only for the robust design of The Hague – Rotterdam area. The models that were used are less advanced than the ones presented in this thesis, although some elements of these models were used. In Appendix F, the application to The Hague – Rotterdam is presented. Despite the fact that less use is made of models, this application still provides good insights into how the non-modelling steps of the design method can be applied to a realistic network.

Because the non-modelling part is emphasized in the design for The Hague – Rotterdam area, this chapter emphasizes the modelling part of the design method for the network of Amsterdam and surroundings. In (Egeter and Snelder, 2010) a robust road network design is made for Amsterdam as well. However, models were not used, and the focus was on a slightly larger area around Amsterdam. The design resulted from several workshops in which experts were involved. In this chapter we use those results as the expert input for our robust design of Amsterdam and surroundings. Of course, when the design method that is presented in chapter 7 is applied in practice, an integrated approach should be followed. Therefore, the results presented in this chapter are purely intended to show how the method works on large-scale networks and which factors are important to consider in practical applications. Although the results might have some practical implications, they cannot be used directly because of the lack of interaction with the experts and because simplifying assumptions were made with respect to the number of model iterations, the future network developments, and future demand. If the method is to be applied in practice, more attention should be paid to these issues, but they are not relevant for showing how the method works. In section 8.3, the simplifications are explained and suggestions are made on how to deal with these problems in practice.

The chapter starts with a description of the network of Amsterdam and its surroundings. Section 8.3 deals with the design standards, section 8.4 deals with the functional analysis, and section 8.5 deals with the design process. The final section summarizes the main findings of the chapter.

8.2 Network description and calibration

Figure 8.1 is a map of Amsterdam and its surroundings. The map shows the road numbers that are used throughout this chapter. Amsterdam is the capital of the Netherlands, and had about 750,000 inhabitants in 2008. A lot of the traffic from the surrounding municipalities travels to, from, and through Amsterdam.

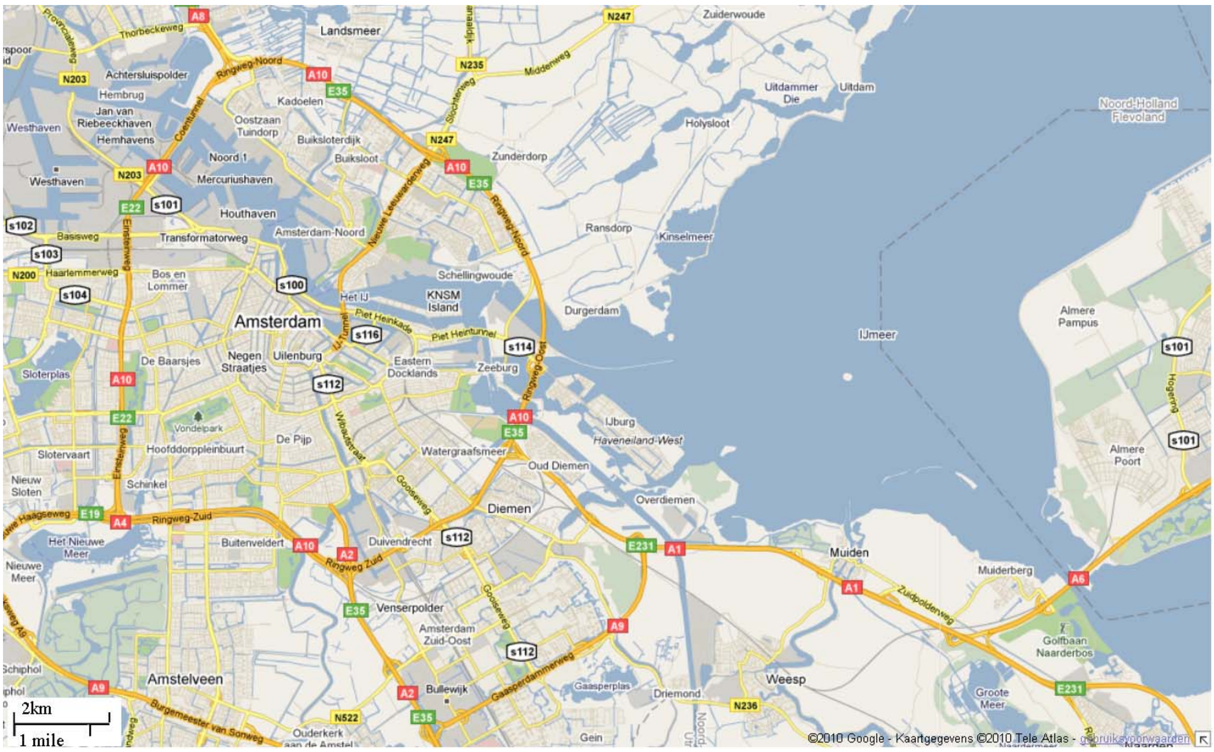


Figure 8.1: Map of Amsterdam and surroundings

Figure 8.2 shows the modelled network of Amsterdam. The network has 211 zones, 1856 links, and 1144 nodes. The total demand for the period 6.00 AM – 10.15 AM is 432,000 vehicles. Figure 8.3 shows the departure time profile. The network and matrices are based on RegioRegie/NRM-networks and matrices. The network and the matrices are loaded into Indy and calibrated using a dynamic OD-estimator (Chen et al., 2010). The model was calibrated using loop detector data from the motorways and several counts on the secondary road network for 2008. The model was validated by comparing its estimated travel times with travel times on several trajectories, and by comparing the locations and severity of congestion with the daily congestion patterns (Duijnisveld et al., to be published). This was done within the framework of analysis for the Schiphol-Amsterdam-Almere project (see Figure 8.7).

Figure 8.2: Number of lanes in the modelled network of Amsterdam

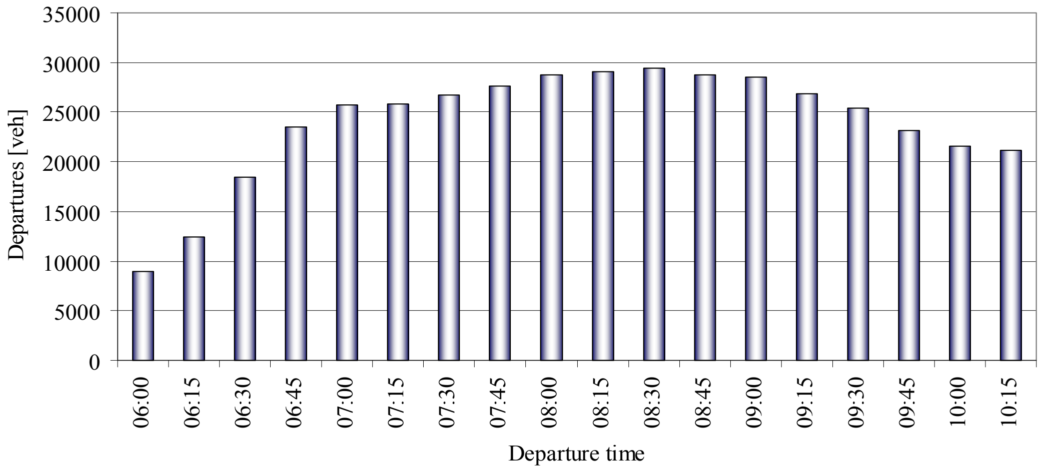


Figure 8.3: Departure time profile for Amsterdam

8.3 Step 1: Design standards for Amsterdam and surroundings

Designing a road network starts with a specification of the design standards by experts. The objective of designing/improving the network, the design variables, and the restrictions are specified. Below the choices that were made in the first step of the design method are presented.

Objective:
We chose to optimize the network according to the objective function that is specified in chapter 6. This means that a combined objective function is used in which the total costs (travel times, travel time reliability/robustness, external costs, and investment and maintenance costs) are minimized. Since optimizing robustness is one of the objectives, the design that is made is called *the robust design*. We chose to optimize the network for the morning peak period 6.00 – 10.00 AM of an average workday. Of course, in practice the other periods should also be taken into account.

Design variables:
We chose to include the following design variables/measures: the spare capacity is optimized by the model, and the modellers can change the structure of the network at the end of the process by adding new alternative routes. After this adjustment, the network is not optimized again and not evaluated either. The spare capacity is optimized only for the motorway links, which are shown in Figure 8.4.

Figure 8.4: Links on which the capacity can be optimized

The other measures can be applied to the other roads as well. We chose to optimize spare capacity only on the motorway links, because the motorways are the roads that probably need the most spare capacity. Of course, this spare capacity can also be created on lower level routes. However, this can be done after the optimization of the spare capacity on the motorways, based on the outcomes of the optimization process and the functional analysis.

Constraints:
We did not include a budget constraint.

Future network and demand level:
Of course, robust networks are designed for the future. In this design it is not really relevant which future year that is, because this chapter is merely intended to show how the method works. The intention is not to make a design that can be used directly in practice. This assumption makes it unnecessary to consider all changes in the network between 2008 and a future year like 2030. In practice, these changes should of course be taken into account. Furthermore, it makes it unnecessary to make an accurate forecast of the future demand. Whereas in practice several future demand scenarios need to be considered, for instance with respect to the development of inhabitants and jobs per region, we make a much simpler assumption: a growth in demand of 20%.

Despite the fact that the planned developments of the network are not considered, we present them below, because they contain relevant information with respect to improvements in the network that are possible, and because this information can be used in the end when our robust design is compared with the existing network. The planned developments for the period 2010 - 2020 are:

- *Coentunnel - Westrandweg (A8/A10/A5):* the Westrandweg is a new road with two lanes in each direction that connects the west side of the ring road of Amsterdam (A10) with the A5 and A9. The Coentunnel (west side A10) is upgraded from two lanes in each direction to three lanes in each direction and two contra flow lanes (see Figure 8.5).

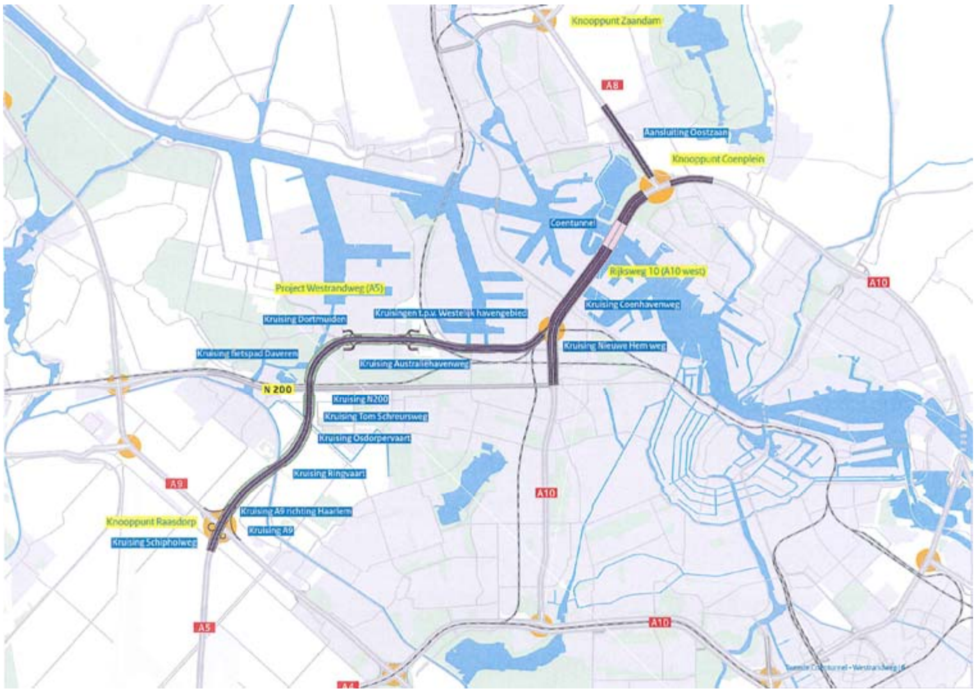


Figure 8.5: Investment 2010-2020 – Coentunnel and Westrandweg

- *Change of the intersection A4 - A9 Badhoevedorp*: a part of the A9 near Badhoevedorp is upgraded to three lanes in each direction and shifted (see Figure 8.6).

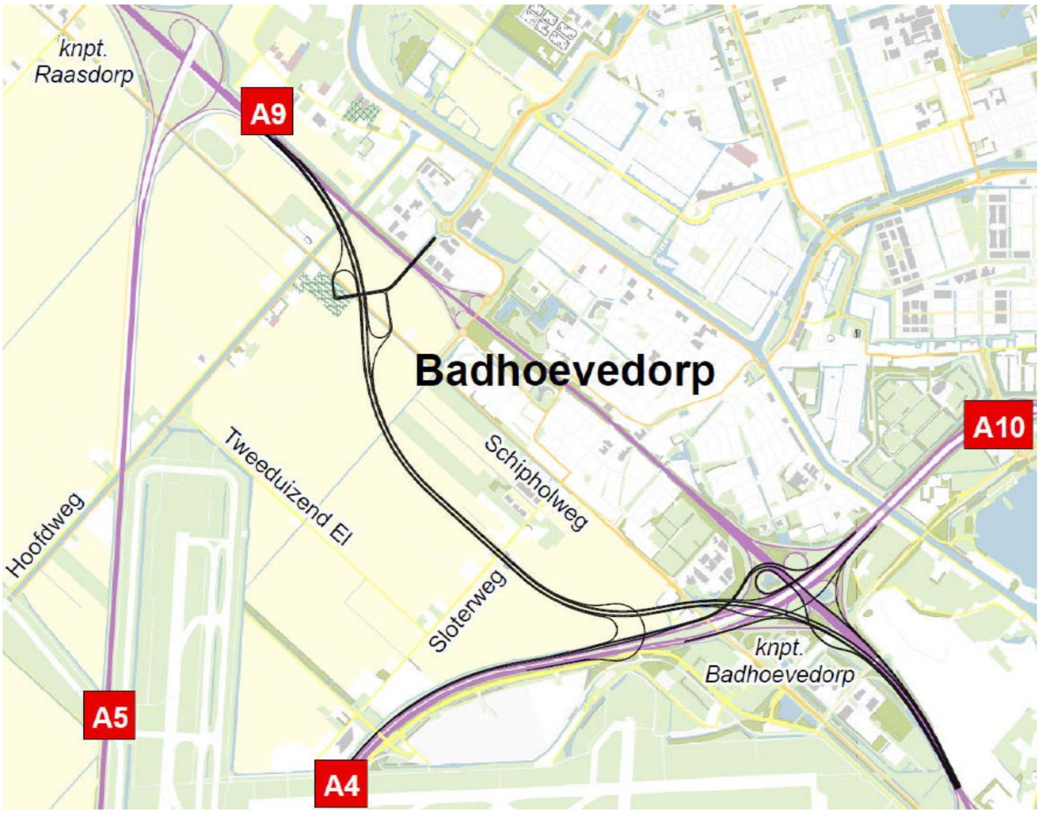


Figure 8.6: Investment 2010-2020 – Badhoevedorp

- *A1-A6-A9 Schiphol-Amsterdam-Almere*: the trajectory Schiphol-Amsterdam-Almere is upgraded. The most important changes are that the A1 between Diemen and Watergraafsmeer and the A10 east are upgraded from three to four lanes in each direction; the A1 between Diemen and Muiderberg is upgraded from 3 lanes in each direction and one contra flow lane, to 5 lanes in each direction and two contra flow lanes; the A9 is upgraded from two lanes plus one peak hour lane in each direction, to five lanes plus one peak hour lane (unbundled in the tunnel) in each direction; the A9 between Badhoevedorp and Holendrecht is upgraded from three lanes in each direction to four lanes in each direction (see Figure 8.7).

Figure 8.7: Investment 2010-2020 Schiphol-Amsterdam-Almere

- *N201 between the A4 and Amstelhoek*: the N201 is adjusted as shown in Figure 8.8.

Figure 8.8: Investment 2010-2020 N201

- *Connection IJburg*: a second connection with IJburg is made.
- Peak hour lanes are constructed on the A4/A10 South between Badhoevedorp and Amstel, on the A9 Velsen – Raasdorp, and on the A1 ’t Gooi.
- The N244 between the A7 and the N247 and the N207 between the A4 and the N205 are upgraded from one lane in each direction to two lanes in each direction.

Costs:
We used the default values for all costs, as specified in Appendix G (Table G.1). They are assumed to be the same for all links. Of course, in practice, the link costs for bridges and tunnels should be set higher.

For all the other parameters, the default settings as specified in Appendix G were used.

Model:
We chose to optimize the capacity under regular conditions and to optimize the spare capacity by using the dynamic traffic assignment model Indy. The process ends with expert suggestions on the improvement of the network. We chose to use a fixed OD-matrix, which implies that the trip distribution model is not used. This choice was made for the following reasons:

- In order to show how the method can be applied to large networks, it is not necessary to model the demand fully correctly. Of course, in practice the demand should be modelled correctly.
- We calibrated the network based on an adjusted OD-matrix and network from the RegioRegie model (originally from the regional model system NRM). We did not have information on how this static OD-matrix was constructed, which implies that we did not have access to matching data about the number of jobs and residents per zone for the year of calibration (2008) and for design year (e.g. 2030). These data are needed to construct and update an OD-matrix with the trip generation and distribution model that is used in our method. If needed, these data can be collected and the method can be applied in practice. However, additional efforts would be needed. If the demand model is used, this model has to be calibrated in such a way that it reproduces the real OD demand as well as

possible, and it would be necessary to project the changes in demand calculated by the demand model on the original matrix.
- The network design is made for a future in which it is not unthinkable that demand regulating measures are taken (such as road pricing) that can steer the demand. Assuming a fixed OD-demand might, therefore, not be very wrong after all.

The maximum number of iterations between the optimization model for regular circumstances and the dynamic traffic assignment model is 5. This loop stops when this maximum number of iterations is reached or when the model outcomes change less than 1%. The maximum number of iterations between the optimization model for irregular circumstances and the dynamic traffic assignment model is 1. By making these choices the model might not yet have converged, which would result in a suboptimal solution. For showing how the model works this is not a problem, but in practice, of course, more iterations should be made.

The population size of the genetic algorithm was set to 50, and the number of generations is 10. In practice, the population size and the number of iterations should be higher.

The costs of vulnerability are computed only for the links with a maximum speed higher than 60 km/hour. In practice, a choice could also be made to simulate incidents on all links. Incidents on local roads can influence the spare capacity that is needed on the higher level roads. However, it is more logical to prevent these effects by including spillback buffers than by adding spare capacity to the motorways.

A choice was made to let at most 15% of the traffic take an alternative route when incidents occur. This 15% was used for all four incident types, as described in section 4.3. In practice, this percentage can be varied for the different incident types.

8.4 Step 2: Functional analysis for Amsterdam and surroundings

In this step an analysis is made of how the current network performs with an increased demand of 20%. At first this is done by looking at the performance of the network under regular and irregular circumstances. Thereafter, a more detailed analysis is made by looking at the percentage of through traffic and by analysing the road structure of the motorways and regional roads.

Figure 8.9 shows the severity of the congestion if no unexpected disturbances occur in the morning peak with an increased level of demand. The darkness of the links indicates the number of minutes of congestion in the morning peak (6.00 AM -10.00 AM). It can clearly be seen that, when the demand increases by 20% compared to 2008, the network does not have enough capacity. Most of the problems occur in the north before the Coentunnel, at the A4 in the direction of Amsterdam, at the A1 in the direction of Amsterdam, and at the ring road. The level of congestion on the roads other than the motorways is not shown. However, a more detailed view on the model outcomes shows that there is insufficient capacity within the city centre as well. The congestion on the ring road A10 is partly a result of spillback effects from the secondary network and local roads, and partly a result of insufficient capacity on the A10.

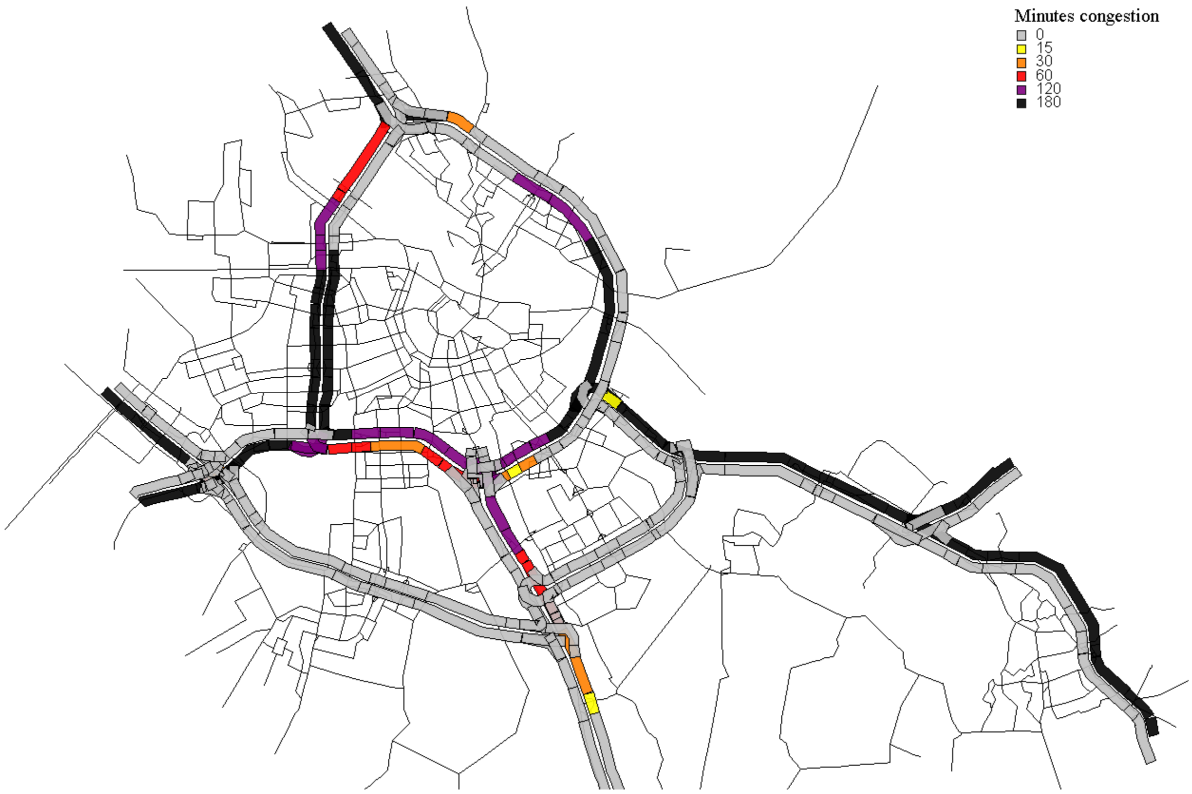


Figure 8.9: Regular congestion in the reference network with increased level of demand (+20%) in the morning peak period 6.00 – 10.00 AM of an average workday

Figure 8.10 shows the average speeds for all the trips that depart at a certain departure time step. The figure shows that the speeds drop from about 60 kilometres per hour to about 10 kilometres per hour, which indicates that the network becomes very congested at the end of the peak period (from a demand perspective). Thereafter, the network slowly recovers.

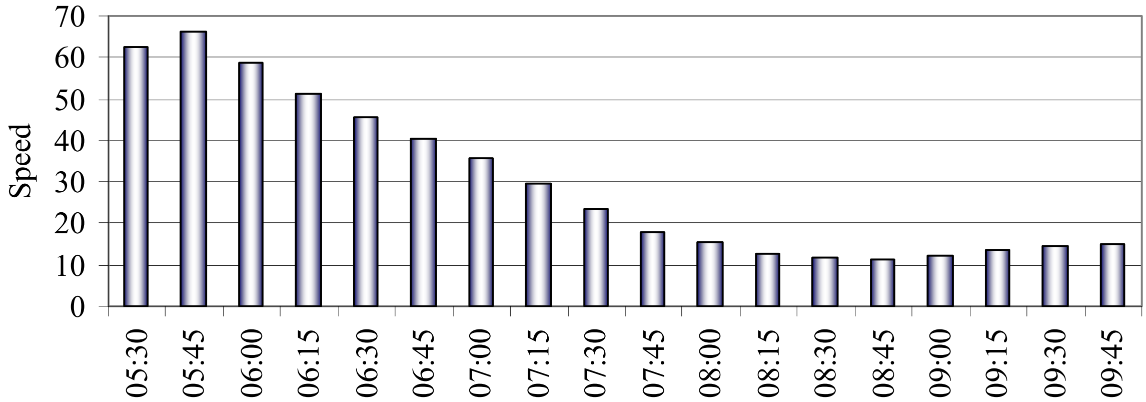


Figure 8.10: Average speed over all trips per departure time step in the reference network with increased level of demand (+20%)

Figure 8.11 shows the vulnerability of each link on which incidents are simulated in the morning peak. The darkness of the links indicates the severity of the lost travel time. When the chance of occurrence of all incidents is taken into account and the value of lost travel time is taken as €15, the cost of the vulnerability of the network in the morning peak is about

€519,000. The vulnerability is high on the A4 between the A10 and the A9, on the A10 south, and on the Coentunnel at the A10 north.

Figure 8.11: Costs of vehicle hours lost as a result of incidents on each link in the reference network with increased level of demand (+20%)

Table 8.1 summarizes the network indicators. The following are more detailed performance measures of the network for different types of traffic. In total, 12% of the trips that are included in the model have an origin and destination outside the modelled area (through traffic), 43% have an origin outside the area and a destination inside the area (incoming traffic), 30% have an origin inside the area and a destination outside the area (outgoing traffic), and 15% stay within the area (regional/internal traffic).

Table 8.1: Network indicators reference network with 20% increased demand

Indicator	Values
Total demand in T [x 1000 pcu]	518
Total travel time in T [x 1000 hours]	432
Total distance travelled in T [x 1000 km]	8253
Average speed [km/hour]	19
Total travel cost regular circumstances in T [x 1000 euro]	7933
Total travel cost irregular circumstances in T [x 1000 euro]	519

Figure 8.12 shows the share of the different traffic types on the motorways. This analysis shows that the motorways have an important function for all types of traffic, and especially for external traffic (incoming and outgoing). On parts of the A10 south and the A10 west, there is a great deal of internal traffic on the motorways.

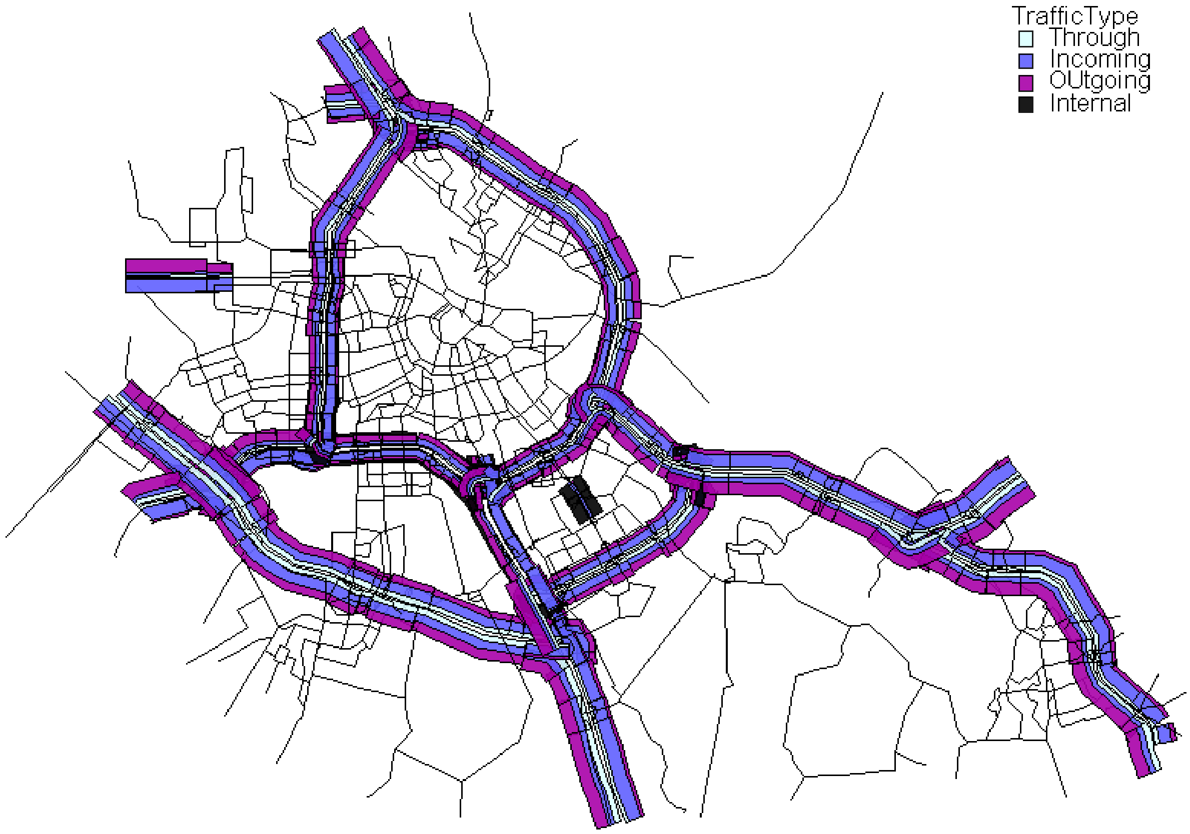


Figure 8.12: Share of traffic types in the morning peak in the reference network on roads with a free speed > 60 km/hour with increased level of demand (+20%)

The following describes the network structure for through traffic and external traffic, and for regional/internal traffic. These two types of traffic have different network requirements:

- Through traffic and external traffic travel relatively long distances and consist mainly of business and freight traffic. The most important elements for this traffic are speed, reliability, and a homogeneous traffic pattern. This implies that, for this traffic, the number of on ramps and off ramps should be minimized.
- The travel patterns and the trip purposes of regional traffic are much more varied than those of the through traffic. The distances travelled vary between 10 and 40 kilometres. Most of the commuting trips take place in this distance class, which implies that the regional traffic is concentrated in the peak period. For regional traffic, circulation is more important than high speeds. Furthermore, the network should be fine-meshed, in order to connect all the origins and destinations.

Egeter and Snelder (2010) present a detailed functional analysis of each road. Below, we summarize their main conclusions. The main conclusions of the functional analysis for the through traffic and external traffic are:

- Through traffic has a relatively small share in the total traffic, which is caused by Amsterdam's being located close to the sea, and its location of Amsterdam with respect to other cities. Incoming and outgoing traffic has a relatively large share, mainly thanks to the economically important destinations within the area, such as the airport and its surroundings, the city centre of Amsterdam city, and the North Sea Channel area.

- The A9 and the A10 are important for the through traffic and the external traffic. The importance of the A9 will increase in the future, because important road infrastructure projects related to the A9 will be completed.
- On the motorways, there is a strong mix between the through traffic, the external traffic, and the regional traffic (sometimes even local traffic on the A10). This is a disadvantage for the quality of the motorways for through traffic. For external traffic, this mixture of functions seems to be an advantage on the one hand, because in this way many origins and destinations are very well connected to the motorways. On the other hand, this is a disadvantage for the external traffic once they are on the motorways.
- The A4, the A2, and the A1 are crucial connections with neighbouring regions. There are no route alternatives for these very important motorways. In future plans, a lot of extra capacity is foreseen on these motorways. However, they will probably stay vulnerable, because extra route alternatives or improved route alternatives on the secondary network are not foreseen.

The main conclusions of the functional analysis for the regional traffic are:
Regional traffic has a varied pattern. However, there is no separate secondary network for this traffic that can operate independently of the motorway network. In fact, the motorway network is also the backbone for this traffic. For regional traffic, the motorway network is not fine-meshed enough, which results in heavily used motorway intersections and on ramps and off ramps. The capacity of the secondary network is not high enough for the volume of the regional traffic. This makes the total network vulnerable.

8.5 Step 3: Design process for Amsterdam and surroundings

This section describes how a robust design of the network of Amsterdam and its surroundings is made. The conclusions of the functional analysis are useful in understanding the model optimization outcomes, and they are an input for the expert modification.

8.5.1 Design an optimal network for the regular situation

In the first step of the design process, the capacity of the motorways is optimized for regular circumstances. The functional analysis showed that there is a lot of congestion on the motorways. The level of congestion is so high that large infrastructure investments are probably beneficial. The optimization algorithm showed that, indeed, a lot of extra capacity is needed in order to improve the travel times under regular conditions. In Figure 8.13 the extra lanes that are needed are shown, and in Figure 8.14 the total number of lanes after optimization is shown. These figures indicate that, especially on the A10 south, a lot of extra lanes are needed. On some parts, three to five extra lanes are needed (which could also be spread over parallel routes). When looking at these results, it should be kept in mind that only five iterations of the optimization model were used. If more iterations were used, the number of lanes that need to be added would definitely change. Most likely, at the places where five lanes are added, fewer lanes are needed, and at some other newly created bottlenecks, some additional lanes need to be added. Therefore, in practice, more iterations should be made. Nevertheless, as can be seen from Table 8.2 and Figure 8.15, the suggested extra lanes decrease the level of congestion and increase the average speed in the network (from 19 km/hour to 25 km/hour). However, it is likely that there are better solutions (network improvements) that would further decrease the level of congestion and have higher benefits. These better solutions could be found by doing more iterations. It should also be noted that congestion that is caused by spillback effects from the secondary network cannot be prevented as long as the capacity of the secondary network is not optimized. In practice, one could choose to optimize these capacities as well, or to adjust these capacities in the expert modification phase.

Figure 8.13: Extra lanes after optimization for regular circumstances

Figure 8.14: Total number of lanes after optimization for regular circumstances

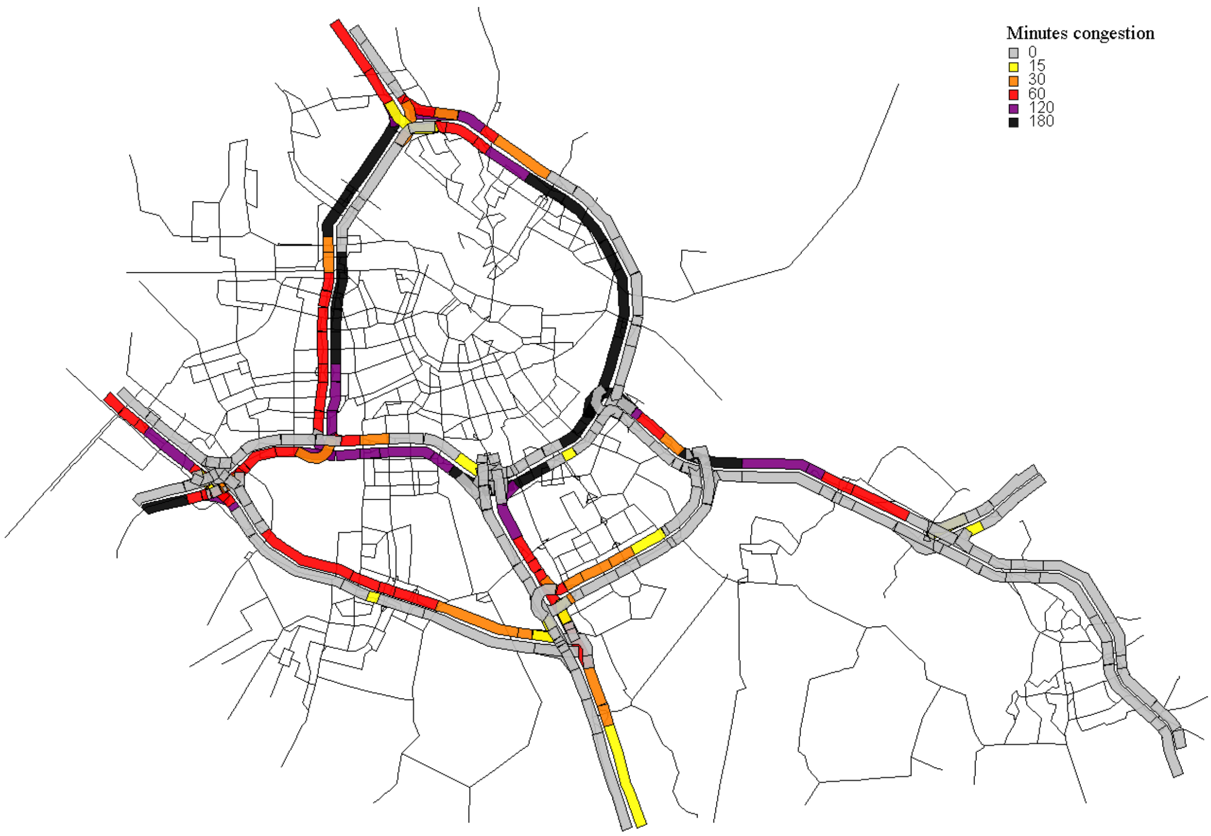


Figure 8.15: Regular congestion after optimization for regular circumstances in the morning peak period (6.00 – 10.00 AM) of an average workday

Table 8.2: Network indicators after optimization for regular circumstances

Indicator	**Value**
Investment costs converted to simulation time period *T* [x 1000 euro]	22
Total demand in *T* [x 1000 pcu]	518
Total travel time in *T* [x 1000 hours]	331
Total distance travelled in *T* [x 1000 km]	8253
Average speed [km/hour]	25
Total travel cost regular circumstances in T [x 1000 euro]	7011
Total travel cost irregular circumstances in *T* [x 1000 euro]	22

8.5.2 Design a robust network for Amsterdam

After the optimization of the capacity under regular circumstances, the additional amount of spare capacity that needs to be added to prevent vehicle loss hours that are caused by incidents is determined. Figure 8.16 and Figure 8.17 show the number of lanes that need to be added and the total number of lanes. Again, this is a sub-optimal result since, in practice, more iterations need to be carried out and the population size and the number of generations need to be increased in the genetic algorithm. The solution that is found is better than adding no lanes, which shows that adding spare capacity can be beneficial. The locations that are found can change when the algorithm is run longer.

Figure 8.16: Extra lanes after optimizing spare capacity

Figure 8.17: Total number of lanes after optimizing spare capacity

Figure 8.18 shows that the average speeds under regular circumstances per departure time improved quite a bit compared to the reference network.
Figure 8.19 shows the regular congestion. It can be concluded that adding spare capacity for irregular circumstances does not change the travel times under regular conditions, much as was to be expected. In fact, as can be seen from Table 8.3, the total benefits under regular conditions are even a bit negative (-11,000 euro).

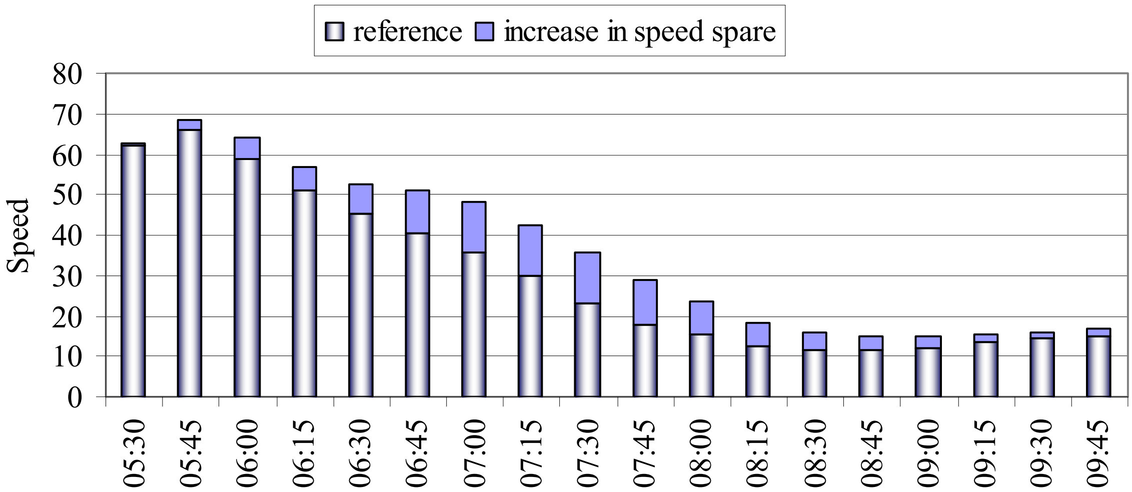


Figure 8.18: Increase of the average speed per departure time as a result of the extra (regular and spare) capacity

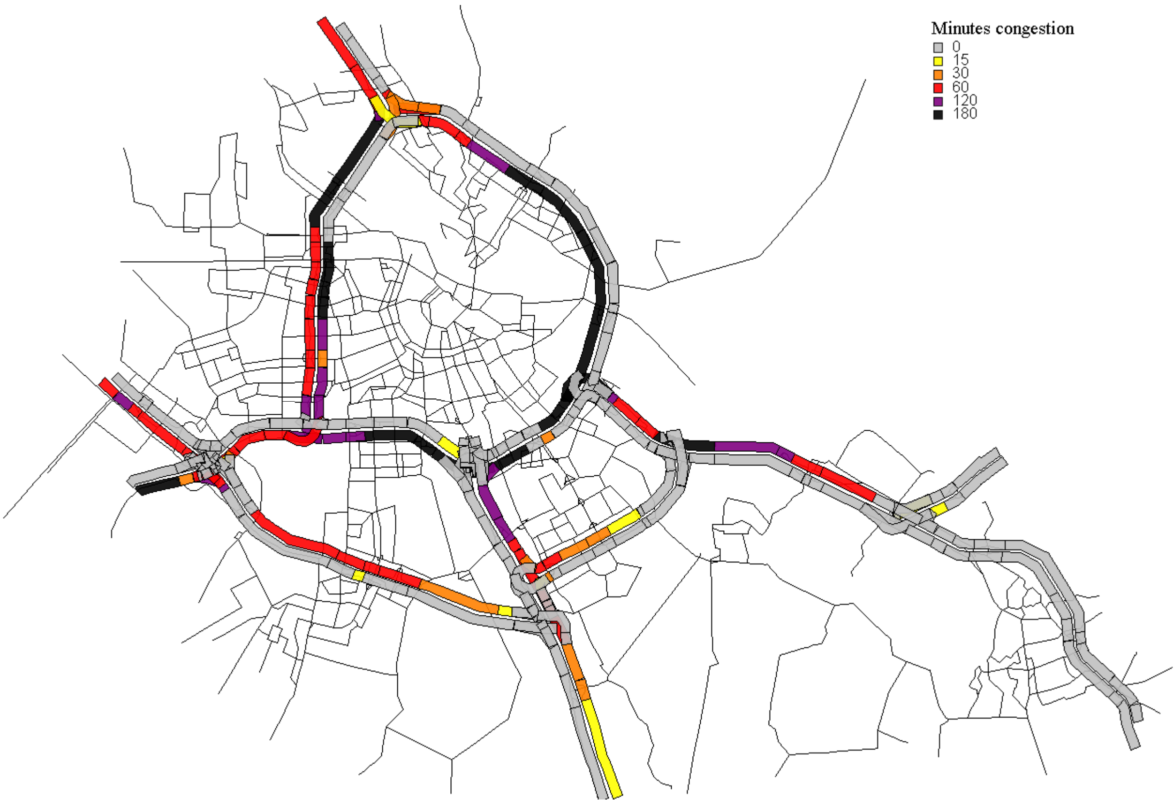


Figure 8.19: Regular congestion after optimizing spare capacity in the morning peak period (6.00 – 10.00 AM) of an average workday

The slightly negative benefits under regular conditions are compensated by the large improvements in robustness. The total costs of vehicle loss hours decrease from €519,000 in the reference network to €13,000 in the robust design. Figure 8.20 shows where these remaining vehicle loss hours as a result of incidents are made. We notice that the costs of vehicle loss hours are lower than can be expected based on the level of congestion under regular disturbances. This is explained by the fact that the incidents are simulated with a start time of 6.45 AM. This is a time at which there is hardly any congestion in the robust network. This implies that, in practice, the incidents should be simulated over a longer time period.

Figure 8.20: Costs of vehicle hours lost after optimizing spare capacity

Table 8.3: Network indicators after optimizing spare capacity

Indicator	**Values**
Investment costs converted to simulation time period *T* [x 1000 euro]	3
Total demand in *T* [x 1000 pcu]	518
Total travel time in *T* [x 1000 hours]	332
Total distance travelled in *T* [x 1000 km]	8253
Average speed [km/hour]	25
Total travel cost regular circumstances in *T* [x 1000 euro]	7020
Total travel cost benefits regular circumstances in *T* [x 1000]	-11
Total travel cost irregular circumstances in *T* [x 1000 euro]	13

8.5.3 Expert modification

In the expert modification phase, the experts have the opportunity to improve the network based on the functional analysis and the model outcomes. As is explained in the introduction of this chapter, not much attention is given to this phase in this chapter. Based on Egeter and Snelder (2010), we do indicate what kinds of changes could be made.

In Figure 8.21 a vision of a robust road network is presented for an area (Stadsregio Amsterdam) that is a bit larger than the area used so far. The most important changes that are suggested are changes in the functionality of some roads. A large part of the A9, the A1, the A2, the A4, and the A6 are pointed out as primary/main motorways. The A5 and the Westrandweg (one of the planned infrastructure projects, see Figure 8.5), the other parts of the A9, and the A10 North and East are indicated as secondary main routes. One of the main reasons for doing this is that the parts of the A10 that have an important function for regional traffic (mainly A10 south) can be changed to a motorway with a function for local traffic. The through traffic, therefore, needs to be guided around Amsterdam via other roads. Furthermore, some roads (many parts of the A9 and small parts of the A4 and the A1) are unbundled, and there are some new roads suggested in the secondary network. The number of lanes per road are not yet determined. A way of doing this is by using the model outcomes. The number of extra lanes that are determined by the model can be spread over different (new) parallel roads, such as the Westrandweg. A shift in the main road from the A10 to the A9 is another example.

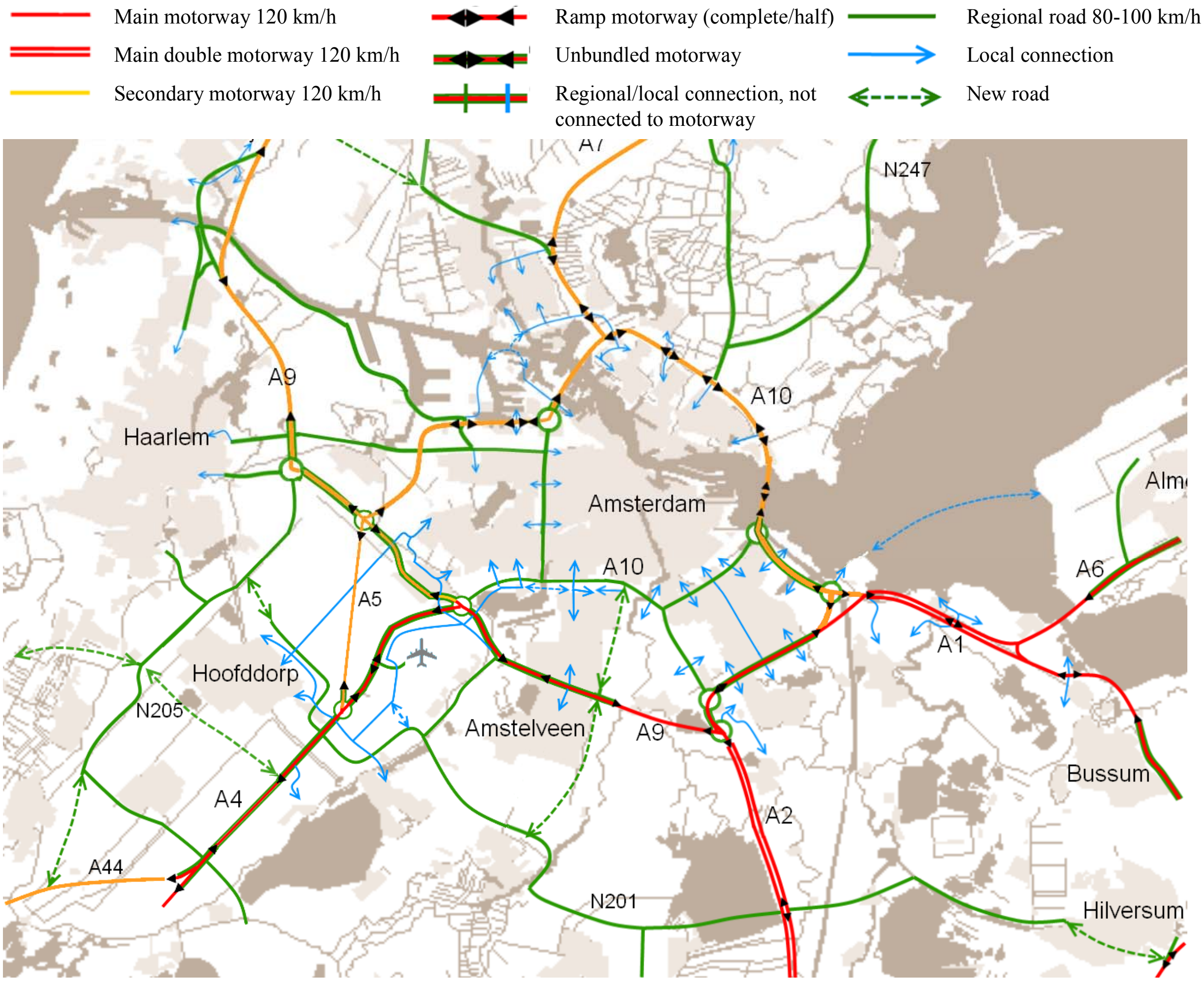


Figure 8.21: Function robust design after expert modification

After this step, a choice could be made to evaluate the robustness of the new design and to make some additional changes to the network.

Finally, when this method is applied in practice, it is advisable to do a sensitivity analysis with respect to the most important parameters and assumptions, as is done for the example network in chapter 7.

8.6 Summary and discussion: Application to the network of Amsterdam

In this chapter, the design method was applied to a large realistic network of Amsterdam and surroundings. As far as is known to the author, this is the first application of a network design method that combines models and expert knowledge in the network design phase on such a large a scale by taking robustness into account. Furthermore, it is the first application in which a dynamic traffic assignment model is used in combination with a marginal incident computation model in order to make a robust road network design. The application of the evaluation method by itself is also a large step forward, since it shows how the vulnerability of a network can be analysed by taking spillback effects and alternative routes into account. By applying the design and evaluation method to a large real-size network, many lessons were learned about how robustness analysis and robust road network design can be done in practice. This application served as an example to answer the 10^{th} research question: How can the method be applied to large scale networks?

From the application, we learned that many choices have to be made with respect to how the method is applied. The choices relate to the objective of the design, the design variables, the constraints, the future developments in demand and supply, the costs, and the model parameters.

From the application, we learned that it is especially important to select the right links for which the capacity is optimized. It is advisable to do this not only for the motorways, but also for the other roads that influence the flows and the level of congestion on the motorways. If this is not done, spillback effects cannot always be prevented in the modelling phase because, in the phase when the network is optimized for regular circumstances, spillback effects are ignored (they can be dealt with in the expert modification phase). Furthermore, it turned out that it is important to spread the incidents that happen over the complete simulation period, especially in highly congested networks. We chose to model the incidents at the beginning of the peak period. However, since the peak period lasted much longer than we expected as a result of the increased demand, the vulnerability of the network at the second half of the simulation period was underestimated. Based on these experiences, it is advisable to reconsider the choices that are made in the first step after the second step (the functional analysis) is carried out.

Furthermore, it turned out the design method can be applied to large networks. This takes a lot of time, because the modelling part takes a long time and because expert workshops have to be organized. In practice, there is a relatively large interval time between the two to four workshops that are organized with the experts, which should be enough time to carry out the modelling steps. Nevertheless, the computation times are an issue that needs to be addressed in the future. As an indication, we mention that the computation time of an Indy run with the link transmission model on the network of Amsterdam and surroundings takes about 12 hours on a normal PC. An Indy run in combination with the vertical queuing model that is used in the optimization phase of the capacity under regular conditions takes less time: about 6 hours. The analyses with the marginal incident computation model go faster, but still take some time. The analysis for the population size of 50 with 10 iterations took a bit less than a day. As a consequence, the complete computation time of the analyses that were carried out in this chapter took about a week. If the number of iterations is increased and the population size that is used in the genetic algorithm is increased, the computation times will increase further. Temporal results are stored, which enables us to stop after each modelling step and continue afterwards. This is an advantage, because initial choices can be reconsidered. Nevertheless, this application showed that we have to be selective when choosing the model parameters. There are many possibilities for reducing the computation time. Adding a stop criterion to the Indy runs that stops the model when convergence is reached instead of waiting until the maximum number of iterations is reached is an important one. The number of iterations within Indy was set high enough to be sure that convergence was reached. However, when the optimization process gets further, the level of congestion decreases, and convergence is reached within many fewer iterations than specified. A second promising improvement is using parallel computing within the genetic algorithm in such a way that multiple networks/chromosomes can be evaluated at the same time. Thirdly, at the moment and especially in the future, there are probably much faster (super)computers that can be used. Finally, it is likely that a lot of computation time can be saved by making the algorithm run more efficiently, by for instance reducing the number of times that files have to be read and results written to files.

The application of the modelling steps of the design method showed that large improvements in network performance can be obtained with a positive benefit-cost balance. The table below summarizes these results.

Table 8.4: Summary network indicators for Amsterdam and surroundings

Indicator	**Initial network**	**After reg. optimization**	**After sparecap optimization**
Investment costs converted to simulation time period *T* [x 1000 euro]	-	22	3
Total demand in *T* [x 1000 pcu]	518	518	518
Total travel time in *T* [x 1000 hours]	432	331	332
Total distance travelled in *T* [x 1000 km]	8253	8253	8253
Average speed [km/hour]	19	25	25
Total travel cost regular circumstances in *T* [x 1000 euro]	7933	7011	7020
Total travel cost irregular circumstances in *T* [x 1000 euro]	519	(not computed)	13

These results were obtained by using a small number of iterations, which results in a sub-optimal solution. If the model is run longer, it is very likely that better solutions will be found.

Finally, in the last section an impression was given on how expert judgements can improve the robustness of the design by applying robustness measures in addition to adding spare capacity to the vulnerable roads.

9 Main findings, implications and recommendations

9.1 Introduction

Section 9.2 summarizes the main findings with respect to robust road network design, and the methods that are needed for accomplishing it. In this thesis, the Netherlands was taken as an example, which suggests that the main findings presented in section 9.2 hold only for the Netherlands, or perhaps only for the example areas that were considered. However, although not proven, we have reason to believe that these findings can be generalized to other similar areas all over the world. In addtion, material presented in the thesis has implications for the different stakeholders that are involved in road network design. These implications are summarized in section 9.3. Finally, the work that has been done raises a lot of questions. Therefore, in section 9.4 some recommendations for future research in this area are given.

9.2 Main findings

This section contains two subsections ─ the main findings with respect to robust road network design and the solution algorithm. Each subsection includes several short statements about the findings, followed by the arguments supporting the statement.

9.2.1 Main findings with respect to robust road network design

Road networks are vulnerable – the costs of vulnerability can add up to several billion euros in the Netherlands if no measures are taken.

A data analysis of a single incident on an off ramp showed that small incidents can block a large part of a motorway network for many hours. Many individual travellers experience an unexpected delay of more than one hour as a result. Incidents like this happen quite often. In the future, the effects of those incidents will likely become larger, because the spare capacity in both space and time will be reduced as a result of growth in demand in urbanized areas, if no measures are taken. Our simulations indicated that the costs of vulnerability in 2008 were ranged between €275 million and €1.2 billion. If no measures are taken, by 2030 these costs might increase to between €900 million and €4.1 billion. These results suggest that it would be beneficial to implement robustness measures to reduce these costs.

In principle, road networks can be made robust against all disturbances. Network managers and policy makers need to specify against which disturbances their network is to be made robust, and at what cost.
In this thesis a common framework was presented that gives policy makers and transportation analysts the possibility to discuss issues that are related to road network robustness and vulnerability. Robustness is defined as follows:

> Robustness is the extent to which, under pre-specified circumstances, a network is able to maintain the function for which it was originally designed.

Vulnerability is the opposite of robustness. A network that is vulnerable is not robust, and vice versa.

In principle, a network can be made robust against all kinds of disturbances ranging from predictable to non-predictable disturbances, from regular to irregular disturbances, from disturbances with a temporarily small or large impact/effect to disturbances that have a permanent impact, and from disturbances that have a network-wide effect to disturbances that have a local effect. Network managers and/or policy makers have to decide how robust the road network is to be made against which disturbances. A decision support model could help them make these decisions. In any case, if robustness issues are discussed, it is advisable to clearly indicate against which disturbances a network is to be made robust. The robustness of a road network focuses more on effects than on probabilities. Nevertheless, for network design, probabilities are an important factor for investment decisions. Therefore, this thesis focused on the class of non-predictable regularly occurring (relatively high probability) disturbances that we called *incidents* (including some incidents, like road closures, that belong to the class of non-regular non-predictable disturbances).

A robust road network is a network in which preventive measures are taken and that has redundancy, is compartmentalized, is resilient, and is flexible.
It can be concluded that a network becomes more robust if there are alternative routes available and there is enough spare capacity on all the routes (<u>redundancy</u>). Having this redundancy is not enough; the transported elements must also be rerouted to these alternative routes as fast as possible, which requires some <u>flexibility</u> in the network and some traffic and incident management strategies (<u>resilience</u>). Furthermore, in a robust network, spillback effects must be reduced to a minimum level (<u>compartmentalization</u>). Finally, <u>preventive</u> measures make a network more robust as well. In the next chapter these elements are worked out in more detail.

The conclusion that prevention, redundancy, compartmentalization, flexibility and resilience are elements that make a network robust is supported by the fact that these elements are also found in other transport and non-transport network. The most important lessons learned from the comparison between the road network and other networks are:

- Compared to other non-transport networks, the road network is unique, since passengers are transported and cars are controlled by humans. Humans are not interchangeable, whereas blood cells are. Therefore, <u>human behaviour</u> has to be considered in any robustness analysis of the road network.

- Building roads is more expensive than building infrastructure for most of the other non-transport networks. Not only is it more expensive, but it also has a higher impact on the environment, which makes constructing road networks extra complex. This explains why it is more difficult to make road networks redundant than other non-transport networks. In general, it can be concluded that the competition for space in land use is a main factor that makes it complex to make transport network robust.
- The road network is more susceptible to disturbances than most other networks, because it is a less controlled network. Therefore, in other networks, such as the railway network and the nervous system, more structure related preventive measures can be taken (and are taken).
- The road network is less flexible and less resilient than, for instance, the cardiovascular system and the Internet. This is mainly because blood and information packages can more easily be rerouted and because there is an instantaneous response to disturbances. The diameter of the blood vessels of the cardiovascular system is flexible enough to cope with sudden changes in the volumes transported. Compared to the rail network, the road network offers more flexibility, because it is easier for cars and, therewith for passengers, to switch to other routes (if they are available).
- Spillback effects are more likely to occur in road networks than in other networks, because they are used more heavily than most of the other networks and because rerouting is more difficult than in most of the other networks. This is probably a result of a lack of alternative routes in some parts of the network and limited access points to alternative routes that are available, a lack of accurate information about the available alternative routes, and the fact that not all drivers are willing to take alternatives routes.

Robustness and reliability are related concepts that need to be addressed together

In the end, the traveller wants to have reliable travel times and does not want to be confronted with unexpectedly high travel times as a result of incidents. We showed that robustness and reliability are two related concepts that differ in focus. The focus of robustness is on the network structure and the effects of disturbances with an unexpected effect at a specific moment in time, whereas the focus of reliability is on the traveller and the probability of expected disturbances measured over a longer period. In a robust road network, disturbances lead, on average, to smaller variations in travel times, and, therewith, to more reliable travel times. However, this is not a one-on-one relation. Making the network more robust, does not have to result in more reliable travel times for all travellers. If robustness measures are taken, the reliability of the network should be evaluated by including all kinds of disturbances. However, it is necessary to give special attention to the disturbances with large effects, such as incidents, because these effects are easily underestimated if traditional methods for reliability analysis are used.

The benefits of robustness should be included in cost-benefit analyses

When robustness measures are taken, robustness/reliability benefits should be considered, because those measures are intended to improve the robustness of the network and, therewith, the reliability of travel times. It is obvious that including solely the travel time benefits under regular conditions, as is usually done, is not enough. Different examples showed that, in specific cases, the robustness/reliability benefits can be significant: possibly of equal size to the travel time benefits. This implies that the benefit-cost balance of robustness measures can be much more positive than when the robustness/reliability benefits are not considered. This should make it easier to get the robustness measures implemented.

Many measures can be taken in order to improve the robustness of a road network. However, the benefits of those measures depend on local circumstances. The benefits cannot be approximated by rules of thumb or regression analyses.
A list of measures that are likely to improve robustness were presented in this thesis. These measures were subdivided into measures that can be taken on the travel market, the transport market, and the traffic market. Furthermore, it was explained to which elements of robustness they are related, and for which disturbances they are relevant.

For a selection of measures that match the best with the focus of this thesis, the possible effects of the measures were analysed by using examples. It was shown that having a single route alternative could reduce delays by 3.1% - 77.1%, depending on the level of information that is given. Additional extra paths result in only small improvements. Of course, the results also depend on the quality (spare capacity and travel time) of the alternative routes.

With respect to unbundling, we concluded that unbundling can have a positive effect on robustness. However, the way in which it is done and the locations where it is done need to be chosen carefully, because unbundling can have both advantages and disadvantages for the robustness of a network.

Finally, we showed that adding buffers and spare capacity to the network can have large positive effects on travel times. Buffers have the potential to reduce the average travel time by 12% in the morning peak in the network of The Hague (section 5.5). In other networks and in other time periods, this impact can, of course, be different. The more congestion there is, the higher the benefits are likely to be. This conclusions is based on a regular situation without incidents. However, it is straightforward to show that similar effects occur in the case with incidents. This implies that buffers can make a network more robust. Finally, an example taken from Immers et al. (2004d) showed that the optimal spare capacity of an average motorway link is about 800 pcu/hour. This optimization was done without considering spillback effects.

The robustness of many parts of the road network are likely to be able to be improved with a positive benefit-cost balance.
An application of the design method to a small test network (the network of The Hague – Rotterdam) and to the network of Amsterdam showed that the robustness of a network can be improved significantly if a set of robustness measures is taken in an integrated way. In the network of The Hague – Rotterdam, the improvement was 30%, and in the example for Amsterdam, the costs of vehicle loss hours as a result of incidents was reduced from €519,000 to €13,000(for the simulation time period). The measures not only reduce the vulnerability (or improve the robustness) of the networks, but the benefits also outweigh the costs. Of course, the gains in robustness in these examples cannot be generalized to all networks, because the benefits depend on the initial robustness, the traffic volumes, the network structure, and other circumstances. Nevertheless, these examples show that it is worthwhile to at least investigate whether or not the robustness of certain parts of a network can be improved by taking robustness measures.

A balanced choice needs to be made among improving the travel times under regular conditions (and therewith generating mobility), improving the robustness of a network, and making maximum use of the existing network.
Robustness measures can lead to an increased demand. If the demand increases, the intended increase in spare capacity might be cancelled out. This implies that the robustness of the

network has not improved by the amount that was to be expected if demand effects had not occurred. Both improved robustness and an increase demand can lead to welfare effects. Therefore, they need to be balanced.

9.2.2 Main findings with respect to the solution algorithm

Evaluating the robustness of a network in network design requires an accurate and fast evaluation method.

In this thesis, many indicators were presented that refer to one or more elements of robustness. We chose to use the indicator ‘vehicle loss hours caused by incidents’, mainly because this indicator describes the different elements of robustness best. Depending on the way it is computed, this indicator can consider spillback effects (compartmentalization), route alternatives (redundancy), and resilience. If flexibility is included in the network, these flexible infrastructures can be considered (depending on the type of model that is chosen). Furthermore, the indicator ‘vehicle loss hours caused by incidents’ is relatively easy to explain and the can be valued by using a value of time. It was also explained that monitoring the indicator (based on data) is complex, but possible to a certain extent. The most complex part of this indicator is to find a model by which this indicator can be computed in an acceptable computation time.

The dilemma in the model choice for evaluating robustness in network design is that a choice has to be made between accuracy (ideally using a dynamic traffic assignment model with detailed congestion modelling, including spillback effects, with multiple types of route choice behaviour during incidents, and with accurate intersection modelling) and computation time. The most accurate models generally take the longest computation time. Using a rule of thumb, for instance, takes hardly any computation time, but is not accurate. For some applications, a rule of thumb can be good enough to get a quick impression of the robustness of a network. However, to make a well balanced decision about robustness measures in network design, it would be better to look for a method/model that deals with the above mentioned requirements in the best possible way and is fast, because many evaluations have to be carried out in network design.

We chose to use the dynamic traffic assignment model Indy to compute an equilibrium under regular circumstances, because this model has an accurate network loading model that models spillback effects according to the simplified kinematic wave theory of Newell. Furthermore, the model can compute an equilibrium route choice and can deal with fixed route choice. En-route route choice is not possible. However, this is not a problem, since the model Indy is used only for a basic run in a situation without disturbances, for which an equilibrium assignment is the most appropriate. Indy is a dynamic model, which makes the modelling of time dynamics possible. And, since Indy is able to work with the marginal incident computation model (MIC) (Corthout et al., 2009), it is able to get an estimate of the impact of incidents very quickly. Four incident types were defined, for which the effects are computed for all links. An approximation algorithm was added to the MIC-module to deal with the use of alternative routes. We showed the quality of the method by making a comparison with loop detector data for five incidents. Finally, we showed how robustness/reliability benefits of robustness measures can be computed with this evaluation method, and how they can be included in a cost-benefit analysis. This method focuses on computing the robustness/reliability benefits for four different incident types. Other methods, such as the Monte Carlo simulation based model SMARA, are available for getting an indication of the benefits related to other disturbances. Both methods can be combined in a straightforward way. However, that was not done in this thesis.

Combining expert knowledge and advanced modelling techniques makes it possible to design a large scale robust road network with many design variables (= robustness measures) that is likely to be supported by all the stakeholders that were involved in the design process.

- This makes it possible to enrich the model input and outcomes with knowledge from the stakeholders that can never be included in a model.
- This makes it possible to include all kinds of robustness measures in the optimization process. It is very difficult to include this many design variables in a full model approach.
- This creates support for the network design during the design phase and might prevent a lot of discussion afterwards.

The fact that extensive use of models is made as well is an advantage, since:

- This gives useful input for the part in which the experts are involved, because the evaluation part of the models shows how vulnerable the existing network is and which locations of the network are vulnerable. It does the same for new designs. The optimization part of the models gives suggestions about how much spare capacity is to be added on which locations.
- This gives quantitative support to the decisions that are made.

The proposed solution algorithm is a compromise between accuracy and computation time in order to be applicable to large-scale networks. Thanks to some simplifications, it is possible to design a good robust road network with a positive cost-benefit balance. However, this network is not necessarily optimal, given the objective function.

The solution algorithm was designed in such a way that it can be applied to large-scale networks. In order to make the algorithm applicable to large networks, some simplifications had to be made in the optimization algorithm:

- The capacity under regular conditions and irregular conditions is optimized separately. This might result in adding too few lanes, because there are cases in which the travel time benefits under regular conditions and the reliability benefits are individually not high enough to justify an extra lane, whereas the combined benefits might be sufficient to justify an extra lane.
- The regular capacity is optimized with a link based heuristic. During this optimization process the network is only evaluated by means of dynamic traffic assignment with vertical queues in between the different iterations.
- The spare capacity is optimized with a network based genetic algorithm. During this optimization process the network is only evaluated by means of dynamic traffic assignment with a link transmission model in between the different iterations.
- The marginal incident computation model combined with an approximation algorithm for alternative routes is a simplified model for analysing the effects of incidents. For instance, it does not consider downstream effects and it also does not consider effects on alternative routes.

An extensive test of the model on a small network shows that, despite the first three simplifications, the model is able to find the global optimum for this network. Thanks to the simplifications, the number of times that a full model run with a dynamic traffic assignment needs to be done stays small, which makes an application of the model to large networks possible.

Finally, the method was made in such a way that the modules can be used independently of each other. Just showing the vulnerable links might, for instance, be very useful to policy makers by itself.

9.3 Implications for stakeholders

The findings that were presented in the previous section have a great deal of implications for policy makers, network managers, road authorities, road users, and other stakeholders. Most of the implications could have already been read between the lines of the previous section. For reasons of clarity, however, they are summarized below.

For policy makers, we have justified the inclusion of robustness analysis and robustness measures in policy plans. We showed that they should consider robustness in addition to measures that, for instance, aim to reduce travel times under regular conditions, improve reliability, and minimize negative external effects. We also explained that different measures can have contradictory effects on the different objectives. Making maximum usage of the road network under regular conditions leaves no spare capacity for situations when disturbances occur and, therefore, is not robust. Road pricing might create more spare capacity in the peak period, but might also make the peak periods longer, which reduces the spare capacity over time. This implies that, even if the structure of the road network is made more robust, strategies are still needed that maintain a certain amount of spare capacity in the network, which can be used when disturbances occur.

In order to use that spare capacity when disturbances occur, accurate information should be given to the right persons at the right time. This implies that appropriate dynamic traffic management strategies should be implemented in order to make maximum use of the more robust road network in case of disturbances. This is mainly a role for network managers and the industries that develop intelligent transport systems, such as navigation systems.

Many elements of this thesis have already found their way into practice in the Netherlands. For example, a vulnerability analysis is now being made of networks being planned for the future, and the robustness/reliability benefits of different future networks are also being analysed. The first elements of the design method that combines expert knowledge with models was used for developing a vision of a robust road network for the area The Hague – Rotterdam in cooperation with the ANWB. This vision includes measures that can be taken to improve the robustness of a road network and shows how this can be done. The complete design method that is presented in this thesis can be used to do the same in a more advanced and more structured way. This would enable policy makers to get a clear vision on how to improve the robustness of their road networks.

For traffic analysts and researchers, this thesis points the way to a lot of future research that needs to be carried out. The following sections give some recommendations for such future research. Furthermore, traffic analysts have a role in making sure that robustness analysis becomes a standard element in network evaluation, and that robustness measures are considered in network designs for the future.

This thesis might give road users new insights into why they experience unexpected delays when disturbances occur, and how difficult it is to prevent such delays from happening. It might also give them hope that these kind of delays will happen less often, if future transport policies include robustness measures. If this happens, they might be faced with some new situations. For example, situations in which they wait a few minutes in a buffer, or maybe even at home, but then continue their trips with less congestion might occur more and more often. New and better route alternatives may be implemented, and different route guidance strategies may be applied, which might affect their preferred behaviour in case of incidents. Even trips under regular conditions might be affected, because short distance trips may become less dependent on the motorways.

Other stakeholders, such as interest groups, should form their opinions about robustness (if not yet done) and, if they are in favour of robustness, help where possible in making road networks more robust.

9.4 Recommendations for future research

The work that has been done on designing robust road networks has answered the ten research questions that were presented in the introduction:

1. What is the importance of robustness for the road network?
2. How is robustness defined?
3. Against which disturbances should the network be made robust?
4. Which elements determine the robustness of a road network?
5. What is the relationship between robustness, travel times, and travel time reliability?
6. Which indicators can be used to measure robustness?
7. How can robustness be evaluated?
8. Which measures can be taken to improve the robustness of a road network, and what are their effects?
9. How can robustness be integrated into a network design method?
10. How can the method be applied to large-scale networks?

The answers to these questions were given as completely as possible. However, it turned out that additional knowledge still needs to be obtained on certain aspects. Furthermore, it turned out that certain research advances still need to be made. Therefore, several recommendations for future research are described below.

Behavioural responses to disturbances: in this thesis, we showed that, for the robustness of a road network, it is important that alternative routes are available. In order to accurately forecast the effects of disturbances, and therewith evaluate the robustness of a network, it has to be known how many people and which people choose which alternative route under which conditions. Some work on this has already been done, as described by Knoop (2009). However, this is just the beginning. In the future, much more traffic data driven research, or stated and revealed preference research, is needed to show the differences in route choice behaviour for different trip purposes, different trip distances, different origin and destination regions, different characteristics of people, and different information strategies. Furthermore, it is not only important to know which alternative routes are chosen, but it is also important to know if and how people adjust their departure time, mode choice, destination choice, and maybe even in the long term their location choice, as a response to disturbances. Once more is known about these behavioural responses, the evaluation models can be improved in such a way that they take these responses into account in a better way. The approximation algorithm for alternative route choice during incidents is probably the first model that needs to be improved.

Robustness for different user groups: related to the previous point is the fact that a distinction needs to be made between different user groups not only in the evaluation, but also in the design. In this thesis, we distinguished to a certain extent between long distance and short distance trips, and between through traffic, external traffic, and regional/local traffic. However, these distinctions were made only in the non-modelling part of the design phase. It is conceivable that a robust road network for freight traffic differs from a robust road network for passenger traffic. In other words, a combined design (as is done in this thesis) might result in a different network than when both user groups are dealt with separately and then combined. The same is true for short distance and long distance traffic. Of course, it could also be the case that the resulting combined network design is the same as when the different user groups are not distinguished. However, this is something that needs to be investigated.

Information: the previous two points suggest that improved information provision during disturbances might be useful. It is important to give timely and accurate information to the right people. For instance, someone who makes a long distance trip might have a good route alternative for his/her trip. However, if the point where the choice between the two routes is to be made is close to the origin, the alternative route would be unreachable, and therefore useless, if the driver is not informed about a disturbance and the possible effects of the disturbance before he/she passes the point where the routes split. One complicating point is that it is difficult to give good information because, if everybody takes the same alternative route, which is not equipped for that, the delays might get worse rather than better. Therefore, it is important to investigate which information strategies can best be applied in which situations in order to make optimal use of a robust road network. This also requires an improvement of the models that are used to forecast the traffic flows, since the response of drivers to information needs to be better included in the models.

Relation between a robust network structure and 'robust traffic management': besides information provision, other traffic management strategies should be developed in order to use the robust road network in the best possible way. In case of disturbances, drivers should be guided to the alternative routes, and the flows on the alternative routes should be managed in the best possible way. This implies, for instance, that intersections should be designed in such a way that they can deal with a sudden increase in traffic volumes in case of disturbances. Flexible signal settings can help with this. Research should be done to develop 'robust traffic management strategies' that improve the flows over the robust road network structure in case of disturbances.

Value of time for unexpected delays: a lot of work is being done in order to find out how people value unreliability. However, the focus of most of the research projects is on unreliability as a result of many regularly occurring disturbances. It is unknown how people value unexpectedly large delays. Therefore, more work should be done in order to obtain more knowledge about the value of time in these circumstances. This information is needed in order to balance the investment costs for reliability improvements against the resulting robustness/reliability gains.

Cost-benefit analysis: in this thesis, we emphasized that the robustness/reliability benefits should be included in the cost-benefit analysis of infrastructure (and other) measures. In section 4.4, a method was shown by which this can be done. However, it requires more efforts to investigate if this is a method that can be standardized in practice or needs to be adjusted, improved, or simplified to become common practice. In our opinion, a disturbance based approach to cost-benefit analysis is the right approach, because this enables us to distinguish between different values of time for different kinds of disturbances, and avoids the problem of trying to distinguish between reliability and robustness, as is currently done in the Netherlands. However, this requires that different values of time for different disturbances or effects of disturbances are available, that a distinction in the data can be made among different disturbances, and that the effects of different disturbances can be measured. Finally, it is important that in the cost-benefit analysis an integral decision can be made between robustness measures and other measures, such as road pricing and dynamic traffic management measures. This requires still a lot of research.

Data availability: in order to get more insights into robustness and to calibrate the evaluation and design models better, it is important to have data about the effects of disturbances for different user groups and data related to different types of disturbances. In this thesis, we mainly used aggregated data for the motorways. We managed to link these data to incident data, to a certain extent. However, much better linked data are needed about traffic counts and speeds, OD-travel patterns, the network structure, and all kinds of disturbances (like the weather conditions, roadwork, incidents, holiday traffic, and events). Furthermore, these data should be available, not only for the motorways, but for whole trips (from door to door), or at least on the secondary network.

Efficient modelling of other disturbances: the focus of this thesis was on incidents. However, many other disturbances can occur. There are models available, like SMARA, that can deal with those disturbances, but not yet in a dynamic setting. Therefore, the available models should be improved or combined, or new, fast models should be developed that are able to forecast the effects of disturbances other than incidents. They should be fast, in order to be able to be used in network design. Tools like these can help in determining whether networks that are robust against incidents differ from networks that are robust against other disturbances, and, if so, how they differ.

Efficiency of the algorithm: the modelling part of the design method consists of many different sub-modules. Although the method was developed in such a way that the computation time stays within acceptable bounds, the computation time can still be greatly reduced, for instance by:

- Limiting the number of times that input files are read and output files written. The current model makes use of several sub-models, which communicate through input and output files. If the codes of those modules are more integrated, their communication can be improved by making more use of the memory. Since, the reading and writing of the input and output files takes a large part of the total computation time, great improvements in computation time can be expected.
- The computation time can be further reduced by using parallel computing techniques. This can be done for instance, for the many fitness evaluations in the genetic algorithm that is used for optimizing the spare capacity.
- Finally, the model Indy runs with a fixed number of iterations. This number of iterations needs to be set high enough to be sure that convergence is reached. As a consequence, the model practically always converges many iterations before the pre-set number of iterations is reached. This implies that the computation time can be reduced by implementing a duality gap and/or path shift based stop criterion.

Prioritization of measures: applying the design method produces a vision of a robust road network. However, the measures that are included in the design cannot all be implemented at the same time. Therefore, a method should be developed by which the measures can be prioritized. This will help in actually producing a robust road network quickly.

Learning from the intermediate results: the design method consists of two optimization models. In our analysis we mainly looked at the final outcomes. However, the intermediate results might provide very useful insights for network design as well. The optimization algorithm for regular circumstances is a kind of 'bottleneck solving' approach. Therefore, the intermediate results might show which bottlenecks occur in which order in the future when a bottleneck solving approach is used in network design. The genetic algorithm produces a lot of near optimal results. An analysis of these solutions could also provide insights into the usefulness of individual measures. The best solutions that are produced in different iterations might give useful information about how networks can be developed over time. The intermediate results might also help in deciding how measures that are taken in the future need to be prioritized. Or in other words, the intermediate results can help in deciding which measures should be taken first. Therefore, it is worthwhile to investigate if the intermediate results can really give additional useful insights for network design.

Broadening of scope: the focus of this thesis was on making the structure of road networks more robust against incidents. However, this immediately raises the question of how other infrastructure networks (like the railway network) can be made robust against incidents, how the transport system as a whole can be made more robust against incidents, and how those networks can be made robust against other disturbances as well. Furthermore, it raises the question how non-infrastructure related measures can improve the robustness of infrastructure networks. Finally, the interaction between spatial planning and land use can be investigated, and the external effects (e.g., emissions, noise, and safety) of a robust road network need to be addressed in more detail.

References

4Cast (2005) *Analyse HWN/OWN verkeer met het LMS*, Rapport voor de Adviesdient Verkeer en Vervoer, Leiden, The Netherlands.

Abdulaal, M., L.J. LeBlanc (1979) Continuous Equilibrium Network Design Models, *Transportation Research Part B*, 13, 1, pp. 19-32.

Albert, R., H. Jeong, A.L. Barabási (2000) Error and attack tolerance of complex networks, in: *Nature,* 406, pp. 378-382.

Asakura, Y. (1999) Evaluation Of Network Reliability Using Stochastic User Equilibrium, in: *Journal of Advanced Transportation*, 33, 2, pp. 147-158.

Astarita, V., K. Er-Rafia, M. Florian, M. Mahut, S. Velan (2001) Comparison of Three Methods for Dynamic Network Loading. In: *Transportation Research Record: Journal of the Transportation Research Board*, 1771, pp. 179-190.

Bell, M.G.H., Iida, Y. (1997) *Transportation Network Analysis*, Wiley, Chichester, United Kingdom.

Bell, M.G.H (2000) A game theory approach to measure the performance reliability of transport networks, in: *Transportation Research Part B*, 34, 6, pp. 533-545.

Berdica, K. (2002) An introduction to road vulnerability: what has been done, is done and should be done, in: *Transport Policy*, 9, pp. 117-127.

Berdica K, L.G. Mattsson (2007) Vulnerability – A model-based case study of the road network in Stockholm. In: Murray, A.T. and T.H. Grubesic (eds.), *Critical Infrastructure: reliability and vulnerability*. Springer-Verlag, pp. 81-106.

Besseling, P., W. Groot, A. Verrips (2004) *Economische toets op de Nota Mobiliteit (Economic analysis of the White Paper on Mobility)*, CPB, Document 65, The Hague, The Netherlands.

Billington, R., R. Allan (1992) *Reliability evaluation of engineering systems: Concepts & techniques*, Plenum Press Publish, New York, United States of America.

Bliemer, M.C.J., P.H.L. Bovy (2003) Quasi-Variational Inequality Formulation of the Multiclass Dynamic Traffic Assignment Problem, in: *Transportation Research Part B*, 37, pp. 501-519.

Bliemer, M.J.C. (2005) *INDY 2.0 Model Specifications,* Delft University of Technology, Delft, the Netherlands.

Bliemer, M.J.C. (2007) Dynamic Queuing and Spillback in an Analytical Multiclass Dynamic Network Loading Model, in: *Transportation Research Record: Journal of the Transportation Research Board*, 2029, pp. 14-21.

Boccaletti, S., V. Latorab, Y. Morenod, M. Chavezf, D.U. Hwanga (2006) Complex networks: Structure and dynamics, *in: Physics Reports 424*, pp. 175 – 308.

Bogers, E.A.I., F. Viti, S.P. Hoogendoorn (2005) Joint Modelling of Advanced Travel Information Service, Habit, and Learning Impacts on Route Choice by Laboratory Simulator Experiments, in: *Transportation Research Record: Journal of the Transportation Research Board*, 1926, pp. 189-197.

Bollobás, B. (1985) *Random Graphs*, Academic Press, London, United Kingdom.

Bollobás, B. (1998) *Modern Graph Theory, Graduate Texts in Mathematics*, Springer, New York, United States of America.

Boog, J.J., M.M.M. Linssen, S.C. Oudmaijer, J.F. Suyver (2007) Nulmeting grondroerdersregeling, Graafschades aan kabels en leidingen en hun oorzaken voorafgaand aan de inwerkingtreding van de Wet, EIM, Zoetermeer, The Netherlands.

Bornholdt, S., H. G. Schuster (2003) *Handbook of Graphs and Networks, From the Genome to the Internet*, WILEY-VCH, Weinheim, Germany.

Boyce, D.E., B.N. Janson (1980) A discrete transportation network design problem with combined trip distribution and assignment, in: *Transportation Research Part B*, 14, pp. 147-154.

Bovy, P.H.L., M.C.J. Bliemer, R. van Nes (2006) *Course CT4801 Transportation Modelling*, Delft university of technology, Faculty of Civil Engineering and Geosciences, Transport & Planning Section, Delft, The Netherlands.

Braess, D., A. Nagurney, T. Wakolbinger (2005) On a paradox of traffic planning, in: *Journal Transportation Science*, volume 39, pp. 446-450.

Cantarella, G.E., G. Pavone, A. Vitetta (2006) Heuristics for urban road network design: Lane layout and signal settings, in: *European Journal of Operational Research,* 175, 3 pp. 1682-1695.

Clark, S., Watling, D. (2005) Modelling network travel time reliability under stochastic demand, in: *Transportation Research Part B,* 39, pp. 119-140.

Chen, A., H. Yang, H.K. Lo, W.H. Tang (2002) Capacity reliability of a road network; an assessment methodology and numerical results, in: *Transportation Research Part B*, 36, pp. 225-252.

Chen, Y., H.J. van Zuylen, M.C. Poelman, W. van der Hoeven (2010) Estimating a large-scale dynamic origin-destination matrix with conservation of a priori structure, in:

Proceedings of the Third International Symposium on Dynamic Traffic Assignment, Takayama, Japan.

Choi, B.. S. Moon, Z. Zhang, K. Papagiannaki, C. Diot (2007) Analysis of point to point packet delay in an operational network, in: *Computer Networks*, 51, 13, pp. 3812-3827

Chiou, S (2005) Bilevel programming for the continuous transport network design problem, in: *Transportation Research part B*, 39 pp. 361-383.

Chootinan, P. S.C. Wong, A. Chen (2005) A Reliability-Based Network Design Problem, in: *Journal of Advanced Transportation*, 39, 3, pp. 247-270.

Corthout, R., C.M.J. Tampère, L.H. Immers (2009) Marginal incident computation: an efficient algorithm to determine congestion spillback due to incidents, in: *Transportation Research Record: Journal of the Transportation Research Board*, 2099, pp. 22-29.

CPB (2006) *WLO mobiliteitscenarios met prijsbeleid*, CPB Notitie opgesteld in samenwerking met de Adviesdienst Verkeer en Vervoer van Rijkswaterstaat.

Davis, G. A. (1994) Exact local solution of the continuous network design problem via stochastic user equilibrium assignment, in: *Transportation Research part B*, 28, 1, pp. 61–75.

Deb, K. (2002) *Multi-Objective Optimization Using Evolutionary Algorithms*, Wiley, Chichester, United Kingdom.

DHV (2007) *Onderzoek ontvlechten verkeersstromen*, Amersfoort, the Netherlands.

Drake, R.L., W. Vogl, A.W.M. Mitchell (2009) *Gray's Anatomy for students*, Elsevier Churchill Livingstone, Toronto, Ontario, Canada, ISBN: 139780443069529

Duijnisveld, M., M. Snelder, M. Minderhoud, S. Calvert (to be published) *Kalibratieproces verkeersmodel verkeershinder A10-A1,* TNO-rapport, Delft, the Netherlands.

Dutch Ministry of Transport, Public Works and Water Management (1994) *Projectstudie Ontvlechting*, registration number VB-SE20060152.

Dutch Ministry of Transport, Public Works and Water Management and the Ministry of VROM (2004) *Nota Mobiliteit,* the Hague, the Nethelands.

Dutch Ministry of Transport, Public Works and Water Management (2008) *MobiliteitsAanpak, Vlot en veilig van deur tot deur,* the Hague, the Netherlands.

D'Este, G. M., M. A. P. Taylor (2003) Network vulnerability: an approach to reliability analysis at the level of national strategic transport networks. In: Bell, M.G.H., Y. Iida (Eds.), *The Network Reliability of Transport. Proceedings of the 1st International Symposium on Transportation Network Reliability (INSTR)*. Pergamon, Oxford, England, pp.23–44.

De Romph, E. (1994) *A dynamic traffic assignment model: Theory and applications*, PhD Thesis, Delft University of Technology, Delft, the Netherlands.

Egeter, B., T. Vonk, G. Tertoolen, T. van Rooijen, B. Immers (2007) *ARKO, Architectuur Wegenknooppunten*, TNO-rapport 2007-D-R0958/B, Delft, the Netherlands.

Egeter, B., M. Snelder (2010) ARKO-werkboek, *Een ontwikkelingsbeeld voor het wegennet in de Amsterdamse region op basis van de ARKO-methodiek*, Bart Egeter Advies, Rotterdam, the Netherlands.

Eliasson, J. (2004) Car drivers valuations of travel time variability, unexpected delays and queue driving, in: *Proceedings of the European Transport Conference 2006*, Strasbourg, France.

Eliasson, J. (2006) Forecasting travel time variability, in: *Proceedings of the European Transport Conference 2006*, Strasbourg, France.

Friesz, T.L., H. Cho, N.J. Mehta, R.L. Tobin, G. Anandalingam (1992) A Simulated Annealing Approach to the Network Design Problem with Variational Inequality Constraints, in: *Transportation Science*, Vol. 26, No. 1, pp. 18-26.

Friesz, T. L., G. Anandalingam, N. J. Mehta, K. Nam, S. J. Shah, R. L. Tobin (1993) The multiobjective equilibrium network design problem revisited: A simulated annealing approach, in: *European Journal of Operational Research*, 65, pp. 44–57.

Fosgerau, M., K. Hjorth, C. Brems, D. Fukuda (2008) *Travel time variability. Definition and valuation*. DTU Transport.

Gao, Z., J. Wu, H. Sun (2005) Solution algorithm for the bi-level discrete network design problem, in: *Transportation Research Part B*, 39, pp. 479–495.

Goldberg, D. E. (1989) *Genetic Algorithm in Search, Optimization and Machine Learning*, Addison-Wesley, Boston.

Groothedde, B. (2005) *Collaborative Logistics and Transportation Networks*, PhD Dissertation Delft University of Technology, TRAIL Research School, Delft.

Groot Koerkamp, M. (2007) Who's in charge? - Regulatory capture in de grondroerdersregeling.

Hall, F.L., K. Agyemang-Duah (1991) Freeway capacity drop and the definition of capacity, in: *Transportation Research Record: Journal of the Transportation Research Board*, 1320, pp. 91-98.

Hamilton, C. (1991) Telecommunication network dependability: A baseline on local-exchange network availability, in: *Proceedings IEEE Annual Reliability and Maintainability Symp Press*, Piscataway, N.J.

Harary, F. (1995) *Graph Theory*, Perseus, Cambridge, MA.

He, Y. (1997) *A flow-based approach to the dynamic traffic assignment problem: Formulations, algorithms and computer implementations*. MSc. Thesis, Massachusetts Institute of Technology, MIT Press, Cambridge MA, USA.

Hudyma, Fels (2004) Causes of failure in IT telecommunications networks, in: Proceedings of SCI, Florida, pp.35-38.

Husdal, J. (2004) Reliability and Vulnerability versus Cost and Benefits, Molde University College, in: *Proceedings of the second International Symposium on Transportation Network Reliability (INSTR),* Christchurch and Queenstown, New Zealand*, pp. 180-186*.

Hussain I. (2004) Fault-Tolerant IP and MPLS Networks, Cisco Press.

Iida, Y., H. Wakabayashi (1989) An approximation method of terminal reliability of a road network using partial minimal path and cut set, in: *Proceedings of the Fifth WCTR, Yokohama*, pp. 367-380.

Immers, L.H., P. van Koningsbruggen (2004a) *Het hybride karakter van verkeersmanagement*, Position paper geschreven in opdracht van Verkeerscentrum Nederland, TNO Inro.

Immers, L.H., A. Bleukx, M. Snelder, J.E. Stada (2004b) Betrouwbaarheid van reistijden: het belang van netwerkstructuur en optimale wegcapaciteit, in: *Tijdschrift Vervoerswetenschap*, jaargang 40, nr. 4, pp. 18-28.

Immers, L.H., J.E. Stada, I. Yperman (2004c) Reliability and robustness of transportation networks; problem survey and examples, in: *proceedings of the NECTAR Cluster Meeting on Reliability of Networks*, Amsterdam, March 19-20.

Immers, B., J. Stada, I. Yperman, A. Bleukx (2004d) Towards robust road network structures, in: *Slovak Journal of civil engineering*, XII (4), pp. 10-17.

Immers, L.H., B. Egeter, R. van Nes (to be published) Transport Network Planning: Methodology and Theoretical Notions, in: *Transportation Engineers' Handbook*, Editor: M. Kutz.

Inspectie Verkeer en Waterstaat and Rijkswaterstaat (2007) *Veiligheidsbalans 2007,* the Hague, the Nethelands.

Jacimovic, A. (2010) *Finding the Weakest Link, A comparative analysis of the road-network-specific measures of link vulnerability and a set of extended topological metrics of link vulnerability from the general network theory,* University College Utrecht and TNO.

Jamakovic, A., H. Wang, P. Van Mieghem (2006) Topological Characteristics of the Dutch Road Infrastructure, in: *Proceedings Seminar Infrastructure Reliability*, Delft, 2006.

Jamakovic, A., S. Uhlig, I. Theisler (2008) On the relationship between topological metrics in real-world networks, in: *Networks and heterogeneous media*, Volume 3, Number 2, pp. 345–359.

Janke, M. (2008) Wide area network outage analyses, *[http://blog.lastinfirstout.net/2008/10/wide-area-network-outage-analysis.html]*.

Jenelius, E. (2007) Incorporating Dynamics and Information in a Consequence Model for Road Network Vulnerability Analysis, in: *Proceedings of Third International Symposium on Transport Network Reliability*, The Hague, the Netherlands.

Jou, R.C., Lam, S.H., Y.H., Liu, K.H. Chen (2005) Route switching behavior on freeways with the provision of different types of real-time traffic information, in: *Transportation Research Part A*, 39(5), pp. 445–461.

Johnson, D.S., J.K. Lenstra, A.H.G.R. Kan (1978) *The complexity of the network design problem*, Publication report 7721/MO. Erasmus Universiteit Rotterdam, Econometrisch Instituut, Rotterdam, the Netherlands.

Jonkers, E., J.M. Schrijver, M. Muller (2009) Ietsdoen wellicht niet nodig: tijdverlies als gevolg van vrachtauto-ongevallen door een objectieve bril, in: *Proceedings Colloquium Vervoersplanalogisch Speurwerk*, Antwerpen, Belgium.

Ketelaars, S. (2010) *Effective Unbundling; Robustness by unbundling*, TNO, the Netherlands.

Klem E., M. Barten, M. Bijlsma (2004) *Inspiratieboek Robuustheid, Verkenning van robuustheidvergrotende maatregelen*, Traffic Test bv, Veenendaal, TT04-060.

Knoop, V.L., S.P Hoogendoorn, H.J. van Zuylen (2007) Quantification of the impact of spillback modelling in assessing network reliability, *CD-ROM. Transportation Research Board of the National Academies*, Washington, D.C.

Knoop, V. L., S. P. Hoogendoorn, H.J. van Zuylen (2008) The Influence of Spillback Modelling when Assessing Consequences of Blockings in a Road Network, in: *European Journal of Transportation and Infrastructure Research*, Vol. 8, no. 4. pp. 287-300.

Knoop, V.L. (2009) *Road Incidents and Network Dynamics, Effects on driving behaviour and traffic congestion*, T2009/13, December 2009, TRAIL Research School, the Netherlands.

Knoop, V.L., M. Snelder, H.J. van Zuylen, S.P. Hoogendoorn (2010) Link-level Vulnerability Indicators for Real-World Networks, in: *Proceedings of the 89th Annual Meeting of the Transportation Research Board*, Washington D.C., United States of America, Accepted for publication.

Koo, R., Y. Yim (1998) *Commuter Response to Traffic Information on an Incident*, Technical Report UCB-ITS-PWP-98-26, University of California, Berkely.

Kouwenhoven, M., R. van Grol, E. Kroes (2004) *Ontwikkeling ex ante Instrument Betrouwbaarheid*, RAND Europe.

Kouwenhoven, M., G. de Jong, P. Rietveld (2005) *Reliability ratio's voor het goederenvervoer*, RAND Europe, WR-274-AVV.

Kraaijeveld, R. (2008) *Onderzoek naar de identificatie van kwetsbare wegvakken*, Master thesis, Delft University of Technology, Delft, the Netherlands.

Kuhn, D. R. (1997) Sources of Failure in the Public Switched Telephone Network, in: *Computer*, pp. 31-36.

LeBlanc, L.J. (1975) An Algorithm for the Discrete Network Design Problem, in: *Transportation Science*, pp. 183-199.

Li, M. (2008) *Robustness Analysis for Road Networks, A framework with combined DTA models*, T2008/14, December 2008, TRAIL Research School, the Netherlands.

Li, H. (2009) *Reliability-based Dynamic Network Design with Stochastic Networks*, T2009/3, May 2009, TRAIL Research School, the Netherlands.

Liang, C., C. Wang, H. Luh, P. Hsu (2009) Disaster avoidance mechanism for content-delivering service, in: *Computers and Operations Research*, Volume 36 , Issue 1, pp. 27-39 .

Liew, S.C., K.W. Lu (1992) A framework for network survivability characterization, in: *Proceedings of the IEEE International Conference on Communications, Conference record, SUPERCOMM/ICC '92, Discovering a New World of Communications*.

Liu, Y. (2001) *Spare capacity allocation: model analysis and algorithm*, University of Pittsburgh.

Lo, H.K., Y.K. Tung (2001) Network Design for Improving Trip Time Reliability, in: *Proceedings of the Transportation Research Board, 80th annual meeting*, TRB ID: 01-2204.

Lo, H.K., Y. Tung (2003) Network with degradable links: capacity analysis and design, in: *Transportation Research Part B*, 37, pp. 345-363.

Lomax, T., D. Schrank, S. Turner, R. Margiotta (2003) *Selecting travel reliability measures*. Texas Transportation Institute, Cambridge Systematics Inc.

Luathep, P., A. Sumalee, H.W. Ho, F. Kurauchi (2010) Sensitivity analysis based approach for vulnerability analysis of large-scale road network, in: *Proceedings of the 12th WCTR*, July 11-15, 2010 – Lisbon, Portugal.

Magnanti, T.L., R.T. Wong (1984) Network Design and Transportation Planning: Models and Algorithms, in: *Transportation Science*, Vol. 18, No. 1, 1984, pp. 1-48.

Mahut, M. (2000) *Discrete flow model for dynamic network loading*, Ph.D. Thesis, Département d'informatique et de recherhe opérationelle, Université de Montréal. Published by the Center for Research on Transportation (CRT), University of Montreal.

Martens M.J., T. Vonk, A. Hendriks (2004) *De kwaliteit van regionale wegennetwerken: een vergelijking tussen Engelse, Duitse en Nederlandse regio's*, TNO Inro rapport 2004-21, Delft.

McDonald, J.C. (1992) Public networks - dependable?, in: *Communications Magazine, IEEE,* 30, 4, pp. 110-112.

McDonald, J.C. (1994) Public network integrity: avoiding a crisis in trust, in: IEEE Journal on Selected Areas in Communications, 12, 1, pp. 5 -12.

Meeuwissen, A.M.H. (2003) Exchange between congestion cost and cost for infrastructure, in: *Proceedings of the 19th Dresden Conference on Traffic and Transport*, Dresden.

Meeuwissen, A.M.H., M. Snelder, J.M. Schrijver (2004) *Statistische analyse variabiliteit reistijden voor SMARA*, Delft, TNO Inro, juli 2004, no. Inro V&V/2004-31, 04-7N-112-73401, ISBN-number 90-5986-095-0.

Meng, Q., H. Yang, M.G.H. Bell (2001) An equivalent continuously differentiable model and a locally convergent algorithm for the continuous network design problem, in: *Transportation Research Part B*, 25, pp. 83-105.

Möller, N., S.O. Hansson (2008) Principles of engineering safety: Risk and uncertainty reduction, in: *Reliability Engineering and System Safety*, 93, pp. 776-783.

Muller, M. (2009) *Ex ante evaluatie van de invloed van infrastructuurmaatregelen op de effecten van incidenten*, TNO, Delft, The Netherlands.

Murray-Tuite P.M., H.S. Mahmassani (2004) Methodology for Determining Vulnerable Links in a Transportation Network, in: *Transportation Research Record: Journal of the Transportation Research Board*, No.1882, Transportation Research Board of the National Academies, Washington, D.C., pp. 88-96.

Nagurney, A., Q. Qiang (2009) *Fragile Networks, Identifying Vulnerabilities and Synergies in an Uncertain World*, Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Newell, G.F. (1993) A simplified theory of kinematic waves in highway traffic, Part I: General theory, Part II: Queuing at freeway bottlenecks, Part III: Multi-destination flows, in: *Transportation Research* 27B, pp 281-313.

Nicholson, A., J. Schmöcker, M.G.H. Bell, Y. Iida (2001) Assessing Transport Reliability: Malevolence and User Knowledge, in: *The Network Reliability of Transport*, editors M.G.H. Bell, Y. Iida, Proceedings of the 1st International Symposium on Transportation Network Reliability, pp. 1-22, Elsevier Science, Kidlington.

Ogryczak, W., A. Ruszczynski (1999) From stochastic dominance to mean-risk models: Semideviations as risk measures, in: *European Journal of Operational Research*, 116, pp. 33-50.

OECD/ITF (2009), *Improving reliability on surface transport networks*.

PBL (2006) *Wegen naar economische groei*, NAi Uitgevers, Rotterdam.

Pearman, A.D. (1979) The structure of the solution set to network optimization problems, in: *Transportation Research part B*, 13, 1, pp. 81-90.

Peer, S., C. Koopmans, E. Verhoef (2010) Predicting Travel Time Variability for Cost-Benefit Analysis, in: *Proceedings of WCTR 2010*, Lisbon, Portugal.

Poorzahedy, H., M.A. Turnquist (1982) Approximate algorithms for the discrete network design problem, in: *Transportation Research part B*, Vol. 16B, No. 1, pp. 45-55.

Roetman, F. (2008) *Externe bedreigingen voor telecommunicatienetwerken*, TNO.

Rooij, A.de (2000), *Fysica van Samenwerking, naar een Krachtenfusie van Burgers, Bestuurders, Bureaucraten en Bedrijven*, Rijkswaterstaat, The Hague, the Netherlands.

Santos, B., A. Antunes, E. Miller (2005) *Solving an accessibility-maximization road network design model: a comparison of heuristics*, Publishing House of Poznan University of Technology, pp. 692-697.

Schrijver, J., B. Egeter, B. Immers, M. Snelder (2008) *Visie Robuust Wegennet ANWB*, TNO-rapport 2008-D-R0661/C, Delft.

Smulders, S. (1988) *Modelling and Filtering of Freeway Traffic Flow*, Report OS-R8706, Centre of Mathematics and Computer Science, The Netherlands.

Snelder, M. (2003) *ROADnet: Optimalisatiealgoritme voor het ontwerpen van wegennetwerken: methode en toepassing*, TNO, Delft, the Netherlands.

Snelder, M., A. Wagelmans, J.M. Schrijver, H.J. van Zuylen, L.H. Immers (2007), Optimal Redesign Of The Dutch Road Network, in: *Transportation Research Record: Journal of the Transportation Research Board*, 2029, pp. 72-79.

Snelder, M, J.M. Schrijver, R. Landman, J. Mak, M. Minderhoud (2008) *De kwetsbaarheid van Randstedelijke vervoernetwerken uit verkeerskundig perspectief*, TNO, 2008-D-R0882/A, Delft, the Netherlands.

Snelder, M., J.M. Schrijver, L.H. Immers, B. Egeter (2009a) Architecture for Designing Robust Road Networks, *Proceedings of the 88th Annual Meeting of the Transportation Research Board*, Washington D.C., United States of America.

Snelder, M. (2009b) *A Comparison between Dynameq and Indy*, Cirrelt-2009-48, Montréal, Canada.

Snelder, M. L.A. Tavasszy, T. van Rooijen, W. Spit, E. Devillers, E. Buckman (2009c) *Kansrijke combinaties van de concepten uit de MobiliteitsAanpak: Een economische toets en een uitwerking van de concepten naar gebiedstype*, TNO and ECORYS, TNO-034-DTM-2009-02313, Delft, the Netherlands.

Snelder, M., J. Mak, L.A. Tavasszy (2010a) *Kwetsbaarheidsanalyse NMCA*, TNO, TNO-034-DTM-2010-01984, Delft, the Netherlands.

Snelder, M. L.A. Tavasszy (2010b) Quantifying the full reliability benefits of road network improvements, in: *Proceedings of WCTR 2010*, Lisbon, Portugal.

Snelder, M., B. Egeter, T. van Rooijen, L.H. Immers (2010c) *✱P+R, ANWB-visie op een nieuw type P+R, Uitgewerkt voor de regio Rotterdam - Den Haag*, TNO-034-DTM-2010-03924, Delft, the Netherlands.

Snelder, M., M. Muller, J.M. Schrijver (2010d) *Begrippenkader en indicatoren voor robuustheid*, TNO-034-DTM-2009-05026, Delft, The Netherlands.

Steenbrink, P. A. (1974) *Optimization of Transportation Networks*, John Wiley & Sons, London.

Sumalee, A., D.P. Watling, S. Nakayama (2006) Reliable Network Design Problem: the case with uncertain demand and total travel time reliability, in: *in Proceedings of the Transportation Research Board of the National Academies,* Washington D.C., United States of America.

Suwansirikul, C., T.L. Friesz, R.L. Tobin (1987) Equilibrium Decomposed Optimization: A Heuristic for the Continuous Equilibrium Network Design Problem, in: *Transportation Science*, Vol. 21, No. 4, pp. 254-263.

Tamminga, G.F., J.C. Maton, R. Poorteman, J. Zee (2005) *De Robuustheidscanner, Robuustheid van netwerken: een modelmatige verkenning,* Grontmij Nederland, I&M-99366053-GT/mk.

Tampère, C.M.J., J. Stada, L.H. Immers (2006) Identifying vulnerable sections in a national road network. Application to the road network of Flanders, in: *Proceedings of the seminar Infrastructure Reliability*, TU Delft, June 22 2006.

Tampère C.M.J., J. Stada, L.H. Immers (2007) Methodology for Identifying Vulnerable Sections in a National Road network, *in Proceedings of the Transportation Research Board of the National Academies*, Washington, D.C..

Taylor, M.A.P., D'Este, G.M.D. (2003) Concepts of network vulnerability and application to the identification of critical elements of transport infrastructure, in: *Proceedings of the 26th Australian Transport Research Forum*, Wellington, New Zealand.

Tero, A., S. Takagi, T. Saigusa, K. Ito, D.P. Bebber, M.D. Fricker, K. Yumiki, R. Kobayashi, T. Nakagaki (2010) Rules for Biologically Inspired Adaptive Network Design, in: *Science* 327, pp. 439-442.

Texas Transportation Institute and Cambridge Systems Inc. (2006) *Travel Time Reliability: Making It There On Time, All The Time*.

Tu, H., J.W.C. van Lint, H.J. van. Zuylen (2007) Impact of Traffic Flow on Travel Time variability of Freeway Corridors, in: *Transportation Research Record: Journal of the Transportation Research Board,* 1993, pp. 59-66.

Ukkusuri, S.V., T.V. Mathew, S.T. Waller (2007) Robust transportation network design under demand uncertainty, in: *Computer-Aided Civil and Infrastructure Engineering*, 22, pp. pp. 6–18.

Van Beek, F., H. Flikkema, J. Francke, P. Besseling, W. Groot, H. Nijland, J. Ritsema van Eck (2006) Welvaart en Leefomgeving, Mobiliteit, [www.welvaartenleefomgeving.nl].

Van de Riet, O.A.W.T. (1998) A System's View on the Transport System from a Dutch Policy Perspective, in: *Proceedings of the 4th TRAIL PhD Congress*, TRAIL Research School, Delft.

Van Lint, J. W. C., Van der Zijpp, N. J. (2003), 'Improving a travel time estimation algorithm by using dual loop detectors', in: Transportation Research Record: Journal of the Transportation Research Board, 1855, pp. 41–48.

Van Lint, J.W.C., H.J. van Zuylen, H. Tu (2008) Travel time unreliability on freeways: Why measures based on variance tell only half the story, in: *Transportation Research Part A*, 42, pp. 258–277.

Visser, H. (2009), Dutch inland waterway system plays important part in goods transport, [www.cbs.nl]

Wardrop, J.G. (1952) Some Theoretical Aspects of Road Traffic Research, in: *Proceedings of the Institution of Civil Engineers Part II*, I, pp. 325-362.

Wakabayashi, H., Y. Iida (1992) Upper and lower bounds of terminal reliability of road networks: an efficient method with Boolean algebra., in: *Journal of Natural Disaster Science*, 14, pp. 29-44.

Waller, S.T., A.K. Ziliaskopoulos (2000) The Stochastic Dynamic Network Design Problem, Transportation Research Board, in: Proceedings of the 80th annual meeting, paper 01-2517.

West, D.B. (1995) *Introduction to Graph Theory*, Prentice-Hall, Englewood Cliffs, NJ.

Willinger, W., R. Govindan, S. Jamin, V. Paxson, S. Shenker (2002) Scaling phenomena in the Internet, in: *Proceedings of the National Academy of Sciences*, 99, suppl. 1, pp. 2573–2580.

Wilmink, I., J.M. Schrijver, H.J.M. Puylaert, M.J. Martens, L.H. Immers, A.H. Hendriks, B. Egeter (2002) *IRVS: Design Method for an Integrated Regional Transport System*, TNO-Inro, ISBN: 90-6743-949-5.

Wilmink, I., A. van den Broeke, L. Berghout (2003) De taakverdeling tussen netwerkstructuur en verkeersmanagement, in: *Colloquium Vervoersplanologisch Speurwerk 2003: No pay, no queue? Oplossing voor bereikbaarheidsproblemen*, Delft, CVS.

Xiong, Y., J.B. Schneider (1992) Transportation Network Design Using a Cumulative Genetic Algorithm and Neural Network, in: *Transportation research record 1364*, pp. 37-44.

Yang, H., M.G.H. Bell (1998) Models and algorithms for road network design, a review and some developments, in: *Transport Reviews*, Vol. 18, No. 3, pp. 257-278.

Yperman, I. (2007) *The Link Transmission Model for Dynamic Network Loading, Katholieke Universiteit Leuven*, Faculteit Ingenieurs Wetenschappen, Afdeling Verkeer en Infrastructuur, Leuven.

Yperman, I., Snelder M. (2010) Evaluating network performance using different dynamic traffic models within a DTA framework, in: *Proceedings of the Third International Symposium on Dynamic Traffic Assignment*, Takayama, Japan.

Zolfaghari,A. F.J. Kaudel (1994) Framework for network survivability performance, in: *Selected Areas in Communications, IEEE Journal*, 12, 1, pp 46-51.

Zorpette, G. (1989) Keeping the phone lines open, in: *Spectrum*, 26, 6, pp. 32 – 36.

Appendix A: The robustness of road networks compared to the robustness of non-transport networks

The human body: Cardiovascular and nervous systems[36]

In this section, the cardiovascular and nervous systems are described and compared with a road network. Both systems, as shown in Figure A.1, are networks within the human body.

The cardiovascular system

The cardiovascular system distributes blood through blood vessels. There are two circulation loops: a short lung circulation loop where blood is oxygenated, and a circulation loop through the rest of the body to provide oxygenated blood. The amount of blood varies between 5 and 8 litres. The average adult has about 5 litres of blood, which consists of plasma, red blood cells, white blood cells, and platelets.

There are three major types of blood vessels: the arteries, which carry the blood away from the heart in the body circulation loop; the capillaries, which enable the actual exchange of water and chemicals (like oxygen) between the blood and the tissues; and the veins, which carry blood from the capillaries back towards the heart. In between the arteries and the capillaries are the arterioles. The venules are in between the capillaries and the veins. The total length of the blood vessels is about a 100,000 kilometres, which is 2½ times the circumference of the earth. The total length of the capillaries is by far higher than the total length of the veins and arteries. The capacity of blood vessels is flexible, since the diameter of the vessels can vary. The three types of blood vessels cannot be considered as independent

[36] The medical information in this section is based on interviews with doctor H. Mulder and medical student S.M. Snelder, Gray's Anatomy for students (Drake et al.,2009) and www.wikipedia.org.

systems. As described above, they form a closed loop. The blood flows from the arteries, to the capillaries, to the veins, and again to the arteries, etc.

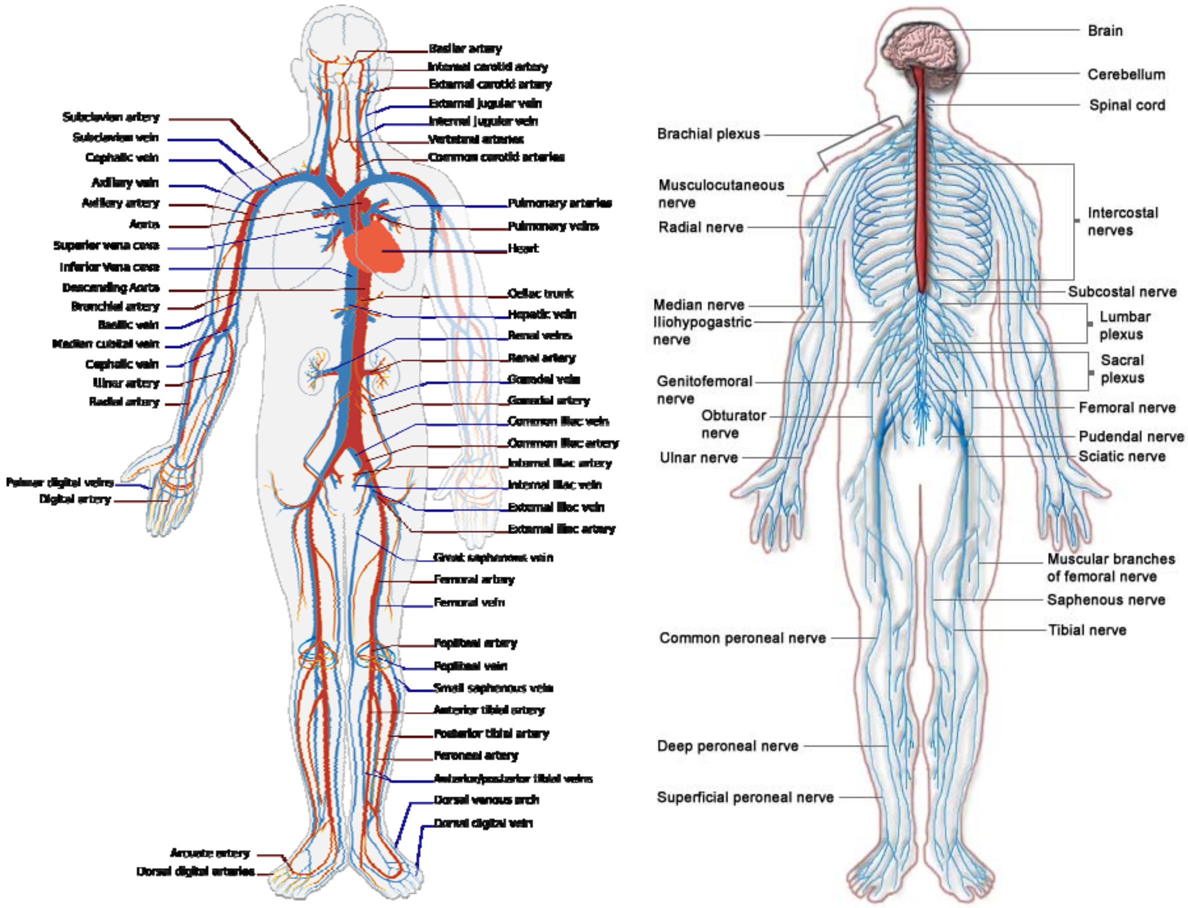


Figure A.1: Cardiovascular system (left) and the human nervous system (right)[37]

The most important function of the cardiovascular system is to transport oxygen (via the blood) to the different parts of our body. Transport of oxygen to the brains and the hart is the number one priority. Oxygenating the other parts of the body has a much lower priority.

There are several reasons why the circulatory system sometimes does not function optimally (disturbances):

- Temporarily or longer lasting (small) changes in blood pressure as a result of a shock, hypertension, a disease, etc.
- Thrombosis is the formation of a blood clot (thrombus) inside a blood vessel, obstructing the flow of blood through the circulatory system. When a blood vessel is damaged, the body uses platelets and fibrin to form a blood clot, because the first step in repairing it (haemostasis) is to prevent loss of blood. If that mechanism causes too much clotting, the vessel's diameter will be corrected by a sophisticated clot dissolution system (fibrinolysis). Eventually the clot breaks free and an embolus is formed. When a thrombus occupies more than 75% of the surface area of the lumen of an artery, blood flow to the tissue is reduced enough to cause symptoms because of decreased oxygen (hypoxia) and accumulation of metabolic products like lactic acid. More than 90% obstruction can result in anoxia (the complete deprivation of oxygen) and infarction (a mode of cell death).

[37] Source: www.wikipedia.org

Blood flow obstruction in the veins of the body results in blood congestion before the obstruction.

- An aneurysm or aneurism is a localized, blood-filled dilation (balloon-like bulge) of a blood vessel caused by a trauma, disease, or weakening of the vessel wall. Due to the aneurysm, the blood flow is converted to a turbulent flow, and, despite the larger diameter of the vessel, the flow speed is reduced. Aneurysms most commonly occur in arteries at the base of the brain and in the aorta (the main artery coming out of the heart, a so-called aortic aneurysm). As the size of an aneurysm increases, there is an increased risk of rupture, which can result in severe bleedings or other complications including sudden death.
- The blood vessels are located as deep as possible within the body. However, at some locations they are very close to skin, which make them susceptible to traumatic impacts.

The cardiovascular system has some regulation mechanisms to recognize and to deal with the above mentioned disturbances:

- Within the blood vessels, there are receptors that monitor the flow through the vessels by continuous pressure management. The so called baroreceptors detect the pressure of blood flowing through them, and can send messages to the central nervous system to increase or decrease total peripheral resistance (diameter of the blood vessels) and cardiac output. Baroreceptors act immediately as part of a negative feedback system called the baroreflex as soon as there is a change from the usual blood pressure, returning the pressure to a normal level. This is an example of a short term blood pressure regulation mechanism. This makes the cardiovascular system very resilient.
- When the baroreceptors are not working, the blood pressure continues to increase, but within an hour the blood pressure returns to normal as other blood pressure regulatory systems take over. In the long term, the blood volume can, for instance, be corrected.
- In a state with very low blood pressure (shock), there is also a redistribution of blood flow to maintain a good flow to the heart and brain. The flow to e.g. the skin virtually stops, resulting in a pale skin.
- For most of the arteries there is backup artery that can temporarily take over its function. This is most obvious in the forearms, the lower legs, and the neck, where there are two arteries that can temporarily take over each other’s function. The capacity of the one that is used can instantaneously be increased when it needs to take over the function of the other artery. In the long term this will cause problems anyway. For capillaries, there is no backup system. The areas that have to be oxygenated by the capillaries can not be reached by other capillaries. Thrombosis in the veins is less of a problem, since the veins carry deoxygenated blood. If one vein is blocked, the circulation most of the time continues via other veins. In terms of a road network: there are many route alternatives available for veins. They are automatically used, since the blood is pressed in the direction in which it has the least resistance.
- Not only the diameter of the vessels is flexible, but the diameter of blood cells is flexible as well. The shape of blood cells can be changed in such a way that they can pass through small passages.
- Formation of new blood vessels (angiogenesis): Angiogenesis is a physiological process involving the growth of new blood vessels from pre-existing vessels. Vasculogenesis is the term used for spontaneous blood-vessel formation, and intussusception is the term for new blood vessel formation by splitting off existing ones. Angiogenesis is a normal and vital process in growth and development, as well as in wound healing. However, it is also a fundamental step in the transition of a tumor from a dormant state to a malignant state.

- Besides these correction mechanisms within the body, bypasses and stents can be created and blood clots can be removed surgically. There are also clot dissolving medicaments.

In summary:
- The cardiovascular system is very flexible. The central nervous system can instantaneously increase or decrease the capacity of blood vessels and increase or decrease the heart rate, and it can also instantaneously give priority to oxygenating certain parts of the body. Furthermore, there are some long term regulation mechanisms available that can, for instance, correct the blood volume.
- There is some temporarily redundancy for most of the arteries. For the capillaries there is no redundancy. For the veins there are many alternatives that can automatically be used optimally.
- The blood vessels are located as deep as possible within the body. However, at some locations they are very close to skin, which make them susceptible to traumatic impacts.

Comparison of the cardiovascular system with the road network.
- If we compare the different types of blood vessels to the different road types, the most appropriate comparison would be that the capillaries are the local roads, the veins are the secondary network, and the arteries are the motorways. With respect to the flow speeds, this comparison goes right. However, with respect to the order in which the types are used, there is not a 100% match. This would for instance imply that blood flows from motorways (arteries) to local road (capillaries) to secondary roads (veins) and back again to motorways etc. In road networks the order is usually local road, secondary road, motorway, secondary road, local road. However, sometimes some road types are skipped or used in another order. In the cardiovascular system the sequence is always the same.
- The road network is designed to transport passenger and freight in cars and trucks (and other modes of transport). These passengers and freight products have specific origins and destinations, whereas blood (which is transported within the cardiovascular system) does not have a specific origin and destination. The only restriction is that all parts of the body should be oxygenated by blood cells (it is not important which blood cell), which implies that the blood should pass by the lungs and all parts of the body. The fact that blood cells do not have specific origins and destinations makes it easier to circulate the blood through different routes. Furthermore, blood cells are flexible and can stick to each other, whereas cars and trucks cannot bump into each other and cannot reshape in order to pass by small passages. This is also a reason why the transport of blood is easier, and more homogeneous flows occur. Besides this, human behaviour plays an important role in driving cars and trucks. This is not the case with blood cells, which also makes it easier to obtain system optimal flows instead of user optimal flows.
- In the cardiovascular system, congestion is less likely to occur than in road networks. This is not only because of the factors described above, but also because of the fact that blood is circulated under high pressure, which forces the blood to find its own way past bottlenecks. As a consequence, spillback effects that occur on the road when queues spill back to other roads do not occur in the cardiovascular system.
- If disturbances occur in the cardiovascular system, there is an immediate response. The baroreceptors in the blood vessels immediately detect changes in pressure, and an automatic nervous signal is sent to the blood vessels to adjust their diameter in order to correct the pressure. The heart rate can be adjusted as well. In road networks the detection of incidents is not instantaneous, but can happen relatively quickly by means of monitoring systems and phone calls. The response takes more time. Cars and trucks have to be removed from the road, and it is not possible to increase the road capacity (of

alternative roads) as is done in blood vessels. There are some mechanisms in the road network by which the capacity of certain routes could be increased by changing the traffic signal priorities and by opening roads whose usage is discouraged by time regulated barrier systems. However, in practice theses systems are not (or hardly ever) adjusted in case of disruptions. Furthermore, it is not possible to reroute all vehicles optimally as is done in veins, since humans cannot be guided in the way that blood cells can.

- When partial blockings of blood vessels occur, buffers are created by extending the diameter of the blood vessels. In road networks, buffers cannot be flexibly created.
- Bypasses can be created surgically (and new blood vessels can be created by the human body itself under specific circumstances), which is by far less expensive than creating new bypassing roads.

The nervous system

The nervous system is a network of specialized cells that send signals from one part of the body to another. The human nervous system consists of two parts ─ central and peripheral. The central nervous system contains the brain and spinal cord. The neurons of the central nervous system are interconnected in complex arrangements, and transmit electrochemical signals from one to another. The peripheral nervous system consists of sensory neurons, clusters of neurons called ganglia, and nerves connecting them to each other and to the central nervous system. Sensory neurons are activated by inputs impinging on them from outside or inside the body, and send signals that inform the central nervous system of ongoing events. The interaction of the different neurons form neural circuits that regulate an organism's perception of the world and its body and behaviour.

The nervous system is characterized by the fact that there is only one way up and down from A to B. Nerves can be damaged by traumatic impacts from outside, or by internal diseases like the autoimmunity disease Multiples Scleroses (MS). Damages are mostly irreversible and cannot be compensated by other nerves. Furthermore, nerves can be oppressed or the blood supply to the nerves (especially the brains) can be blocked. This can cause temporary paralysis. If it takes too long, the paralysis might become irreversible because the nerves get damaged.

The nerves are protected by the myelin sheath. Furthermore, nerves are protected by their location deep within the body (except for the ulnar nerve). The spinal cord is, for instance, protected by the spinal column. Therefore, a lot of preventive measures are taken that more or less compensate for the fact that adaptive strategies are not possible. The places at which nerves come together are, however, not well protected, and are, therefore, vulnerable.

Comparison of the nervous system with the road network.

Compared to the road network, the nervous system has no route alternatives, which makes the system more vulnerable than the road network. On the other hand, the nervous system is much better protected than the road network, which makes it more robust.

Telecommunication networks[38]

Telecommunication is the transmission of signals over a distance for the purpose of communication. Nowadays, we use telephones, television, radio, or computers for

[38] This section is mainly based on an interview with Prof. dr. ir. R. Kooij, http://en.wikipedia.org/wiki/Telecommunication, and http://en.wikipedia.org/wiki/Internet.

telecommunication. In this section, telecommunication networks are described, and a comparison is made between telecommunication networks and the road network.

A basic telecommunication system consists of three elements:
- A transmitter that takes information and converts it to a signal.
- A transmission medium that carries the signal.
- A receiver that receives the signal and converts it back into usable information.

For example, in a radio broadcast, the broadcast tower is the transmitter, free space is the transmission medium, and the radio is the receiver. Often, telecommunication systems are two-way, with a single device acting as both a transmitter and receiver or transceiver. For example, a mobile phone is a transceiver.

Signals can be either analogue or digital. In an analogue signal, the signal is varied continuously with respect to the information. In a digital signal, the information is encoded as a set of discrete values (for example ones and zeros). During transmission, the information contained in analogue signals will be degraded by noise. Conversely, unless the noise exceeds a certain threshold, the information contained in digital signals will remain intact. Noise resistance represents a key advantage of digital signals over analogue signals.

A network is a collection of transmitters, receivers, and transceivers that communicate with each other. Digital networks consist of one or more routers that work together to transmit information to the correct user. An analogue network consists of one or more switches that establish a connection between two or more users.

In an analogue telephone network, the caller is connected to the person he wants to talk to by switches at various telephone exchanges. The switches form an electrical connection between the two users, and the setting of these switches is determined electronically when the caller dials the number. Once the connection is made, the caller's voice is transformed to an electrical signal using a small microphone in the caller's handset. This electrical signal is then sent through the network to the user at the other end, where it is transformed back into sound by a small speaker in that person's handset. There is a separate electrical connection that works in reverse, allowing the users to converse.

In a radio and television broadcast system, the central high-powered broadcast tower transmits a high-frequency electromagnetic wave to numerous low-powered receivers. The high-frequency wave sent by the tower is modulated with a signal containing visual or audio information. The receiver is then tuned so as to pick up the high-frequency wave, and a demodulator is used to retrieve the signal containing the visual or audio information. The broadcast signal can be either analogue (signal is varied continuously with respect to the information) or digital (information is encoded as a set of discrete values).

The Internet is a global system of interconnected computer networks that use the standard Internet Protocol Suite (TCP/IP) to serve billions of users worldwide. Many computer scientists describe the Internet as a "prime example of a large-scale, highly engineered, yet highly complex system" (Willinger et al., 2002). It is a network of networks that consists of millions of private and public, academic, business, and government networks of local to global scope that are linked by a broad array of electronic and optical networking technologies. The Internet carries a vast array of information resources and services, most notably the inter-linked hypertext documents of the World Wide Web (WWW) and the infrastructure to support electronic mail.

In Figure A.2, a part of the Internet is shown schematically. Each line is drawn between two nodes, representing two IP addresses. The length of the line is indicative of the delay between the two nodes. Lines are colour coded according to their corresponding allocation, as follows: Dark blue: net, ca, us; Green: com, org; Red: mil, gov, edu; Yellow: jp, cn, tw, au, de; Magenta: uk, it, pl, fr; Gold: br, kr, nl; White: unknown.

Any computer on the Internet has a unique IP address that can be used by other computers to route information to it. Hence, in principle, any computer on the Internet can send a message to any other computer using its IP address. These messages carry with them the originating computer's IP address, allowing for two-way communication. The information packages that are sent over the Internet are routed with routers. In principle all the packages take the shortest path. However, priority can be given to some packages. Furthermore, if needed, the packages can easily be rerouted. The bandwidth of the fibreglass cables is important for the capacity, and therewith the speed, at which the messages are sent. The network structure is made in such a way that service level agreements can be made. These agreements often require a very high availability/reliability of 99.9%.

The Internet works in part because of protocols that govern how the computers and routers communicate with each other. The nature of computer network communication lends itself to a layered approach, where individual protocols in the protocol stack run more-or-less independently of other protocols. This allows lower-level protocols to be customized for the network situation while not changing the way higher-level protocols operate. A practical example of why this is important is because it allows an Internet browser to run the same code regardless of whether the computer it is running on is connected to the Internet through an Ethernet or Wi-Fi connection.

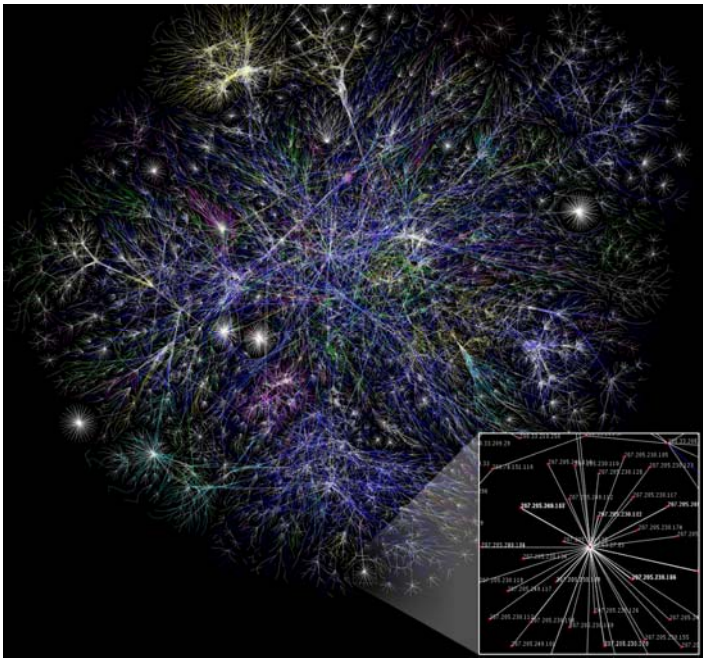


Figure A.2: Partial map of the Internet[39]

[39] Source: http://en.wikipedia.org/wiki/File:Internet_map_1024.jpg

The length of the telecommunication network in the Netherlands is estimated to be between 287,000 kilometres and 365,300 kilometres (Roetman, 2008). In the same report, the frequencies of different disturbances that occur on telecommunication networks are shown (Table A.1). The number of affected locations/households depends strongly on the level at which the disturbance occurs (close to the customer or far away from the customer). Power losses usually have a lower impact than other disturbances, because in most cases this happens at locations close to the customers. The probability that damage occurs close to the customer is larger than the probability that many users are shut down at the same time. Besides that, the locations where many connections come together are often made in a redundant way. The disturbances in the table mostly refer to the network infrastructure. However, bottlenecks can also occur, for instance, at overloaded servers. For a more detailed analysis of the vulnerability of telecommunication networks, see (Kuhn, 1997; Zorpette, 1989; Hamilton, 1991; Zolfaghari and Kaudel 1994; Liew and Lu 1992; McDonald, 1992 and 1994; Liu, 2001; Ogryczak and Ruszczynski, 1999; Choi et al., 2007; Liang et al., 2009; Hudyma and Fels, 2004; Hussain, 2004; Janke, 2008; Boog et al., 2007; Groot Koerkamp, 2007).

Table A.1: Disturbances on telecommunication networks in the Netherlands

	Number of failures per year	Number of failures per location per year	Mean time between failure per location (years)	Number of affected locations/households per failure
Digging damage	2820	0.00772	116	140
Power loss	18324	0.33	3	137
Fire damage	18954	0.00242	413	2129
Vandalism	202410	0.0259	39	1027
Water damage	35560	0.00455	250	6583
Software failures	98100	0.0125	80	2107
Maintenance errors	171674	0.0219	46	1211

Because telecom networks are often made redundant at the physical layer, and contain back-up routes at the routing layer, connections can usually be re-established in from 50 milliseconds to 30 seconds.

In order to deal with temporal overload situations, servers in telecom networks have buffers where requests are temporarily stored until the server is less busy.

The above mentioned adaptation strategies make sure that high service level agreements can be met. Nevertheless, sometimes disturbances still have a large effect. When this happens, the disturbance either occurs at a so called single-point-of-failure that was not discovered before (like power supply points), or the situation that causes the disturbance is very rare (like a heavy earthquake in the Netherlands).

Comparison of telecommunication networks with the road network.
Compared to the road network, telecommunication networks are often made more redundant. A lot of spare capacity is available. Furthermore, this spare capacity can more easily be used, because information packages can easily be rerouted, which is much more difficult compared to rerouting drivers. It is easier to make a telecommunication network redundant than to make a road network redundant, because the construction costs are much lower. Laying down a

cable costs on average about €25,000 – €35,000 per kilometre, whereas constructing a road costs on average about €8 million – €10 million per lane kilometre.

Furthermore, the speeds at which packages are sent, disturbances are recognized, and packages are rerouted are much higher than in the road network. Where a disturbance on the road network can easily cause a delay of several minutes up to more than an hour for drivers, disturbances on the Internet usually do not lead to more than a 30 second delay for the users of the Internet. Of course, in exceptional situations the delay on the Internet (mainly at the application side) can also take much longer.

Appendix B: Comparison among indicators for robustness

In this appendix, a selection of the indicators that are mentioned in section 3.5 are computed for an example network by using different modelling techniques. These computations give additional insights into the indicators. Furthermore, this analysis was done in order to learn whether or not it is possible to approximate the more complex indicators that are computed by the complex model, and therewith most likely to give the best description of robustness, with indicators that are easier to compute. For this purpose, some very simple indicators, like the intensities, are added to the comparison as well. If it is possible to give an accurate approximation of the complex indicators, a lot of computation time can be saved, which makes it easier to analyse the vulnerability of networks and to design robust road networks. The indicators are sorted beginning with simple indicators that can be directly computed based on the outcomes of the national or regional model system (or any other static traffic assignment model) and ending with the most complex index and the most complex computation method.

<u>1: Intensity:</u> the number of vehicles (or passenger car units) that pass a road section in an hour. In Figure B.1 the intensities (also referred to as flows) are presented for the example network. They are the result of model computations with the national model system (LMS). The flows are an important factor in the other indicators that are presented below as well. In general it holds that higher flows cause higher total travel time losses as a result of incidents. Of course, other factors, like the spare capacity on the route that is considered and on alternative routes, are important as well.

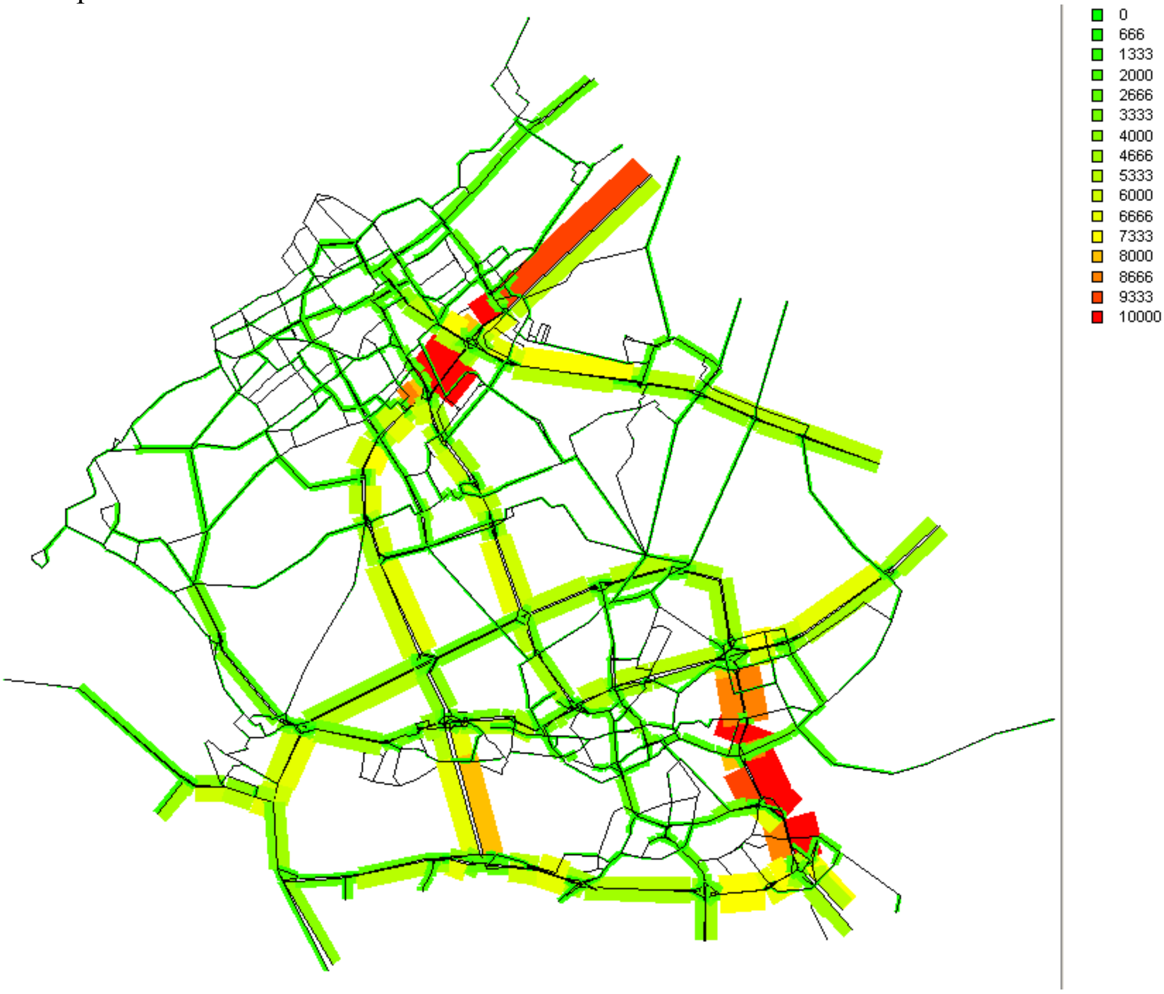


Figure B.1: Intensity (pcu/hour), based on LMS

<u>2: Spare capacity</u>: the spare capacity is the capacity minus the flow. Figure B.2 shows this indicator (both by width and by colour). The motorways have a lot of spare capacity. This is because, in this example network, a scenario is used in which a lot of infrastructure investments in the motorways are made. On the places where there is a lot of spare capacity, small incidents are likely to have small to no effects.

Figure B.2: Spare capacity (pcu/hour), based on LMS

3: I/C-ratio: the ratio between flow and capacity. Roads with a high I/C-ratio are relatively vulnerable, because there is not much spare capacity available. The difference with the previous indicator is that the previous indicator is more oriented towards motorways than this indicator, because regional and local roads have a lower capacity, and therewith have a higher chance to have a little spare capacity in absolute terms. The I/C-ratio is comparable to the speed ratio, because both focus on the level of congestion. Because, in practice, the intensities first increase and thereafter decrease when it becomes more congested (which implies that a low intensity can point to both congested and uncongested traffic states), the speed ratio is a better indicator when dynamic models or real data are used. The speed ratio does not have this problem.

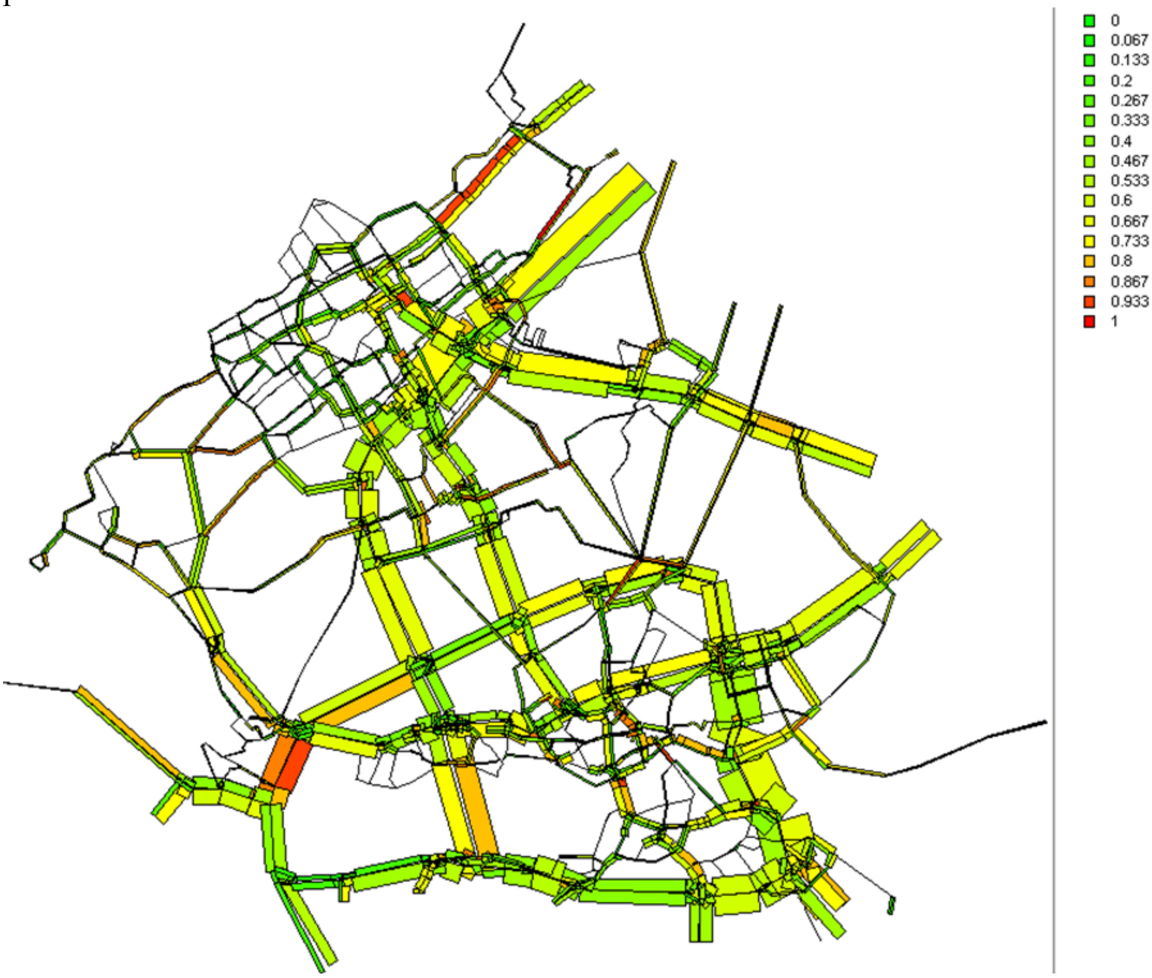


Figure B.3: I/C -ratio, based on LMS

<u>4: Number of OD-pairs that make use of a link:</u> the number of OD-pairs that make use of a link based on the shortest paths is an indicator that comes from graph theory and refers to the centrality/betweenness of a link. If many OD-pairs make use of a link, disturbances lead to delays for many travellers from different parts of the network. Figure B.4 shows this indicator. In order to compute this indicator, an OD-matrix, a network, and an all-or-nothing assignment model is needed.

Alternatives for this indicator can be found by using an equilibrium assignment instead of an all-or-nothing assignment, and by taking the volumes on the link into account to show that roads that are used by the same number of paths are more vulnerable if more traffic uses the link.

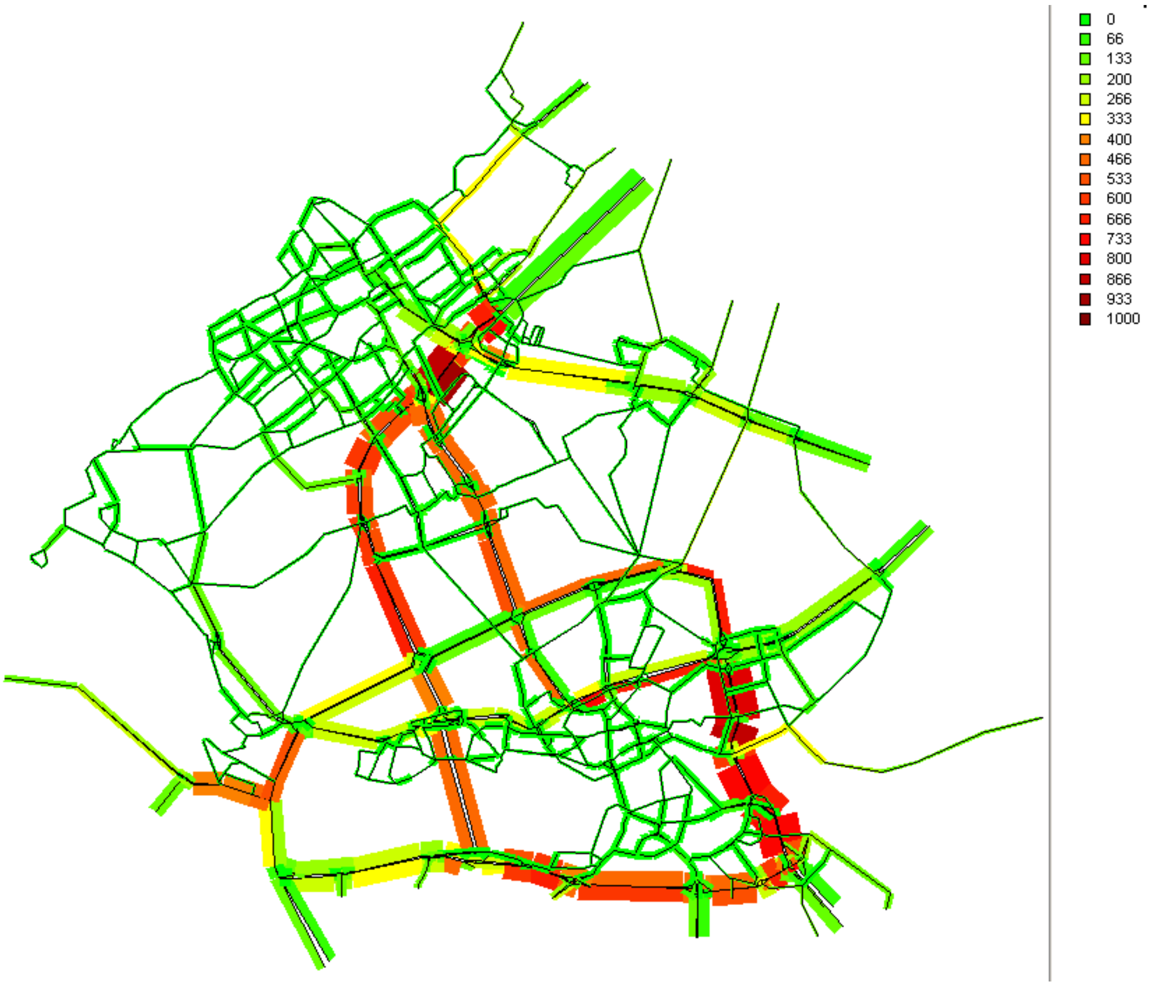


Figure B.4: Number of OD-relations that use a link

5: The availability and quality of alternative routes: this indicator shows to what extent alternative routes are capable of functioning as backup options for the road itself. This indicator is described in section 3.5. This indicator has a variant that does not consider flows and one that does. The one that does not is easier to compute, because that can be done solely based on the network structure. The other one requires an assignment.

There are many variations possible to this indicator. In Figure B.5, a variation is shown in which only 50% of the spare capacity of alternative routes for the motorways are taken into account, because it is assumed that travellers that make use of the motorways prefer motorways over lower level roads. Other assumptions with respect to the road type could be made as well. Furthermore, the distance parameter (the way in which distances are measured – via the network or as the crow flies) and the angle parameter can be varied. Unfortunately, there is not enough data available about the driver's behaviour during incidents that can be used to calibrate the parameters of this indicator.

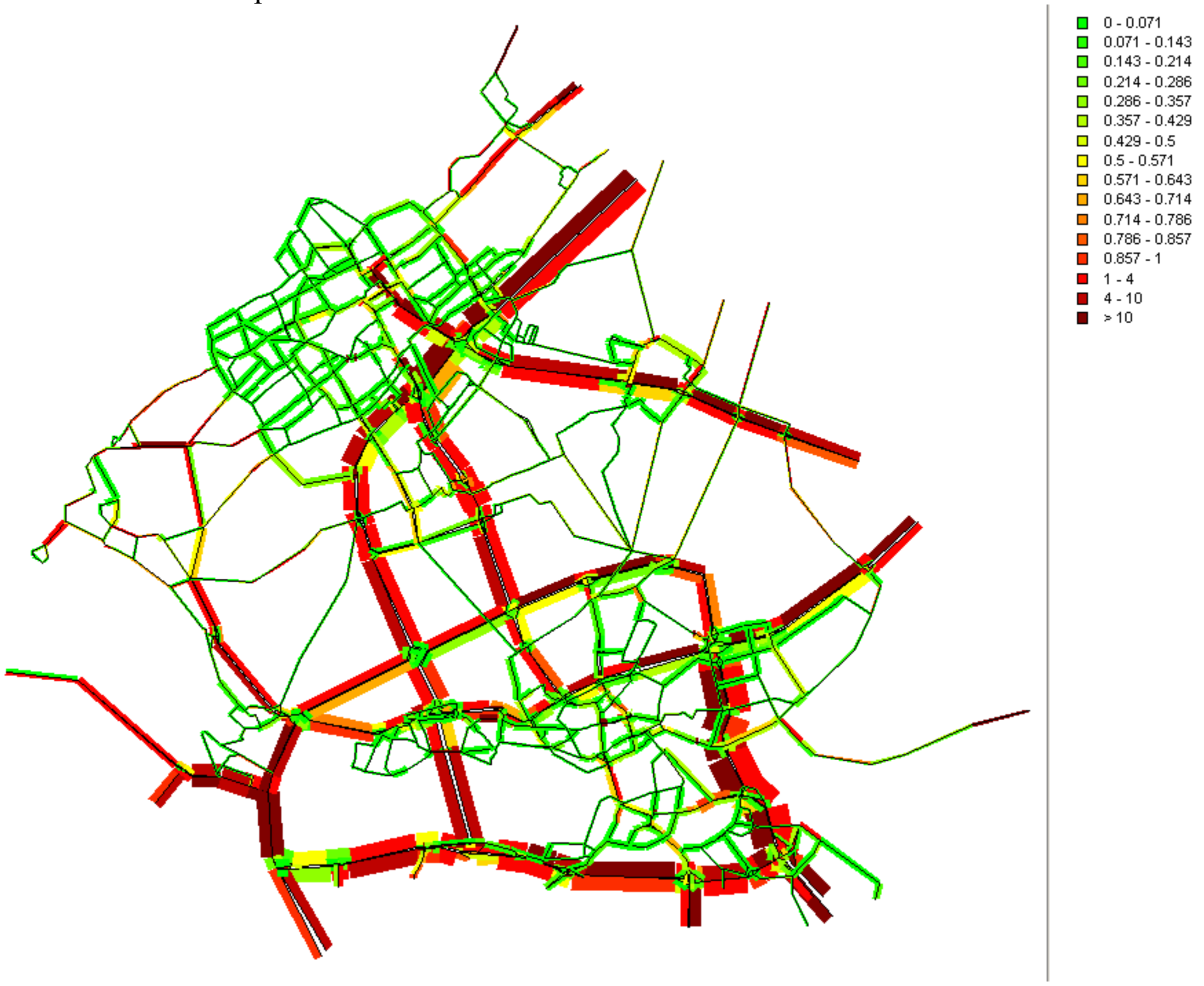


Figure B.5: Ratio between flow and spare capacity on alternative routes, based on LMS (beta = 0.85)

Finally, care should be taken when the spare capacity is computed by subtracting the intensities from the capacities. As explained before, these intensities might be misleading. For instance, if a road is completely congested, the intensity is 0 pcu/hour, which indicates that there is a lot of spare capacity in the network, whereas, in practice, there is not even space for one extra car. Therefore, a correction should be applied. This can, for instance, be done by setting the spare capacity to 0 pcu/hour if the speed ratio drops below a certain threshold.

<u>6: Number of by spillback effects affected vehicles:</u> the number of vehicles that are affected upstream of the link where an incident occurs is approximated by considering all vehicles that drive upstream of the link and can reach the link within two minutes of free-flow travel time (the two minutes could also be another number) (collection of links *AA*). The intensities v_a are corrected by the spare capacity. The idea behind this is that the spare capacity can be used to bypass an incident. Of course, this is useful only when the incident does not block the complete road. We chose this indicator, since we preferred to use an indicator that is independent of the reduction factor of incidents. The indicator is shown in equation A.1. In this equation, $speed_a$ is the realized speed and L_a is the link length.

$$Affectedvehicles_i = \sum_{a \in AA} \max\{0, (v_a - sparecap_a) / speed_a\} * L_a \tag{A.1}$$

This indicator is an approximation for the spillback effects that occur as a result of incidents. It is a simplification of effects that can only be modelled accurately by dynamic models. One of the simplifications is that this indicator implicitly assumes that queues do not spill back more than four kilometres upstream of an incident. This indicator finds the links at which an incident leads quickly to a lot of congestion: the links just downstream of intersections of which the incoming links have high intensities and low spare capacities. The advantage of this indicator is that it can be directly computed based on the outcomes of a single static or dynamic assignment without simulating many incidents.

Figure B.6: Number of vehicles affected by spillback, based on LMS

7: Extra vehicle kilometres travelled as a result of link closure: this indicator shows how much extra vehicle kilometres need to be driven in the complete network when a link is closed. This indicator is computed by doing a static user equilibrium assignment for each closed link and comparing the total vehicle kilometres driven to the situation without incidents. The computation of this model requires an OD-matrix, a network, and an assignment model. Variation of this indicator can be made by using an all-or-nothing assignment model, or even a dynamic traffic assignment model (but that takes a lot of time), or by looking at the travel times or the generalized costs instead of the vehicle kilometres driven. The last is difficult, because travel times tend to get very high at high I/C-ratios, which makes it difficult to get a good estimate of the travel times.

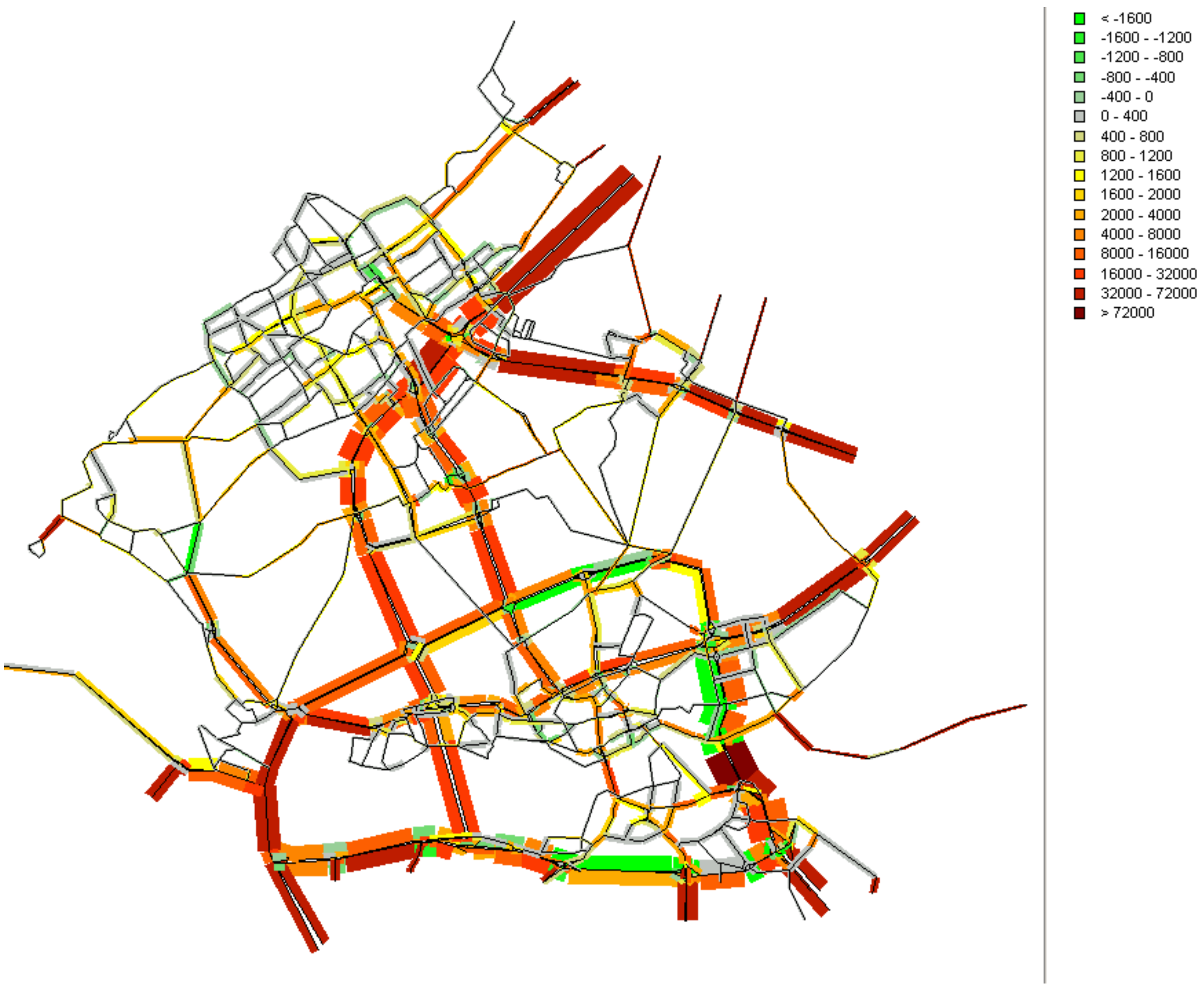


Figure B.7: Extra vehicle kilometres as a result of a link closure

<u>8: Travel time losses as a result of incidents computed with Indy and the MIC-module</u>

This is the indicator and the computation method that are described in chapter 3 and chapter 4. In

Figure B.8 the results are shown for the situation in which the approximation algorithm for alternative routes is not used (indicator 8a). In Figure B.9 the results are shown for the situation in which the approximation algorithm for alternative routes is used (indicator 8b). In these examples, only one incident type is used, in which the capacity is reduced by 50% during a ½ hour in the morning peak.

Figure B.8: Vehicle loss hours caused by incidents, Indy-MIC

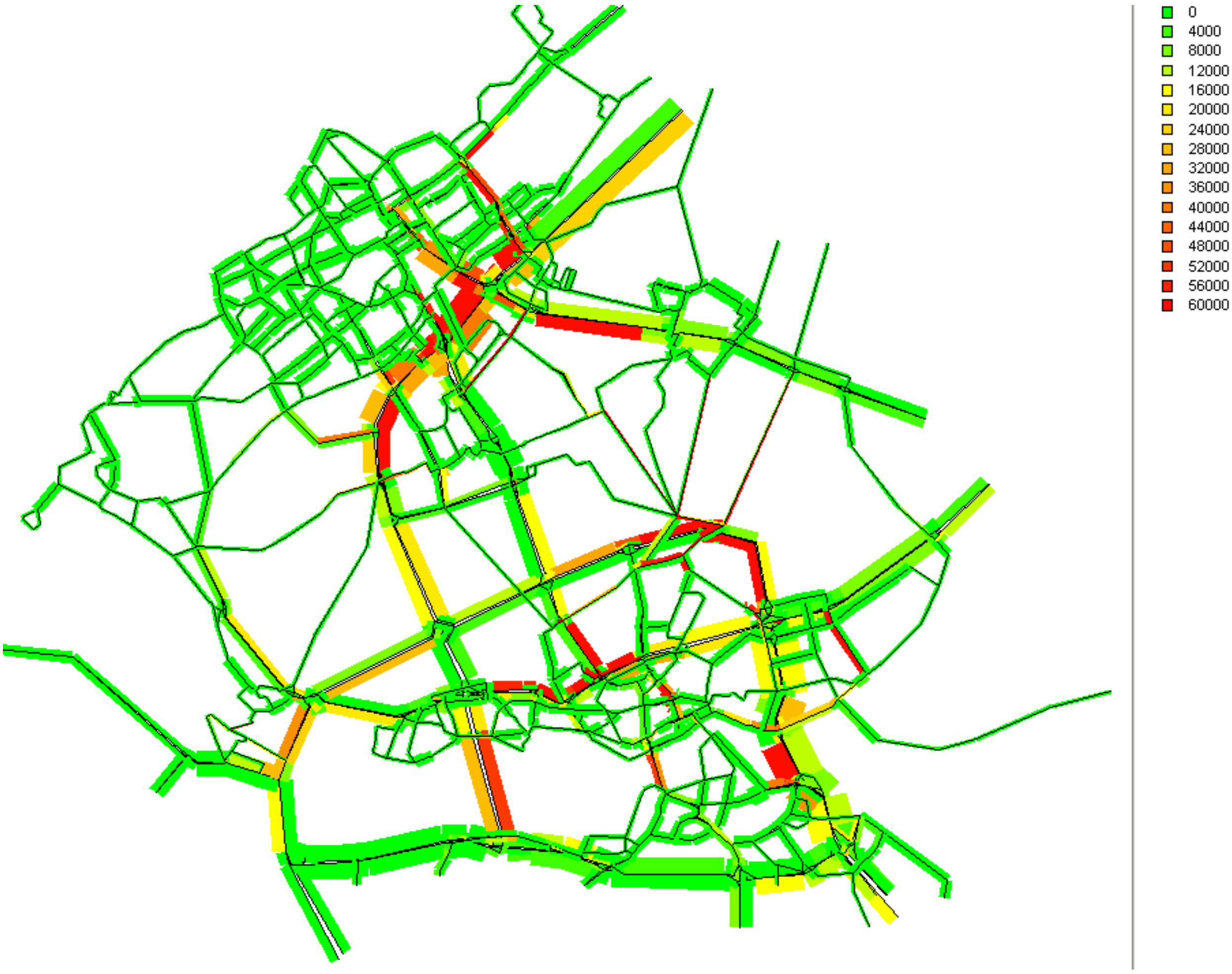


Figure B.9: Vehicle loss hours caused by incidents, Indy-MIC-alt routes

Correlation among indicators.

The question that remains is to what extent the first seven simple indicators can approximate the 8th complex and most accurate indicator. In the table below, the rank correlations among the indicators is shown. The closer the absolute rank correlation is to 1, the better the first indicator correlates to the second indicator. It appears that indicators 3 and 4 give the best approximations of the complex indicator. However, the correlations are so low that it can be concluded that the first seven indicators by themselves do not have enough explanatory power to compute the vulnerability of the links (if we assume that the 8th indicator is correct). Besides that, it is clear that indicators 3 and 4 are so simple that it is very unlikely that they fully describe the concept of robustness. However, this does not exclude the possibility that a combination of indicators can describe the concept of robustness in a reasonable way. A combined regression analysis and decision tree approach did not result in a good approximation either. This leads to the conclusion that it is necessary to use a dynamic traffic assignment model and a method like the MIC module to accurately compute the vulnerability of links. Since this is not always possible in practice, due to limitations in time, budget, or network size, a fallback method is proposed in Appendix B that combines a few indicators that refer to the elements of robustness as described in section 3.4 and should, therefore, from a descriptive point of view, give good insights into the robustness of a network.

Table B.1: Rank correlation coefficient among the different indicators

Indicator	2	3	4	5	6	7	8a	8b
2	1.00							
3	-0.40	1.00						
4	0.09	0.64	1.00					
5	-0.13	0.84	0.70	1.00				
6	0.10	0.29	0.33	0.28	1.00			
7	-0.07	0.50	0.52	0.49	0.12	1.00		
8a	-0.08	0.65	0.63	0.59	0.46	0.42	1.00	
8b	-0.08	0.63	0.62	0.58	0.44	0.42	0.98	1.00

Appendix C: Overview of methods used in the Netherlands for robustness and reliability analyses

Below a short description is given of the methods that are used in the Netherlands to quantify robustness and reliability. Note that this list may not be complete. Additionally, other static or dynamic models can be made suitable for modelling robustness. By means of example, in Appendix C a comparison is made between two dynamic traffic assignment models: Indy and Dynameq.

Rules of thumb:

- Properties: no model.
- Advantages:
    - No need for lengthy and complex model calculations (once the indexes are known).
    - Unambiguous.
- Disadvantages:
    - Difficult to derive.
    - Difficult to summarise all aspects of robustness in a single index.
    - Difficult to take into account different types of disruptions (different accidents, weather conditions, etc.); (duration, location, seriousness, etc.); (traffic conditions).
- Explanation: rules of thumb form the quickest but least accurate method. For reliability, for example, often 25% of the travel time is stated as benefits in cost-benefit analyses in the Netherlands. This is a simple and inaccurate rule of thumb that makes no distinction between network properties and network occupancy. Such an index does not yet exist for the effects of major disruptions. An index for robustness could maybe be derived based on accident analysis. To make this possible, ideally for different types of incidents (or actually also other sorts of disruptions, such as heavy rain showers), the characteristics of the incident (duration, location, etc.), the traffic conditions (intensity, speed, etc.), other circumstances (e.g. the weather), and the effects (extra vehicle kilometres, extra travel

time, etc.) should be known. A lot of data are available, but whether this is sufficient to be able to derive accurate rules of thumb is not certain. A second option is to use model simulations to derive a rule of thumb, rather than to use data. The advantage of this is that all types of disruptions can be simulated. This only needs to be done once. Thereafter, the rule of thumb could be used (if an accurate rule of thumb is found). The choice of model is of importance here. An important aspect is the degree to which the model can effectively determine the effect of the simulated disruption.

Method applied in the National Market and Capacity Analysis (NMCA) (Snelder et al., 2010a)

- Properties: based on single user equilibrium assignment with the national model system (LMS).
- Advantages:
    - No additional model calculations with, for instance, a dynamic traffic assignment model are needed.
    - Matches with the elements that make a network robust. This gives a direct indication of which measures are likely to be the most beneficial.
    - Can be applied to large networks of the size of the Netherlands, or maybe even larger.
- Disadvantages:
    - Simplified calculation methods used, which makes the indicator less accurate.
    - Looks at relative changes. These relative changes will, therefore, need to be linked to an absolute value.
- Explanation: in this method, a vulnerability analysis is carried out in order to get a first reasonable accurate indication of where which robustness measures have the highest potential. The method was required to be applicable to large size networks. It should, therefore, have a short computation time. The method was also required to give an as accurate as possible description of robustness. A choice was made combine four indicators that refer mostly to redundancy and compartmentalization:
    (1) The speed ratio (similar to indicator 3 of Appendix A).
    (2) The availability and quality of alternative routes (indicator 5 of Appendix A).
    (3) Number of vehicles affected by spillback effects (indicator 6 of Appendix A).
    (4) The chance of disturbances.

  These four indicators determine together the robustness of a road section and all four can be directly computed based on the outcomes of the LMS. The four indicators are combined to a vulnerability score by the following formula: Vulnerability = (4)*(3)*(2)/(1).

SMARA

- Properties: static macroscopic model, original route choice or new equilibrium, vertical traffic queue, no intersection modelling.
- Advantages:
    - Determines the entire travel time distribution.
    - Determines the chance and effect, so that evaluation is relatively simple.
    - Original route choice and new equilibrium possible.
    - Can, in principle, be applied to every transport model, if an origin-destination matrix and network are available.
    - Can be used to look at reliability and robustness in a consistent manner.
- Disadvantages:
    - Static: this implies that spillback effects and time dynamics are not considered.

  - o Calculation time relatively long: between several hours and a day for the whole of the Netherlands (depending on the route choice assumptions).
- Explanation: This model is designed to determine the travel time. With a static equilibrium assignment model, an average morning peak period, non-peak period, or evening peak period is first calculated. Then the intensity and capacity are varied using Monte Carlo simulation. This is done based on static distributions of both the demand (variation as a result of seasonal influences and events) and the supply (variation as a result of different weather conditions, accidents, roadworks). The effect of the variation in supply and demand can be determined in two ways: with the original route choice, and with a new equilibrium. Although this model is designed to look at the reliability of travel time, this model can also be used to look at the effect of the 5% biggest disruptions. Its disadvantage is that it is a static model, and congestion spillback is therefore not optimally taken into account. The advantage is that reliability and robustness can be looked at in a consistent fashion. In (Meeuwissen et al, 2004) and Appendix D, the model is described in more detail.

Robustness Scanner
- Properties: static macroscopic model, original route choice + assumptions regarding shift in routes, vertical traffic queue, no intersection modelling
- Advantages:
  - o Easy to link to LMS and NRM (national regional modelling systems).
  - o Static and therefore reasonably quick.
- Disadvantages:
  - o Assumptions regarding use of alternative routes: all the road users that experience a delay as a result of capacity restriction choose another route (average = 53%). This results in there being no more delays on the original route.
  - o Static -> no congestion spillback.
- Explanation: the Robustness Scanner (Tamminga et al., 2005) uses a static traffic model and, therefore, does not describe the effects of congestion spillback. The method first identifies the vulnerable road sections, based on the chance of an accident and the amount of traffic that is delayed by an incident. Then it is ascertained which road users take which alternative routes. The advantages of this method are related to the use of a static model that can generate results quickly and easily (Kraaijeveld, 2008). The model makes an important assumption about detour behaviour: it assumes that all road users that would experience delay as a result of the capacity restriction choose another route. On average, in this model, it turns out that 53% of the road users choose an alternative route (Kraaijeveld, 2008). In the model, the road users that do drive via the incident location will not experience any delay, because the remaining capacity is just enough for the number of road users using that route.

Tampère et al. (2006) use a similar approach, with the difference that they use a dynamic assignment from the start, and that they use different selection criteria for possible vulnerable links.

Indy
- Properties: dynamic macroscopic model, original route choice, possibly with assumptions concerning shifts to alternative routes or new equilibrium, horizontal traffic queue with first-order wave theory LTM, no intersection modelling.

- Advantages:
    - o Different forms of route choice behaviour are possible (with the exception of en-route route choice).
    - o Takes congestion spillback into account -> congestion is located in the right place.
    - o Possible to calculate different types of disruptions (including variation in duration, location, and capacity reduction of incidents).
    - o Takes into account traffic flow over time.
- Disadvantages:
    - o Long calculation time -> relatively few disruptions can be calculated. The extension with Indy, Indy-MIC (Marginal Incident Computation), is an exception to this (see section 4.3).
    - o Requires a lot of data to calibrate the model.
- Explanation: Using this model, it is possible to determine the effects of an incident in various ways:
    a. Equilibrium + 1 iteration with capacity restriction: here, in the case of an incident, everyone chooses the same route they would have chosen in 'normal' circumstances (user equilibrium). No one, therefore, changes his route (neither does anyone change means of transport or departure time).
    b. Indy equilibrium + shifting routes + 1 iteration: here, a predefined percentage of road users that experience more than a pre-specified percentage of delay as a result of an incident choose another route.
    c. Indy equilibrium + MIC (Corthout et al., 2009): here, no one changes his or her route. The effect of an incident is analytically determined in the upstream direction. Based on a situation of equilibrium, it is possible to determine quickly the effect of an incident (less than 0.1% of a complete simulation), because no new model run is required to be performed. This makes it possible therefore to calculate a couple of thousand incidents at different locations and with different characteristics in a few minutes up to an hour (depending on the network size and level of congestion). A disadvantage of this method is that downstream effects cannot be determined. A second disadvantage is that calculations are performed only with original routes. Alternative routes are not taken into account.
    d. Indy new equilibrium: this variation assumes that everyone is completely informed about the incident and is able to adapt his or her route choice accordingly. A totally new equilibrium is created. In practice, this will not happen, but it provides extra insight into the range in which the effect of the incident will lie.

Marple-e

- Properties: dynamic macroscopic model, en-route route choice, horizontal traffic queue without first-order wave theory, including intersection modelling.
- Advantages:
    - o En-route route choice is possible (but also demands assumptions concerning the amount of information that is available and the response to that information).
    - o Takes congestion spillback into account.
    - o Possible to calculate different types of disruptions (including variations in duration, location, and capacity reduction).
    - o Takes into account traffic flow over time.

- Disadvantages:
    - Long calculation time.
    - Requires a lot of data to calibrate the model.
    - Less detailed congestion modelling than in Indy.
- Explanation: As with Indy, Marple can be used to determine the effect of an incident based on a new equilibrium. Additionally, Marple allows the route choice to be modified during a journey, based on available information. This form most closely approaches the route choice of people in practice, but demands assumptions concerning the moment at which information becomes available, and the response to this information.

Appendix D: Comparison between Indy and Dynameq

In the past few years, many dynamic models have been developed, and they are being applied more and more often. Traditionally, these models are classified as microscopic, mesoscopic, and macroscopic. These models differ in the level of detail in the network (e.g. with or without explicit lane modelling) and the level on which the traffic dynamics are modelled (vehicle level, flow level, or a combination of both). All models aim to reproduce the traffic situation in the best possible way within their intended scope. It is interesting to make a comparison among models, to learn how differences in model specification lead to different outcomes. This appendix summarizes the results, which are extensively described by Snelder (2009b). The intention is not to compare all dynamic models with each other, but to show the differences between Indy and Dynameq. Indy is a macroscopic dynamic traffic model that has been developed in the Netherlands and Belgium by TNO, the Delft University of Technology, and the Catholic University of Leuven. Dynameq is based on a traffic simulation model that was designed to produce reasonably accurate results with a minimum number of parameters and a minimum of computational effort (Astarita et al., 2001) (Mahut, 2000). However, the underlying structure of the model has more in common with microscopic than with mesoscopic approaches, as it is designed to capture the effects of car following, lane changing, and gap acceptance. Dynameq was developed in Canada by the University of Montreal and Inro.

The aim of the comparison is three fold:

- *Identify the differences in the specifications of both models*: Dynameq and Indy are two different dynamic assignment models. One of the most important differences is that the first is lane based and the second link based. A comparison of differences in specifications will indicate what the differences exactly are.
- *Find the differences in computational efficiency*: since both models are specified in a different way, they converge in different ways to an equilibrium. Besides that, the

computation time will differ per iteration, because different network loading algorithms are used. The question is how large these differences are, and what causes them.
- *Show the differences in the outcomes of models and their practical implications*: "What goes wrong if no lanes are modelled?"

First, a comparison is made between the model specifications, in order to get an understanding of the similarities and differences between the models. The comparison of the specifications of Dynameq and Indy showed that there are several important similarities and differences. The models are comparable in the sense that both are equilibrium models in which paths are generated, path choice plays a role, and dynamic network loading takes place in an iterative process. The newest and most accurate network loading model in Indy (LTM) works according to Newell's kinematic wave theory, and so does the network loading algorithm in Dynameq. The main differences are:
- Dynameq generates paths in the first iterations of the simulation, whereas Indy generates paths before the simulation starts.
- Dynameq has a deterministic path choice, whereas Indy has a stochastic path choice.
- Dynameq is lane based, whereas Indy is linked based.
- Dynameq models individual vehicles, whereas Indy models aggregate flows (per path). This enables Dynameq to model gap acceptance at intersections and lane changing behaviour, which cannot be done by Indy.
- Dynameq has a more detailed intersection model than Indy. It can deal with traffic signals and priority flows, whereas Indy can only approximate this by introducing a maximum link outflow capacity.
- Dynameq is event based, whereas Indy works with fixed time steps.

We ran both models on three networks with several scenarios to show how the above mentioned differences influence the model outcomes. The first test network was a network in which delays at intersections were excluded and, in principle, two equal paths are available for the single OD-pair. To get an idea of what the network looks like, the network of the first scenario is shown below.

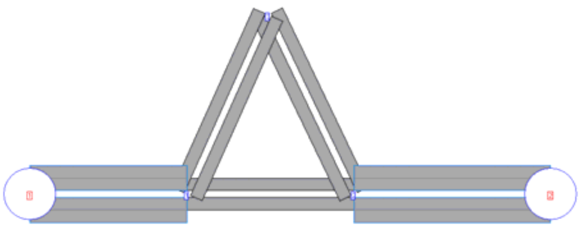


Figure D.1: Network test case 1

It was shown that:
- The stochastic path choice of Indy in combination with generating trips can have advantages over deterministic path choice, since already in the first iteration the traffic is spread over all the pre-specified paths. In this way, Indy converges faster to an equilibrium than Dynameq. This general conclusion is illustrated by the fact that in the first scenario Dynameq does not find a second path, because there is no congestion on the first path. Therefore, in Dynameq the capacity of the network is not used fully, which results in the fact that it takes Dynameq twice as long to get all the traffic over the network. It must be said that this is an extreme example, because in larger networks links will be used by multiple paths from multiple OD-pairs, which always causes some delays

and, therefor,e extra paths to be generated. The differences in path choice also became clear in the second scenario, where Indy finds a perfect equilibrium in which 50% of the traffic uses each path. The spread in Dynameq after the first 30 iterations is 53%-47%.

- In the second scenario, Dynameq does find two paths, which makes the results of Dynameq and Indy more comparable. In fact, if Indy uses the paths of Dynameq, the results are almost identical. This shows that both network loading models are the same if delays at intersections and lane changing behaviour do not play a role, as was to be expected based on the model specifications.
- In this test network the links are very long, which allows for a high time step in Indy. The time step could even be set to 180 seconds. However, a time step of 60 seconds is used. Even with this time step, Indy is about 4 times as fast as Dynameq. This is probably caused by the fact that Indy is a macroscopic model and, therefore, does not have to keep track of individual vehicles and the way in which they behave. On the other hand, if the links, or even only one of the links, would have been shorter, a smaller time step had to be chosen, which would increase the computation time of Indy. In that case Dynameq becomes faster than Indy, which is illustrated better on the Bakersfield network (shown below). This shows the consequences of having event based or time step based models.

The second test network is a network with four zones and flows between all zones. There is only one path between each OD-pair. Therefore, path choice does not play a role. The network is shown below.

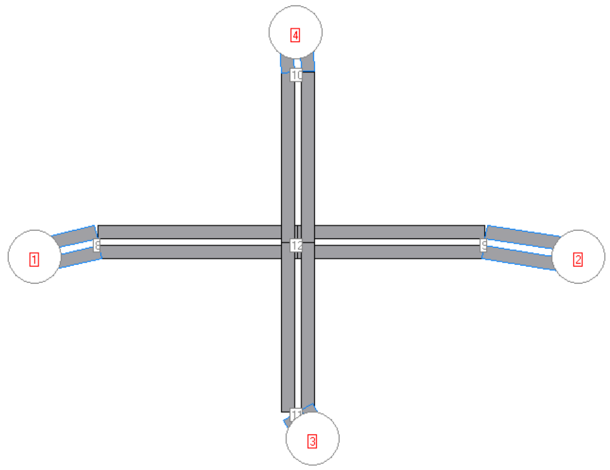


Figure D.2: Network scenario 1 and 2, test case 2

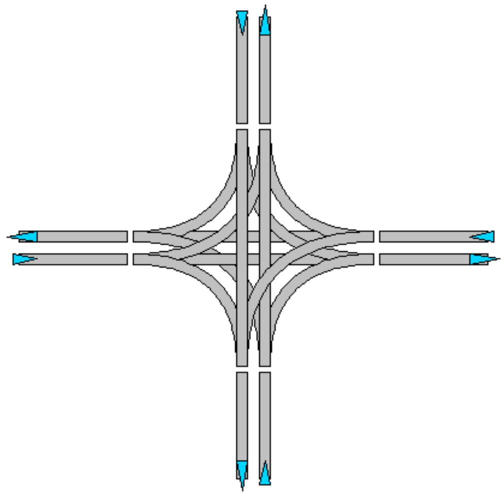


Figure D.3: Intersection, test case 2

This example emphasizes the differences between both models at intersections:

- In the first scenario, the intersection is unsignalized. In this case the differences in model outcomes are very large. The total travel time in Indy is about twice as low as in Dynameq, and the network is empty almost an hour earlier. The explanation for this is that Indy does not consider delays at intersections that are caused by waiting for gaps that are needed to cross an intersection. In Indy, cars can virtually drive over each other, which is of course not realistic. The fact that the total travel time is twice as low shows that the delays at intersections can add up to a substantial amount. On the other hand, the flows in this network are high. If there is less traffic in the network, the delays at the intersections also reduce.

- The two scenarios with signalized intersections showed that it is possible to model signalized intersections by outflow constraints, because the results of Indy and Dynameq are exactly the same for both scenarios. This is, however, possible only in the situation in which there are no conflicts in lane usage, because in those situations the outflow is a result of the arrival pattern of traffic on the intersection. Therefore, the maximum outflow cannot be computed based on the link capacity, green times, and cycle lengths. For those situations, an approximate outflow capacity has to be found.

Finally, both models are compared on a realistic network: the Bakersfield network. This network is shown below. It has 36 signalized intersections (black nodes).

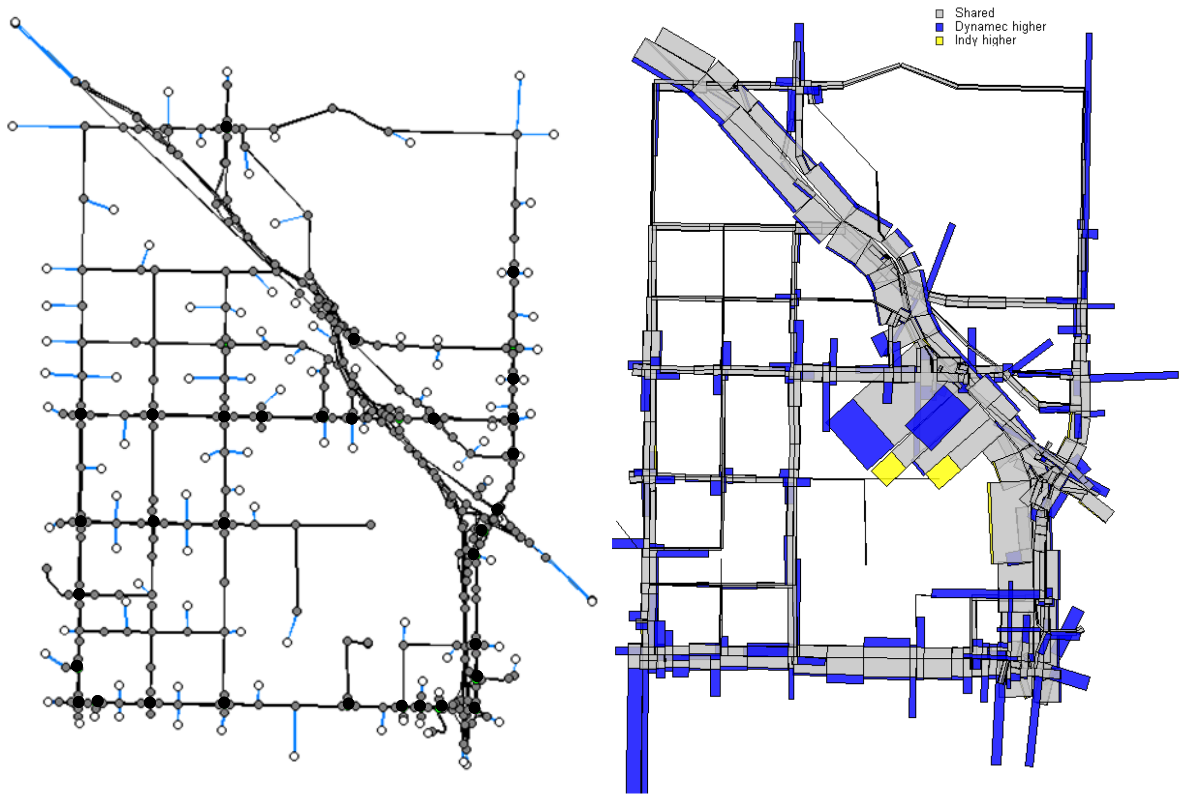


Figure D.4: Bakersfield network (left) and differences in model outcomes (density) in the case in which Dynameq is run with signalized intersections and Indy is run with the paths of Dynameq

All the differences mentioned above become clearer in this network. Besides that, some additional differences become clear:

- In the Bakersfield network, congestion occurs on the motorway due to lane changing behaviour. This congestion is recognized by Dynameq. Since Indy does not model the described lane changing behaviour, it does not find the congestion on the motorway. An approximate outflow restriction is imposed in Indy on the link upstream of the motorway junction where the congestion starts. However, this appears not to capture the dynamics in the traffic flow caused by lane changing behaviour completely.
- Dynameq is run with signalized and unsignalized intersections, and Indy is run twice with the path and path flows of these runs and once with its own paths and path choice. The average speeds in the Dynameq outcomes of the signalized network are about 12 km/hour lower than in all the other model runs ─ also in the model run in which Indy uses the paths

and path flows produced by the same Dynameq run. This is explained by the fact that Indy does not model the delays at the intersections. In the case in which Dynameq is run with unsignalized intersections, the average speeds come very close to the average speeds computed by Indy.

- The results of the three model runs with Indy are very close to each other. This suggests that the equilibrium that is found by Indy is close to the equilibrium that is found by Dynameq. It is remarkable that the equilibrium run of Indy is in fact closer to the equilibrium run of Dynameq with signalized intersections than the equilibrium run with Dynameq with unsignalized intersections, because Indy does not use signals. From the total vehicle kilometres driven, it can be seen that the slightly shorter paths (shorter in distance) are chosen if the signals are not used. This suggests that the signals on the shorter paths cause delays that make travellers choose longer (in distance) paths.
- The development of the average speed, the total travel time, and the average lane density in the network over time for the five different model runs is more or less the same over time in all Dynameq and Indy runs. This indicates that, despite all the before mentioned differences, the model outcomes also show a lot of similarities.
- A comparison over time per link showed that there is a very high correlation between Dynameq and Indy if the inflow and outflow are compared. In the case in which Indy uses the paths of Dynameq, this is logical. Although, even in that case, delays can cause differences in inflow and outflow over time (but not over links). The fact that the inflow and outflow have an R^2 above 0.95 indicates that the equilibrium route choice of both models does not differ much. However, the R^2 of the speeds is low (0.24). For a large part, this is caused by links at intersections. The R^2 of the density is even worse (0.18 or 0.19) than the speed in case of signalized intersections. However, in the case of unsignalized intersections, the R^2 goes up to 0.67 and 0.76. This emphasizes the need to model delays at intersections explicitly. The right hand figure of Figure D.4 shows that the differences are indeed, for a large part, located near intersections.
- Dynameq converges slower and keeps higher gaps in the later iterations. The first cause for this is that Dynameq models more delays (and therefore bigger differences between paths). This hypothesis is strengthened by the fact that the gaps in the case without signals are already much lower (below 10%) than in the case with signals, in which the gaps go up to 70%. A second cause is that Dynameq hasn't found all the paths yet in the first 10 iterations, which causes the high gaps in the first iterations. Finally, the fact that Dynameq uses a deterministic assignment results in the fact that the traffic is less spread over all available paths.
- In the equilibrium run with Indy, a time step of 5 seconds was used. This time step is slightly larger than the free flow link travel time of 141 links. The smallest free flow travel time is 1.3 seconds. This implies that, for these links, the link lengths had to be extended during the simulation in such a way that they have a free-flow travel time of 5 seconds. With this time step, the computation time per iteration is still almost 10 times as high as in Dynameq, which is likely to be caused by the fact that Dynameq is event based, and many links have a much higher free flow travel time than 5 seconds. A second explanation might be the number of paths used. In total, there are 2988 OD- pairs with a demand larger than 0. Indy generated 8399 paths, and there is flow on all these paths. Dynameq generated 19692 paths, but there is flow on only 6200 paths. Since the computation time of Indy depends on the number of paths that are used in the evaluation phase, this could also be an explanation for the longer computation times. In the comparison in which the Dynameq paths are used, a time step of 1 second is used in Indy. Therefore, in these runs the link lengths did not have to be extended. For these runs, only one iteration was needed. This iteration took 37 minutes, which is more than 5 times (9.4) higher than the case with

a time step of 5 seconds. The explanation for this might well be the number of paths. Although only 6200 paths are used, all 19692 paths are considered in Indy.

The above mentioned differences lead to the following recommendations for Dynameq and Indy:

Dynameq:

- Generating the paths before the simulation starts or storing the paths of a previous model run can reduce the number of iterations that are needed to reach an equilibrium and, therefore, decrease the computation time. It might be worthwhile to investigate if this is possible.
- Stochastic route choice instead of deterministic route choice could lead to faster convergence as well. However, changing this has bigger implications, since it is a fundamental change in the assumed route choice behaviour.

Indy:

- The modelling of delays at intersections can be improved. For the level on which Indy is currently mostly used (high level, with mainly motorways) this is less important than for cases in which local networks are used. However, it could still be large improvement. This requires more input data, but could make the calibration easier in the end. A first improvement, which is probably relatively easy, would be the introduction of an option to include maximum outflow constraints per movement instead of per link. This prevents links having to be split into three separate links to replicate the structure of a node. Other improvements that are needed to include the gap acceptance principle might be investigated as well, which is already being done by the University of Leuven.
- Including lane changing behaviour is not possible in Indy. It might be worthwhile to investigate how this behaviour could be approximated.
- A practical suggestion is not to use the adjusted link capacities from networks that are used in static assignment models and adjust those capacities a bit further in the calibration, but to reset the capacities to the level that is to be expected based on free-flow speeds, average vehicle lengths, and a response time. In the calibration, the maximum outflow capacities can then better be adjusted instead of the capacities themselves.
- Investigate the possibility and the gains in computation time of switching from a time step based network loading model to an event based model. This is probably relatively easy, since the University of Leuven already has an event-based version of LTM.
- Remove paths that are not used or barely used during the simulation, in order to reduce the memory usage and increase the computation speed.

Appendix E: Monte Carlo simulation model: SMARA

In this Appendix, the model SMARA is explained. SMARA calculates the travel time distribution by means of Monte Carlo simulation. The model predicts the bandwidths of travel times from door to door and on the link level by randomly selecting different demand and supply situations. A statistical analysis has been carried out to get the distribution for the following four situations:

- Demand, generic: influence of seasons.
- Demand, specific: events
- Capacity, generic: weather.
- Capacity, specific: accidents and road maintenance.

Factors and probabilities are determined for these four situations by means of statistical analysis. The factors indicate how much the demand differs from the nominal demand and how much the capacity of the road differs from the capacity under normal circumstances. Furthermore, for different weather conditions there are speed factors as well. The factors are divided into several classes. The probabilities show how often the different classes occur. Below, it is explained how the statistics are determined. The statistical analysis is based on data from before 2004. Of course, an update could be made. However, it is likely that these statistics will not change much over time.

The statistics for the influence of seasons are based on large scale surveys that were carried out in the Netherlands (OVG/MON) in 1999, 2000, and 2001, in which people were asked to keep a diary of all the trips that they made within a certain period. Three years were used, in order to increase the reliability of the data. All trips were selected that were made on weekdays excluding the public holidays in the periods 7.00 - 9.00 hour and 16.00 - 18.00 hour for the peak period, and in the periods 6.00 – 7.00 hour, 9.00 - 16.00 hour, and 18.00 - 24.00 hour for the off-peak period. For the peak and off-peak periods the total number of trips per day was determined. Twenty classes were distinguished in such a way that the horizontal distance between the classes (factors number of trips) is equal. The nominal demand is the median of the daily demand. The variations in demand caused by seasonality are applied to the complete origin-destination matrix.

The statistics for events are based on the top 45 events in the Netherlands. For these events, the number of cars per hour and the probability of the event (related to the number of opening days) are determined. Meeuwissen et al. (2004) explain in detail how this is done. The number of visitors or cars per event is not uniformly registered for all events. Therefore, for each event different conversions had to be made, from yearly or daily numbers of visitors to numbers of visitors per hour, and from numbers of visitors to numbers of cars. Furthermore, the arrival times had to be determined, since the arrival times are not spread uniformly over time. For the events for which the arrival time distribution was not known, it was assumed that all visitors arrive within a time frame of two hours. Only the events with a probability lower than 40% that occur on weekdays are selected. The other events either occur that often that they are already included in the nominal situation, or they occur on weekend days and are therefore not relevant for our weekday analysis. For each event, it is known in which region it takes place. The origins are spread in proportion to the regular demand in that region.

The statistics for the five different weather conditions (fog, rain + darkness, rain + daylight, dry + darkness, dry + daylight) are based on the weather registration of the Royal Netherlands Meteorological Institute (KNMI), known capacity reduction factors for rain and darkness, and some assumptions that had to be made because of a lack of data. It is, for instance, assumed that the remaining capacity in case of fog is slightly less than that in case of rain and darkness (0.89 versus 0.92). Furthermore, it is assumed that the probability of rain is equally spread over the day. Since we only had data about the motorways, a distinction between road types is not made. Finally, assumptions had to be made about the speed factors, which are shown in Table E.1 The nominal capacity is the capacity under average weather conditions. Therefore, the capacity factor of dry weather during daytime is higher than 1. The variations in capacity caused by different weather conditions are applied to all the links in the network at the same time.

The statistics for incidents are based on different incident registration databases, and on the number of vehicle kilometres driven per road type. Conversions had to be made from total numbers of incidents to the number of incidents per road type and number of lanes per road type. Meeuwissen et al. (2004) explain in detail how this is done. We distinguished four incident types: car breakdown, accident that blocks the hard shoulder, accident that blocks one or more lanes, and rubbernecking. There is no registration of congestion caused by rubbernecking. Therefore, it is assumed that the chance of rubbernecking is equal to the chance of both accident types (in general, car breakdowns do not cause rubbernecking). Furthermore, it is assumed that the capacity reduction is 50% of the weighted capacity reduction of both accident types. The variations in capacity caused by incidents are separately determined for each link in the network.

The statistics for roadworks are based on roadwork registration databases. For 17 different types of roadworks, the capacity factors and the probabilities of occurrence were determined for the peak and off-peak periods and for different road types. The variations in capacity caused by roadworks are separately determined for each link in the network.

Table E.1, Table E.2, and Table E.3 present the factors and probabilities for incidents and different weather conditions. Further data can be found in (Meeuwissen et al., 2004).

Table E.1: Factors and probabilities for different weather conditions

	Capacity factor	Speed Factor	Probability Peak	Probability Off-peak
Fog	0.89	0.50	1.81%	1.03%
Rain + darkness	0.92	0.80	1.81%	2.97%
Rain + daylight	0.95	0.85	6.52%	5.42%
Dry + darkness	0.98	0.95	19.52%	32.05%
Dry + daylight	1.02	1.05	70.34%	58.52%

Table E.2: Capacity factors for different incident types classified by number of lanes

Lanes	Car Break down	Accident that blocks the hard shoulder	Accident that blocks one or more lanes	Rubbernec king
1	0.95	0.81	0.05	0.83
2	0.95	0.81	0.28	0.85
3	0.99	0.83	0.40	0.87
4	0.99	0.85	0.49	0.89

Table E.3: Probabilities for different incident types classified by number of lanes

	Probabilities peak period (x10-6)			
Road type	Car Break down	Accident that blocks the hard shoulder	Accident that blocks one or more lanes	Rubber-necking
Motorway	2.87	0.51	0.13	0.64
Main road	11.2	1.98	0.50	2.48
Secondary road	61.9	11.0	2.75	13.8
	Probabilities off-peak (x10-6)			
Road type	Car Break down	Accident that blocks the hard shoulder	Accident that blocks one or more lanes	Rubber-necking
Motorway	2.25	0.40	0.10	0.50
Main road	9.20	1.64	0.41	2.05
Secondary road	54.0	9.60	2.40	12.0

First, the 'nominal/regular' travel times (no disturbances) are computed with a static assignment model. The demand matrix is assumed to be fixed and can, for instance, be taken from any traditional four-step traffic and transport model. Thereafter, the variation to this nominal situation is calculated by randomly sampling disturbances from the distributions of weather conditions, roadworks, incidents, seasonalities, and events. This implies that, in every iteration, the capacity reduction (or increase) of all links (as a result of different weather conditions) or specific links (as a result of road works or incidents) is determined, and that also the demand between specific OD-relations (as a result of events) or all OD-relations (as a result of seasonality) is varied. The outcomes of this Monte Carlo simulation approach consists of travel times for each Monte Carlo iteration for all OD-pairs and on all links, from which various statistics can be computed. Some statistics, like the standard deviation, are computed automatically.

Two modes of route choice behaviour in case of disturbances can be used. In the first mode, a complete new equilibrium is found, which implies that everybody has the opportunity to deviate from the routes that they would use in the case without disturbances, and that they have complete information about the disturbances. The second option is that nobody changes routes. This fixed route choice behaviour matches with the situation in which nobody had information, route alternatives are not available, or nobody wants to make use of those alternatives. Of course, neither extreme is realistic, since, in practice, a few people always change their behaviour. However, using the two extremes gives insight into the bandwidth of the results. Furthermore, there is a lack of information about the actual route choice of drivers in case of disturbances, which makes it difficult to calibrate theoretically better route choice models, such as en-route route choice models.

Appendix F: A robust road network design for the area The Hague - Rotterdam

In this appendix, it is explained how a robust road network design is made for the network of 2020 for the area The Hague – Rotterdam in the Netherlands. In this appendix, the policy network of 2020 is referred to as “policy network”. This network is the starting point for the robust design. This robust road network is referred to as “robust network”. Schrijver et al. (2008) and Snelder et al. (2009a) provide an extensive description of this robust network design. The method used in this appendix does not match completely with the method that is proposed in this thesis. However the non-modelling steps are quite similar. Therefore, this appendix serves as an example to illustrate how the non-modelling design steps work.

Functional analysis

In this section the functional analysis is described. The dynamic traffic assignment model Indy is used for the functional analysis of the policy network of 2020. In Indy, spillback effects are considered. The functional analysis starts with a simulation of an average morning between 6.00 AM and 10.30 AM in 2020. The network and OD-matrix from the static NRM-model are used. The model has been calibrated based on the model outcomes of the NRM-model. For the dynamics of time and the congestion locations, a simple calibration procedure is used, which implies that it was checked if the congestion locations and moments on which the congestion occurs are more or less similar to the average daily situation on the network of The Hague-Rotterdam. There does not have to be an exact match, because the network, the level of demand, and the demand pattern are different in 2020 compared to 2008. In total, about 1.1 million car and truck trips are made in the period 6.00 AM to 10.30 AM. The following departure fractions are used for each half hour time period: 0.08, 0.1, 0.12, 0.14, 0.16, 0.15, 0.11, 0.08, and 0.07. This distribution is based on the measured departure time distribution of 2006. It is assumed that road pricing, which was intended to be introduced before 2020, changes the demand pattern in such a way that more drivers depart outside the peak period.

Figure F.1 shows the existing road network with the names of the roads. Figure F.2 and Figure F.3 show the number of lanes and the maximum speeds for the policy network of 2020. This network has 5424 links, 3335 nodes, 485 zones, and 5595 lane kilometres.

Figure F.1: The network of The Hague- Rotterdam of 2008

Figure F.2: Number of lanes in the modelled network of 2020

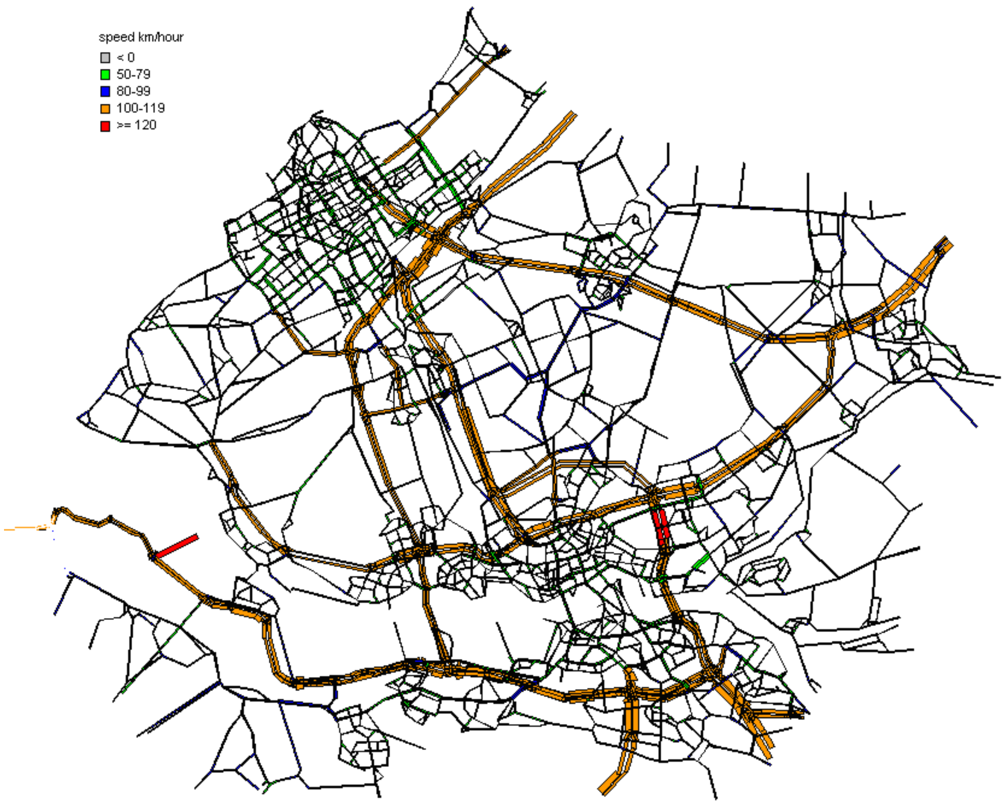


Figure F.3: Maximum speed in the modelled network of 2020

The simulations with Indy estimate that the total number vehicle kilometres is 14.26 million, the total travel time is 422,000 hours, and the average speed is 33.8 km/hour. Figure F.4 shows the speed ratio at 8.30 AM. The dark links are congested links. From this figure it can be concluded that there are several congested roads in the morning peak. Figure F.5 shows the vulnerability indicator that is introduced in section 3.5 (equation 3.1). From this figure, it can be concluded that parts of the A4, the A12, the A13, and the A15 are vulnerable (indicator >1). This is caused by the fact that there are no good route alternatives available. Some links at the border of the study area areindicated as vulnerable as well. However, since they are on the border, they are not necessarily vulnerable, since there might be good route alternatives outside the study area. Figure F.6 shows the results of a simulation with a static assignment model of an incident on each link. The OD-matrices for the period 7.00–9.00h was used for this simulation. The vehicle loss hours are measured as the difference between the total travel time in the situation with an incident on the selected link and the situation without an incident. The incidents are simulated by reducing the capacity by 90%. This figure more or less indicates the same roads are vulnerable as in Figure F.5, but adds parts of the ring road of Rotterdam to the list.

Figure F.4: Speed ratio at 8.30 AM without incidents in the policy network

Figure F.5: Vulnerability index in the policy network

Figure F.6: Vehicle loss hours caused by an incident on each link in the policy network

Now that the most problematic links in the situation without disturbances and the most vulnerable links in case of an accident are known, the cause of these problems has to be found. In this process, urban/regional traffic (44%), incoming and outgoing traffic (45%), and through traffic (11%) are distinguished. (Above, between the brackets, the share of each type of traffic for the area The Hague – Rotterdam was given.)

An analysis of the routes that are used by the urban/regional traffic showed that this traffic uses many different road types, each with its own speed limits, capacity, types of crossings, and layouts. However, the differences among these road types are no longer functional. The car driver just chooses the fastest route via a random combination of road types. For many urban/regional trips, the travellers have to make use of a limited number of motorways (A12, A13, A15, A20), for which there are no route alternatives. This makes the network structure vulnerable, as indicated before. Because the distance between motorways is too large compared to the spatial orientation of the cities, choosing shortcuts via routes that are not suited for this traffic becomes very attractive, which is not desirable for the villages through which those routes go.

With respect to the through traffic, there are two main flows: one from Amsterdam to Belgium in the north-south direction, and one from the harbour of Rotterdam to Germany in the east-west direction. Both flows make use of the same motorways that are used by the urban/regional traffic, which results in low speeds, unreliable travel times, and a disturbed flow pattern, especially on the motorways that pass Rotterdam and The Hague. Furthermore, for important routes like the A4 between the crossing A4/A13 and Leidschendam and the west side of the A15, there are no route alternatives.

The incoming and outgoing traffic is a mixture of the two types that are described before. Therefore, they experience similar problems.

Design process and design results

From the functional analysis, we concluded that the main problem is not congestion as such, but the fact that the urban/regional traffic is forced to use the motorways, which they have to share with the other traffic and for which there are no route alternatives. The selected design standards indicated that the structure has to be changed towards:

- A separate, robust, and well-balanced network for the urban/regional traffic with a maximum speed of 80 km/hour. The choice for 80 km/hour was made in order to avoid that an increase of the demand will occur when the road infrastructure is improved. According to Immers et al. (2004d), the optimal spare capacity is about 800 pcu per hour per lane. The functional analysis showed that this spare capacity is not always available, which implies that for those links or screen lines where not enough spare capacity is available, extra capacity should be added to make the network more robust. However, since there is latent demand in the existing network, the new capacity will be filled with new cars, which would imply that there is still not enough capacity. In order to avoid this, a choice was made to improve the network in such a way that the network quality does not improve too much. This can be accomplished by reducing the speed on well chosen routes. Therefore, the choice was made to give the urban/regional network a maximum speed of 80 km/hour instead of the 100 or 120 km/hour on the motorways that are currently used by the urban/regional traffic. A well-balanced network implies that route alternatives with an equal speed and number of lanes (mostly 2 lanes in each direction) and with an equal distance between the routes (about 5 km) should be created. The chance of congestion on the urban/regional network is smaller, because of the extra spare capacity. If, nevertheless, congestion occurs, the introduction of buffers avoids spillback to other routes. Applying these principles to the network of The-Hague-Rotterdam results in the network for the urban/regional main roads that is shown in Figure F.7.

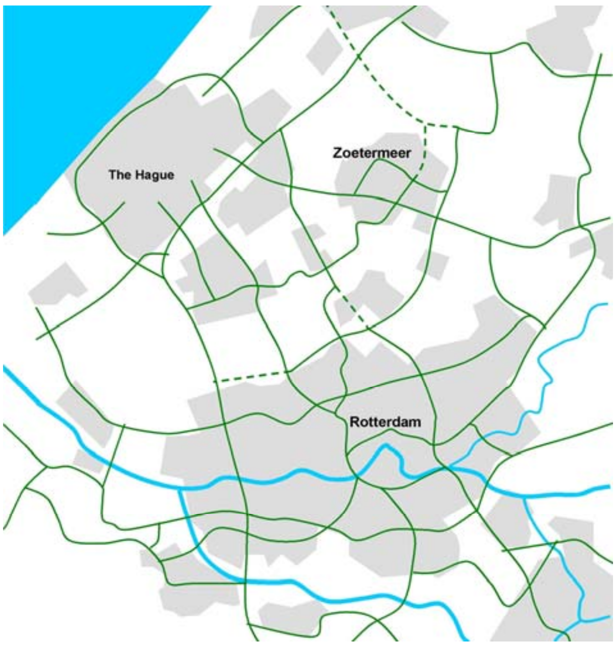


Figure F.7: Robust road network design for the urban/regional traffic

- Since the through traffic is only 11% of the total traffic, only a limited number of main routes/motorways has to exist in the network. Since the through traffic does not want or need to be disturbed by the urban/regional traffic, a separate infrastructure for this traffic can be constructed with a long distance (about 10 km) between the off-ramps and on-ramps. In the existing network this is often not more than 2 km. A maximum speed of 120 km/hour for these roads is proposed. For the through traffic, there are two main flows: one in the north-south direction, and one in the east-west direction. The main east-west route goes via the existing A15. For the main north-south route, there are two options. The first goes via the existing A4, including the missing parts of the A4. The second option goes via the existing A16 and continues with a new road that connects the A16 and the A4. The advantages of this route are that urban areas are avoided and that this route offers a completely new route alternative. However, since the expected realization time is quite high, we have chosen the first option. Since this option does not have a complete route alternative, route alternatives should be created via the urban/regional network. For these flows, congestion in the normal situation should be avoided. This can be accomplished by buffers on the on-ramps and off-ramps. The main routes have two lanes in each direction. Besides these two main routes, several additional roads offer extra routes for the incoming and outgoing traffic to and from the region. Furthermore, several backup options over the urban/regional network are offered that can be used in case of disturbances. Traffic management can help in route guidance in these situations. Figure F.8 presents the network options for the through traffic.

a: b:

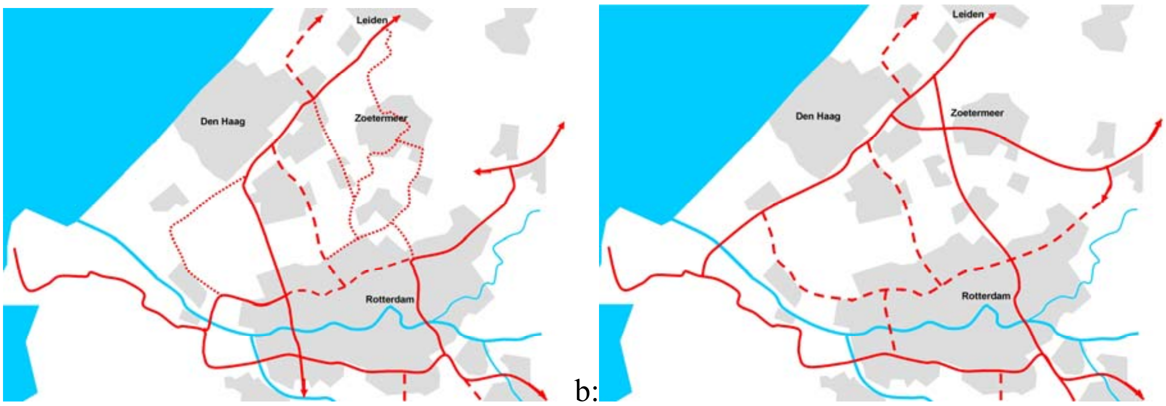


Figure F.8: Two options for a robust road network design for the through traffic

- A robust road infrastructure consists not only of links. A suitable configuration of nodes is also necessary. For the through traffic network, these nodes are drawn in detail. The exact configuration of these nodes can be found in (Schrijver et al., 2008). In the urban/regional network, the crossings between the roads can be of any type. The choice for a roundabout, fly-over(s), and/or traffic lights depends upon the situation.

- Buffers can make the network more robust. They can prevent congestion to spill back to other roads, and they can be used to regulate the inflow to the main routes. Since the network for through traffic and the network within an urban area should be free of congestion, the buffers are located on the regional part of the urban/regional network. An example of this is shown in Figure 5.11.

- Good transport options from and to public transport are part of a robust network, since this offers extra choice options and backup options. In Figure F.9, the transfer options (✱P+R points) at the city borders of Rotterdam and The Hague are shown. These transfer options are easy to reach from the main roads, and offer an easy transfer to the fast and frequently scheduled public transport. Snelder et al. (2010c) present an extensive description of how this design was made (in Dutch).

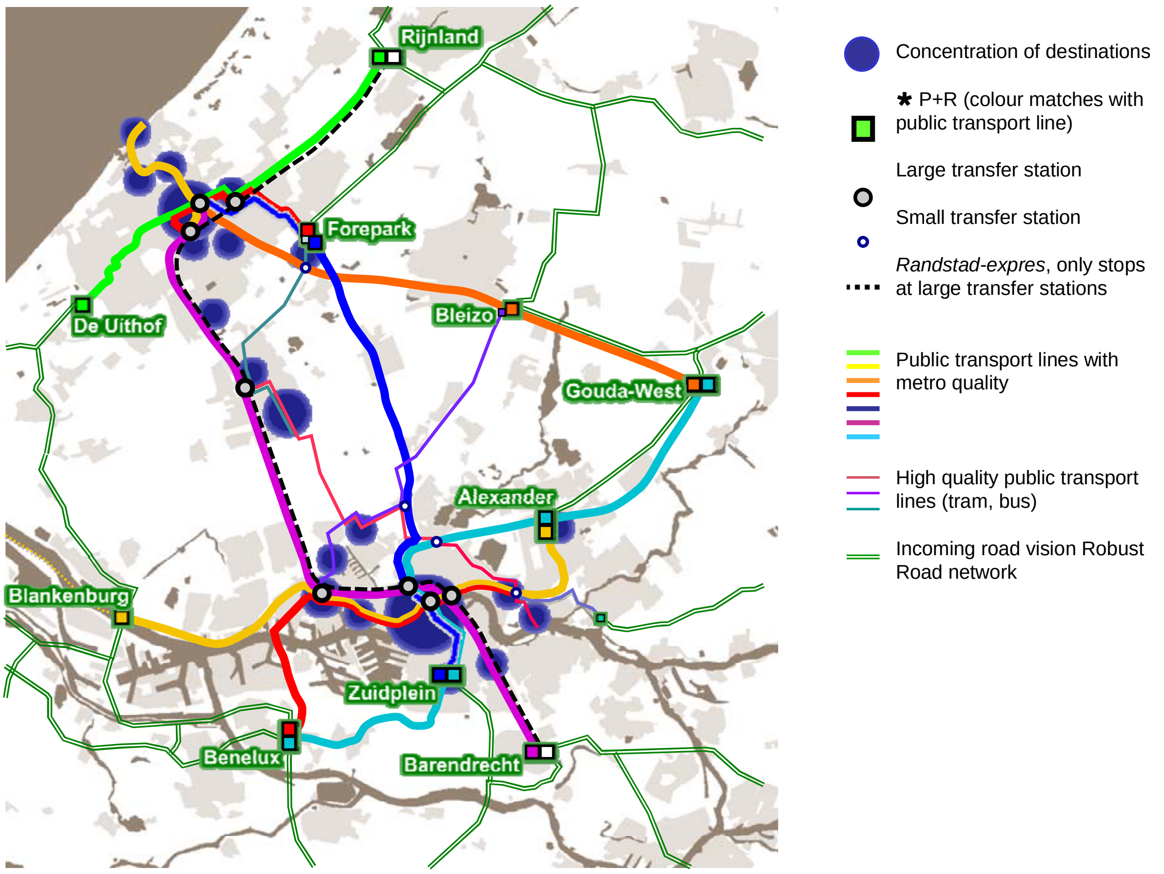


Figure F.9: Connection with the public transport network

In Figure F.10, the complete network is presented. Figure F.11 shows the difference between this network and the policy network of 2020. Sometimes the robust network (links and nodes) uses less space than the policy network (indicated in gray), and sometimes it uses more space (indicated in black). The new design contains more or less the same number of lane kilometres. The main difference is that the investments are allocated more to the urban/regional network. The following network changes are required:

- Some of the motorways that are mainly used for urban/regional traffic are downsized to urban/regional roads,
- Some of the local roads are upgraded to the urban/regional level,
- Some of the motorways are split into parallel structures for the through traffic and the urban/regional traffic, such that both types of traffic do not disturb each other,
- some of the nodes are restructured to fit their new function,
- some missing links in the network for the through traffic, and mainly in the network for the urban/regional traffic, are added,
- buffers are added at on-ramps and off-ramps, at the borders of cities, and at some other important locations in the urban/regional network,

- flexible infrastructure elements (like movable barriers) are added that help in rerouting traffic in case of non-regular situations like incidents,
- transfer points to and from public transport are added.

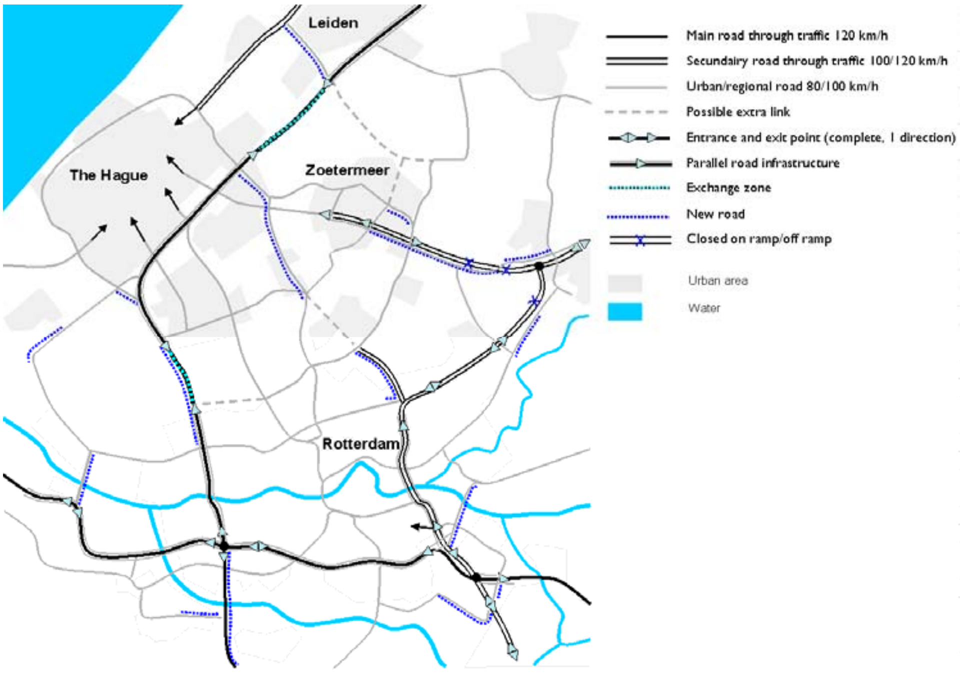


Figure F.10: Complete robust road network design

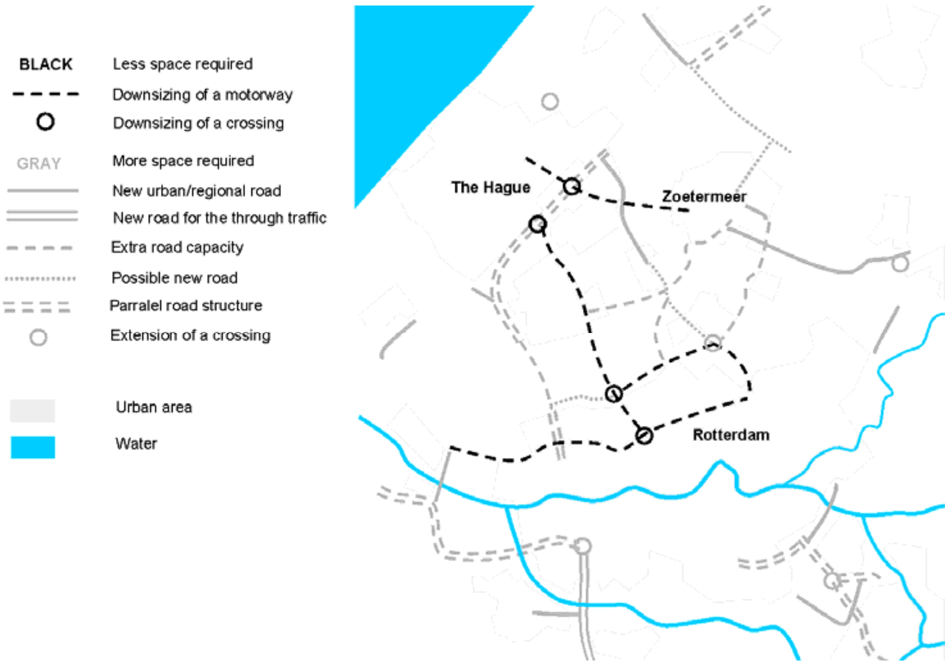


Figure F.11: Changes of the robust network compared to the policy network

How robust is the robust road network?

The performance of the robust road network presented above was tested with the model Indy. The exact same model settings and model inputs were used as with the functional analysis of the policy network. The only exception to this is the network itself, which is adjusted according to the new design. This implies, for instance, that the travel demand was not changed, which is in line with one of the design goals of not creating extra demand. Furthermore, this assumption is necessary for making a comparison of the network performance with the policy network. The robust network has 5522 links, 3430 nodes, 485 zones, and 5729 lane kilometres. The robust network, therefore, has only 2.4% more lane kilometres than the policy network.

The simulations with Indy estimated that the total number of vehicle kilometres is 14.14 million (-0.8%), the total travel time is 412,000 hours (-2.3%), and the average speed is 34.3 km/hour (+1.6%). The numbers between the brackets show how these numbers differ from those of the policy network. Figure F.12 shows the speed ratio at 8.30 AM. A comparison of this figure with Figure F.4 shows that the traffic is spread more equally over the route alternatives in the robust network. The flow on the N13 (former A13) is, for instance, lower. The road between the A13/A16 is used less in the robust network, and there is less congestion on the A16. At a few places there is still congestion. The A20 and the parallel road at the A12 are examples of this.

Figure F.12: Speed ratio at 8.30 AM without incidents in the robust design

Figure F.13 shows the vulnerability indicator. Compared to the policy network (Figure F.5), the A13, A15, and A12 are less vulnerable. Figure F.14 shows the results of a simulation with a static assignment model of an incident on each link. This figure illustrates even more clearly that the robust network is really more robust, since the vehicle loss hours are clearly less extreme than in the policy network. In fact, the average number of vehicle loss hours is 29% lower (95 hours compared to 135 hour), and the standard deviation of the vehicle loss hours is lower as well (240 hours compared to 339 hours).

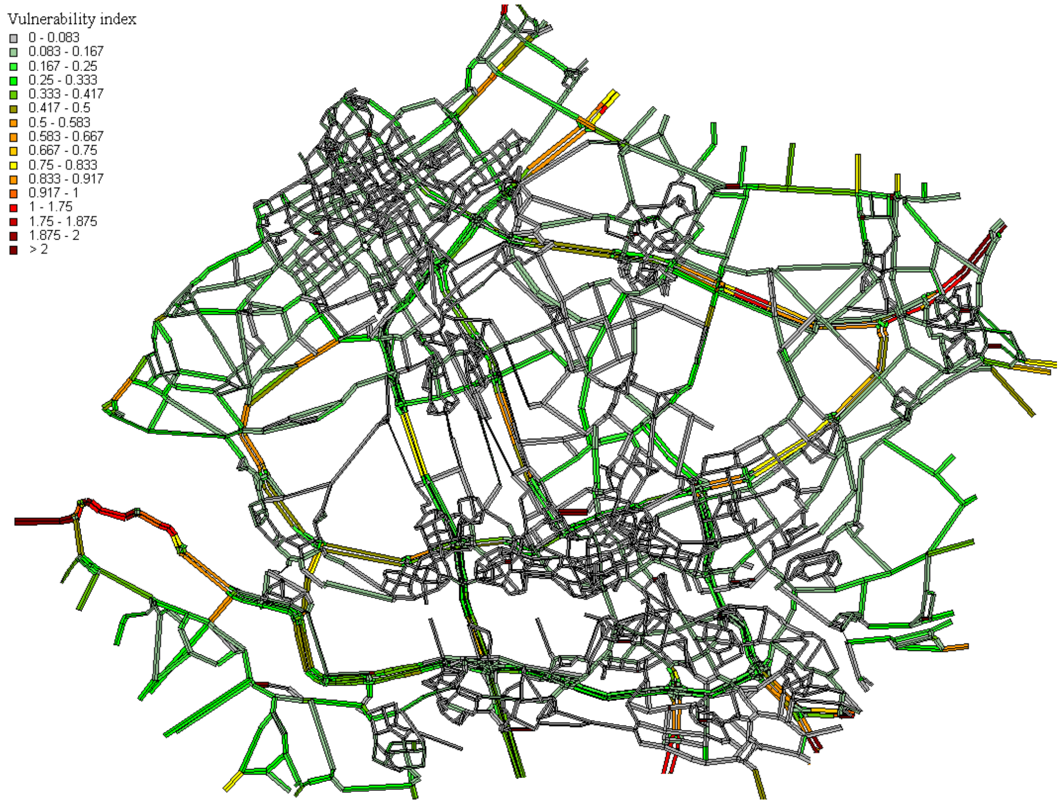


Figure F.13: Vulnerability index in the robust design

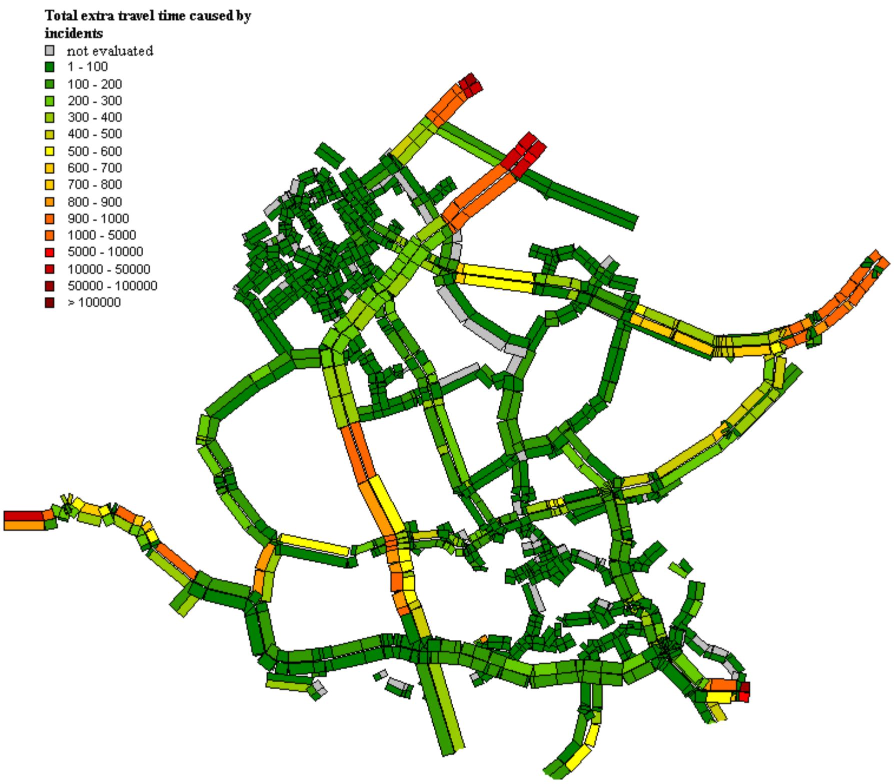


Figure F.14: Vehicle loss hours caused by an incident on each link in the robust design

Since the number of vehicle kilometres (and therewith the emissions) in the robust network is 0.8% lower than in the policy network, the speeds of the urban/regional traffic are lower, and the traffic is spread more equally over different routes, the air quality is most likely. Furthermore, since the urban/regional traffic uses roads with a maximum speed of 80 km/hour instead of 100/120 km/hour, fewer kilometres are driven, and accidents cause less congestion in the robust network, which reduces the chances of secondary incidents, the robust network should be safer. Note that we did not calculate the exact costs of constructing the robust network, because the network needs to be specified in more detail before this can be done.

Appendix G: Model parameters

This appendix presents tables containing the model parameters, their descriptions, and their default values. It is important to notice that the parameters that are described in this appendix do not cover all the decisions that have to be made, since the design method involves a lot of non-modelling decisions as well. The design-related choices with respect to the costs, the decision variables, and the algorithm are presented respectively in Table G.1, Table G.2, and Table G.3.

Table G.1: Cost parameters

Cost parameter	Description	Default
$vot	Average value of time [euro/hour]	15
$vod	Average cost per kilometre. Combination of fuel, car ownership cost per km, road pricing cost, and external costs. Can specify per link or as a general parameter [euro/km]	0.8
$discountrate	Discount rate {0 …1}	0.055
$economic_lifetime_infra	Economic lifetime of infrastructure [years]	30
Fixed_costs	Fixed infrastructure costs. Can be specified per link [euro/km/week]	5298.8
Variable costs	Variable infrastructure costs. Can be specified per link [euro/pcu/km/week]	1.703

Table G.2: Decision variable parameters

Decision variable	Description	Default
$max_extraregularcap	The maximum number of extra lanes that can be added in the optimization under regular conditions {1,2,3…..,n}	5
$max_extrasparecap	The maximum number of extra lanes that can be added in the optimization for incidents {1,2,3…..,n}	2

Table G.3: Algorithm parameters

Algorithm parameters	Description	Default
$GenerateSkimPT	If True, a job is run to generate a matrix with generalized traffic costs for modes other than cars/trucks based on assumptions that can be adjusted {True; False}	False
$DoTripdistribution	If True, the demand model is run {True; False}	True
$DoInitialIndyAssignment	If True, an initial assignment for Indy is done {True; False}	True
$DoIndyAssignment	If True, Indy assignments are done in between the optimization steps {True; False}	True
$DoOptimizeCapRegular	If True, the capacity under regular conditions is optimized {True; False}	True
$DoOptimizeSpareCap	If True, the spare capacity for incidents is optimized {True; False}	True
$NumberOfTripGenIterations	The maximum number of iterations between the demand model and Indy {1,2,3…..,n}	10
$NumberOfIndyIterations	The number of iterations in Indy between route choice and network loading {1,2,3…..,n}	10
$NumberOfOptmize-RegularCapIterations	The number of iterations between the optimization algorithm for regular conditions and Indy (and the demand model) {1,2,3…..,n}	10
$NumberOfOptmize-SpareCapIterations	The number of iterations between the optimization algorithm for incidents and Indy (and the demand model) {1,2,3…..,n}	2
$populationsize	The population size that is used in the genetic algorithm {1,2,3…..,n}	50
$ga_iterations	The number of iterations in the genetic algorithm {1,2,3…..,n}	20
$eps_demand	The stop criterion: relative change in total demand for the iterations between the demand model and Indy {0,…,1}	0.01
$eps_travelcost	The stop criterion: relative change in total travel cost for the iterations between the demand model and Indy {0,…,1}	0.01
$eps_totalcostreg	The stop criterion: relative change in total costs for the iterations between the optimization algorithm for regular conditions and Indy (and the demand model) {0,…,1}	0.01
$eps_totalcostspare	The stop criterion: relative change in total costs for the iterations between the optimization algorithm for incidents and Indy (and the demand model) {0,…,1}	0.01

Table G.4 describes the decisions that have to be made with respect to the dynamic traffic assignment with Indy. These choices are mostly related to the time steps that have to be used and the locations where the results are stored. In Table G.5 the parameters for the MIC module are described, and in Table G.6 the parameters for the alternative route choice approximation algorithm are described. These tables are, thus, considered with evaluation related choice.

Table G.4: Indy parameters

Indy parameters	Description	Default
$purpose	Number of trip purpose {1,2,3 …..n}	1
$mode	Number of mode {1,2,3…..n}	10
$time	Time step for which network is specified	1000
$user	Number of user class {1,2,3…..n}	1
$result	Number of location where results are stored {1,2,3…..n}	1
$iteration	Number of iteration where results are stored {1,2,3…..n}	1
$timestep_demand	Demand interval [minutes]	10
$timestep_indyoutput	Aggregation time for which output is stored [minutes]	5
$min_time	Start time for which the results are included in the capacity optimization [minutes] {1000 …..}	1000
$max_time	End time for which the results are included in the capacity optimization [minutes] {1000 …..}	1180
$reference_timestep _demand	Time that is used as reference time [minutes]	1090
$networkloading	Network loading model that is used {vertical queuing = NONE; horizontal queuing = DYNAMIC; LTM = PHYSICAL}	PHYSICAL
$effectivevehiclelength	Effective vehicle length used for computation of lane jam density and possibly for the lane capacity if not specified in the initial network [meter]	8
$responsetime	Response time used for computation of lane capacity if not specified in the initial network [seconds]	1.25
$lanejamdensity	Lane jam density [pcu/km/lane]	1000/ $effective-vehiclelength
$jamspeed	Minimum speed in queues [km/hour]	10
$min_outflow	Minimum outflow of links [pcu/hour]	600

Table G.5: MIC parameters

<table>
<tr><th>MIC parameters</th><th>Description</th><th>Default</th></tr>
<tr><td>Incident list</td><td>List of links on which an incident is simulated</td><td>All links</td></tr>
<tr><td>$nincidenttypes</td><td>Number of incident types to be modelled {1,2,3…..n}</td><td>4</td></tr>
<tr><td>Incident types</td><td>Start time [hours], end time [hours] reduction (share capacity left {0.. 1}, incident probability per 100,000 km</td><td><table>
<tr><th>Incident-type</th><th>start</th><th>end</th><th>reduction</th><th>prob.</th></tr>
<tr><td>1</td><td>0.75</td><td>1</td><td>0.95</td><td>0.34</td></tr>
<tr><td>2</td><td>0.75</td><td>1.5</td><td>0.8</td><td>0.06</td></tr>
<tr><td>3</td><td>0.75</td><td>1.5</td><td>0.3</td><td>0.015</td></tr>
<tr><td>4</td><td>0.75</td><td>1.5</td><td>0.85</td><td>0.07</td></tr>
</table></td></tr>
</table>

Table G.6: Parameters for the approximation algorithm for alternative routes

Alt. routes parameters	Description	Default
$parameter_dist	Distance parameter {	0.85
$perc_altroutes	Maximum percentage of traffic that takes an alternative route in case of disturbances {0…1}	0.15
$maxangle	The maximum absolute angle between the original link and the alternative link [degrees] {0..90}	60
$maxdist	The maximum distance as the crow flies between the original link and the alternative link. Used to narrow down the search space. [kilometre] {0…n}	25
$distance_vulnerability	The way in which the distance between the original link and the alternative link is measured {"DIJKSTRA" = shortest distance over network}; "SHORTEST" = distance as the crow flies}	"DIJKSTRA"
$scale_factor_to_km	Scale factor of coordinate system used in network to kilometre. For instance, from meter kilometre = 0.001 {0…..n}	0.001

Summary

The Dutch road network is, like many other road networks in the world, congested in the morning and evening peaks. The locations of congestion are quite often the same; this makes it relatively easy to take the delay of this regular congestion into account when planning a trip. However, as a result of unforeseen disturbances, also unexpectedly large delays occur. A data analysis of a single incident on an off ramp showed that small incidents can block a large part of a motorway network for many hours. Many individual travellers experience an unexpected delay of more than one hour as a result. Incidents like this happen quite often. In the future, the effects of those incidents will likely become larger, because the spare capacity in both space and time will be reduced as a result of growth in demand in urbanized areas, if no measures are taken. Our simulations indicated that the costs of vulnerability in 2008 were ranged between €275 million and €1.2 billion. If no measures are taken, by 2030 these costs might increase to between €900 million and €4.1 billion. These results suggest that it would be beneficial to implement robustness measures to reduce these costs.

Road networks are vulnerable – the costs of vulnerability can add op to several billions of euros per year in the Netherlands if no measures are taken.

In order to reduce the vulnerability, different measures need to be taken that make the road network more robust. Our main objective was to develop a network design method for designing road networks that are robust against incidents which is applicable to real sized networks and can be used in practice. We focussed on the robustness of the road structure. This can be seen as a first step towards the development of a robust mobility system (including public transport and multimodal trips) in which also non-structure related measures are taken.

Before this could be done, a common framework had to be developed that gives policy makers and transportation analysts the possibility to discuss issues that are related to road network robustness and vulnerability. Robustness is defined as follows:

Robustness is the extent to which, under pre-specified circumstances a network is able to maintain the function for which it was originally designed.

Vulnerability is the opposite of robustness. A network that is vulnerable is not robust, and vice versa. In principle, a road network can be made robust against all kinds of disturbances ranging from predictable to non-predictable disturbances, from regular to irregular disturbances, from disturbances with a temporarily small or large impact/effect to disturbances that have a permanent impact, and from disturbances that have a network-wide effect to disturbances that have a local effect. Network managers and/or policy makers have to decide how robust the road network is to be made against which disturbances. This is what is meant with the clause 'pre-specified circumstances' in the definition of robustness. Of course, it is preferable to make the network robust against as many disturbances as possible. However, this is not always possible because, for instance, making the network robust against earthquakes requires other investments than making the network robust against incidents. The robustness of a road network focuses more on effects than on probabilities. Nevertheless, for network design, probabilities are an important factor for investment decisions. Therefore, this thesis focused on the class of non-predictable regularly occurring (relatively high probability) disturbances that we called *incidents* (including some incidents, like road closures, that belong to the class of non-regular non-predictable disturbances).

In principle, road networks can be made robust against all disturbances. Network managers and policy makers need to specify against which disturbances their network is to be made robust, and at what cost.

A definition of robustness is not enough to design a robust road network. The question what makes a road network robust, needs to be addressed as well. A comparison between the road network, the railway and inland waterway network, the cardiovascular and nervous system and telecommunication networks showed that a network becomes more robust if there are alternative routes available and there is enough spare capacity on all the routes. However, having this redundancy is not enough. The transported elements must also be rerouted to these alternative routes as fast as possible, which requires some flexibility in the network and some traffic and incident management strategies. Furthermore, in a robust network, spillback effects must be reduced to a minimum level. Finally, preventive measures make a network more robust as well. An analysis of incidents that happened in one part of the Netherlands underlined that five elements (prevention, redundancy, compartmentalization, resilience, and flexibility) are likely to make a network more robust. This is important to realize, because this gives direction to the indicators to use for robustness analysis and the measures that can be taken to improve the robustness of the road network.

A robust road network is a network in which preventive measures are taken and that has redundancy, is compartmentalized, is resilient, and is flexible.

For the traveller, the term robustness is irrelevant. In the end, the traveller wants to have reliable travel times and does not want to be confronted with unexpectedly high travel times as a result of incidents. In the developed framework, we showed that robustness and reliability are two related concepts that differ in focus. The focus of robustness is on the network structure and the effects of disturbances with an unexpected effect at a specific moment in time, whereas the focus of reliability is on the traveller and the probability of expected disturbances measured over a longer period. In a robust road network, disturbances lead, on average, to smaller variations in travel times, and, therewith, to more reliable travel times.

However, this is not a one-on-one relation. Making the network more robust, does not have to result in more reliable travel times for all travellers. If robustness measures are taken, the reliability of the network should be evaluated by including all kinds of disturbances. However, it is necessary to give special attention to the disturbances with large effects, such as incidents, because these effects are easily underestimated if traditional methods for reliability analysis are used.

Robustness and reliability are related concepts that need to be addressed together.

In order to include robustness in the network design an indicator is needed that somehow relates to reliability. In this thesis, many indicators were presented that refer to one or more elements of robustness (prevention, redundancy, compartmentalization, resilience, and flexibility) or to travel times or reliability of travel times. We chose to use the indicator ‘vehicle loss hours caused by incidents’, mainly because this indicator describes the different elements of robustness best. Depending on the way it is computed, this indicator can consider spillback effects (compartmentalization), route alternatives (redundancy), and resilience. If flexibility is included in the network, these flexible infrastructures can be considered (depending on the type of model that is chosen). Furthermore, the indicator ‘vehicle loss hours caused by incidents’ is relatively easy to explain and the can be valued by using a value of time. It was also explained that monitoring the indicator (based on data) is complex, but possible to a certain extent. The most complex part of this indicator is to find a model by which this indicator can be computed in an acceptable computation time.

Robustness is measured by the vehicle loss hours caused by incidents. This indicator was chosen mainly because it describes the different elements of robustness best.

The dilemma in the model choice for evaluating robustness in network design is that a choice has to be made between accuracy (ideally using a dynamic traffic assignment model with detailed congestion modelling, including spillback effects, with multiple types of route choice behaviour during incidents, and with accurate intersection modelling) and computation time. The most accurate models generally take the longest computation time. Using a rule of thumb, for instance, takes hardly any computation time, but is not accurate. For some applications, a rule of thumb can be good enough to get a quick impression of the robustness of a network. However, to make a well balanced decision about robustness measures in network design, it would be better to look for a method/model that deals with the above mentioned requirements in the best possible way and is fast, because many evaluations have to be carried out in network design.

We chose to use the dynamic traffic assignment model Indy to compute an equilibrium under regular circumstances, because this model has an accurate network loading model that models spillback effects according to the simplified kinematic wave theory of Newell. Furthermore, the model can compute an equilibrium route choice and can deal with fixed route choice. En-route route choice is not possible. However, this is not a problem, since the model Indy is used only for a basic run in a situation without disturbances, for which an equilibrium assignment is the most appropriate. Indy is a dynamic model, which makes the modelling of time dynamics possible. And, since Indy is able to work with the marginal incident computation model (MIC) (Corthout et al., 2009), it is able to get an estimate of the impact of incidents very quickly. Four incident types were defined, for which the effects are computed for all links. An approximation algorithm was added to the MIC-module to deal with the use of alternative routes. We showed the quality of the method by making a comparison with loop detector data for five incidents. Finally, we showed how robustness/reliability benefits of robustness measures can be computed with this evaluation method, and how they can be included in a cost-benefit analysis.

Evaluating the robustness of a network in network design requires an accurate and fast evaluation method that considers spillback effects and alternative routes. A macroscopic dynamic traffic assignment model is combined with a marginal incident computation model and an approximation algorithm for alternative route choice in order to evaluate the robustness in acceptable computation time.

When robustness measures are taken, robustness/reliability benefits should be considered, because those measures are intended to improve the robustness of the network and, therewith, the reliability of travel times. It is obvious that including solely the travel time benefits under regular conditions, as is usually done, is not enough. Different examples showed that, in specific cases, the robustness/reliability benefits can be significant: possibly of equal size to the travel time benefits. This implies that the benefit-cost balance of robustness measures can be much more positive than when the robustness/reliability benefits are not considered.

The benefits of robustness measures should be included in cost-benefit analyses.

This raises the question which measure can be taken to reduce the vulnerability or to increase the robustness of the road network and how these measures could be applied. A list of measures that are likely to improve robustness were presented in this thesis. These measures were subdivided into measures that can be taken on the travel market, the transport market, and the traffic market. Furthermore, it was explained to which elements of robustness they are related, and for which disturbances they are relevant. For a selection of measures that match the best with the focus of this thesis, the possible effects of the measures were analysed by using examples.

It was shown that having a single route alternative could reduce delays by 3.1% - 77.1%, depending on the level of information that is given. Additional extra paths result in only small improvements. Of course, the results also depend on the quality (spare capacity and travel time) of the alternative routes. With respect to unbundling, we concluded that unbundling can have a positive effect on robustness. However, the way in which it is done and the locations where it is done need to be chosen carefully, because unbundling can have both advantages and disadvantages for the robustness of a network. Finally, we showed that adding buffers and spare capacity to the network can have large positive effects on travel times. Buffers have the potential to reduce the average travel time by 12% in the morning peak in the network of The Hague. In other networks and in other time periods, this impact can, of course, be different. The more congestion there is, the higher the benefits are likely to be. This conclusion is based on a regular situation without incidents. However, it is straightforward to show that similar effects occur in the case with incidents. This implies that buffers can make a network more robust. Finally, an example taken from Immers et al. (2004d) showed that the optimal spare capacity of an average motorway link is about 800 pcu/hour. This optimization was done without considering spillback effects.

Many measures can be taken in order to improve the robustness of a road network. With respect to the network structure, adding spare capacity, creating parallel routes that are spread in a balanced way over the network, adding buffers and unbundling are the measures with the highest potential to improve the robustness of a road network.

However, the benefits of those measures depend on local circumstances. The benefits can not be approximated by a rule of thumb or regression analyses.

The question that remains is where in the network which measures can best be taken. This so-called robust road network design problem is a complex problem from a practical and theoretical point of view for instance because many different stakeholders are involved that all have different partly overlapping and partly contradicting objectives. Furthermore, many disturbances can occur on many locations and different driver groups respond differently to that. Finally, a lot of different measures can be taken on many different locations. We formulated the robust road network design problem as a bi-level problem in which a the top level a network is designed which is robust against incidents and in which at the lower level the route choice and trip choice are determined given the network structure. The presented formulation differs from formulations of the Network Design Problem in literature, because it combines destination choice, mode choice and dynamic route choice in the lower-level problem and it includes short-term variations in supply caused by incidents in the Network Design Problem.

The method that has been developed to solve the robust road network design problem combines expert knowledge and advanced modelling techniques which makes it possible to design a large scale robust road network with many design variables (= robustness measures) that is likely to be supported by all the stakeholders that where involved in the design process. The fact that experts get the opportunity to interfere in the optimization process can be considered as a big advantage since this makes it possible to enrich the model input and outcomes with knowledge of stakeholders that can never be included in a model. Furthermore, this makes it possible to include all kinds of robustness measures in the optimization process. It is very difficult to include this many design variables in a full model approach. Finally, it creates support for the network design during the design phase and might prevent a lot of discussion afterwards. The fact that extensive use of models is made is an advantage since this gives useful input for the part in which the experts are involved, because the evaluation part of the models shows how vulnerable the existing network is and which locations of the network are vulnerable. It does the same for new designs. The optimization part of the models gives suggestions about how much spare capacity is to be added on which locations. Finally, this gives quantitative support to the decisions that are made.

The network design method consists of three steps, with feedback loops. The steps are briefly described below:

- Step 1: *Design standards*. Designing a road network starts with specifying the design standards. The objective of designing/improving the network, the design variables, and the restrictions are specified.
- Step 2: *Functional analysis*. A functional analysis is an analysis in which the network performance of the reference network is evaluated. This is done based on expert knowledge of the different roads in the network and based on an evaluation model that evaluates the robustness of all the links in the network. The evaluation model first determines the travel times and congestion under regular conditions, and analytically determines the cost of the travel time losses by four different incident types that occur with a certain probability on each link. An approximation algorithm for alternative route choice during incidents is added to that evaluation method.
- Step 3: *Design process*. In this step the robust road network is designed. The design standards and the results of the functional analysis are used as a starting point. A model can be used to get an idea of where adding capacity would be the most useful. At first the

'optimal' number of extra lanes for the regular situation is determined. Thereafter the 'optimal' number of extra lanes of spare capacity for improving the robustness is indicated. After this modelling exercise, a group of experts (these could be stakeholders and/or traffic engineers) have the opportunity to adjust the suggested network. They can indicate which alternative routes should be used, where capacity should be added, and how much capacity should be added there. Furthermore, they can add all kinds of other measures to improve the robustness of the road network. Some design principles were proposed that can help the experts in their task.

After adjusting the network, the performance of the network needs to be evaluated again with the same evaluation model as is used in the functional analysis. Furthermore, this evaluation indicates what the remaining problems in the network are. A choice could be made to go back to the design steps and change the network design slightly to eliminate or reduce these problems. This makes designing an iterative process.

Combining expert knowledge and advanced modelling techniques makes it possible to design a large scale robust road network with many design variables (= robustness measures) that is likely to be supported by all the stakeholders that where involved in the design process.

The design method was extensively tested for a small test network in order to show the quality and the sensitivity of the method. It was shown that for this small test network a good solution (even the global optimum) could be found. Thanks to the simplifications the number of times that a full model run with a dynamic traffic assignment needs to be done stays limited which makes an application of the model to large networks possible.

The design method was also applied to a large realistic network of Amsterdam and surroundings. As far as is known to the author, this is the first application of a network design method that combines models and expert knowledge in the network design phase on such a large a scale by taking robustness into account. Furthermore, it is the first application in which a dynamic traffic assignment model is used in combination with a marginal incident computation model in order to make a robust road network design. The application of the evaluation method by itself is also a large step forward, since it shows how the vulnerability of a network can be analysed by taking spill back effects and alternatives routes into account. By applying the design and evaluation method to a large real-size network a lot of lessons were learned about how robustness analysis and robust road network design can be done in practice. These lessons relate to the level of detail of the network, the calibration, the simulation of incidents, the parameter settings of the models and the results that can be expected. The application of the modelling steps of the design method showed that large improvements in the network performance can be reached with a positive benefit-cost balance. The total travel costs under regular conditions were reduced by 12% and the total travel costs under irregular conditions (incidents) were reduced by 97% which implies that they were almost eliminated. These results were already obtained by using a limited number of iterations which results in a sub-optimal solution. If the model is run longer it is very likely that better solutions will be found.

Finally, an impression was given on how expert judgements can further improve the robustness of the design by applying all kinds of other robustness measures than adding spare capacity to the vulnerable roads.

The robustness of many parts of the road network can likely be improved with a positive benefit-cost balance. For a large size realistic network of Amsterdam and surroundings it was shown that the total travel costs under regular conditions can be reduced by 12% and the total travel costs under irregular conditions (incidents) can be reduced by 97% which implies that they can almost be eliminated.

Robustness measures can lead to an increased demand if the travel times under regular conditions improve. This increase in demand was not considered in the test case for the network of Amsterdam. However, it was considered in the small test network. It was shown that robustness measures can lead to an increased demand. If the demand increases, the intended increase in spare capacity might be cancelled out. This implies that the robustness of the network has not improved by the amount that was to be expected if demand effects had not occurred. Both improved robustness and an increase demand can lead to welfare effects. Therefore, they need to be balanced. The presented design method makes this possible.

A balanced choice needs to be made between improving the travel times under regular conditions and therewith generating mobility, between improving the robustness of a network and between making maximum use of the existing network

The findings that are presented in this thesis have a lot of implications for policy makers, network managers, road authorities, road users and other stakeholders. For policy makers, we have justified the inclusion of robustness analysis and robustness measures in policy plans. We showed that they should consider robustness in addition to measures that, for instance, aim to reduce travel times under regular conditions, improve reliability, and minimize negative external effects. We also explained that different measures can have contradictory effects on the different objectives. For instance, making maximum usage of the road network under regular conditions leaves no spare capacity for situations when disturbances occur and, therefore, is not robust. Therefore, it is needed to make a balanced choice between measures that aim to make maximum usage of the capacity under regular conditions and robustness measures. Furthermore, structure related measures need to be accompanied by accurate information, dynamic traffic management and incidentmanagement strategies in order to make maximum use of a robust road structure in case of disturbances. Industires that develop intelligent transport systems, like navigations systems, can contribute to this.

For traffic analysts/consultants and researchers, this thesis offered a lot of tools and insights by which robustness analyses can be done and by which robust road networks can be designed. Finally, this thesis points the way to a lot of future research that needs to be carried out. For the other stakeholders, like interest groups, there is a task to form an opinion about robustness (if not yet done) and, to help were possible in making the network more robust.

Finally, this research intents to contribute to the realisation of a robust road network. In the end, the road users benefit from that because they are less often confronted with large unexpected delays.

- Maaike Snelder, 1 December 2010

Samenvatting

In veel landen behoren files in de spitsperioden tot het gewone straatbeeld. De plekken waar de files staan en de tijdstippen waarop ze optreden zijn redelijk voorspelbaar. Met de extra reistijd als gevolg van deze files kan dus tot op zekere hoogte goed rekening worden gehouden. Naast vertragingen door deze reguliere files kunnen verstoringen zoals ongelukken tot onverwacht grote vertragingen leiden. Een data-analyse van een klein ongeluk op een afrit heeft aangetoond dat kleine incidenten een groot deel van het wegennetwerk kunnen blokkeren en dat automobilisten hierdoor gemakkelijk vertragingen van meer dan een uur kunnen oplopen. Dit soort ongelukken gebeuren in Nederland ongeveer 2000 keer per jaar. Het effect van deze verstoringen wordt in de toekomst waarschijnlijk groter als maatregelen uitblijven omdat door een toename van de vervoervraag in en tussen stedelijke gebieden de reservecapaciteit in het wegennetwerk afneemt. Indicatieve simulaties hebben uitgewezen dat de kosten van een kwetsbaar wegennetwerk in 2008 tussen 275 miljoen Euro en 1,2 miljard Euro lagen. Indien maatregelen uitblijven, kunnen deze kosten in de toekomst (ongeveer 2030) oplopen tot bedragen tussen 900 miljoen Euro en 4,1 miljard Euro per jaar.

Wegennetwerken zijn kwetsbaar – de kosten van kwetsbaarheid kunnen oplopen tot enkele miljarden per jaar in Nederland als maatregelen uitblijven.

Om de kwetsbaarheid van het wegennetwerk te verminderen moeten verschillende robuustheidsmaatregelen worden genomen. De doelstelling van dit onderzoek was het ontwikkelen van een praktisch toepasbare netwerkontwerpmethode waarmee het wegennetwerk robuust kan worden gemaakt voor incidenten. De focus van dit onderzoek lag op de robuustheid van de wegennetwerkstructuur. Dit kan worden beschouwd als een eerste stap in het ontwikkelen van een robuust mobiliteitsysteem (inclusief het openbaar vervoer en multimodale verplaatsingen) waarbij ook niet-structuurgerelateerde maatregelen moeten worden genomen.

Om deze ontwerpmethode voor een robuust wegennetwerk te kunnen ontwikkelen is eerst een begrippenkader opgesteld. In dit begrippenkader is robuustheid als volgt gedefinieerd:
Robuustheid is de mate waarin een netwerk kan blijven functioneren onder voorgedefinieerde omstandigheden.

Kwetsbaarheid is het tegenovergestelde van robuustheid. Een netwerk dat kwetsbaar is, is niet robuust en omgekeerd. In principe kan een netwerk robuust worden gemaakt tegen allerlei verstoringen variërend van voorspelbare tot onvoorspelbare verstoringen, van reguliere tot irreguliere verstoringen, van verstoringen met een tijdelijk klein effect tot verstoringen met een langdurig/blijvend groot effect en van verstoringen die een effect hebben op het hele netwerk tot lokale verstoringen. Het is aan de netwerkbeheerders en/of beleidsmakers om te bepalen hoe robuust een netwerk moet worden gemaakt tegen welke verstoringen. Dit bedoelen we met de zinsnede 'voorgedefinieerde omstandigheden' in de definitie. Uiteraard geldt wel dat het de voorkeur verdient om het netwerk robuust te maken tegen zo veel mogelijk verstoringen. Dit is echter niet altijd mogelijk omdat bijvoorbeeld andere investeringen nodig zijn om een netwerk robuust te maken voor aardbevingen dan voor ongelukken. Voor robuustheid is het gevolg van verstoringen van groter belang dan de kans op verstoringen. Het gaat er om dat als een verstoring optreedt, het netwerk blijft functioneren. Desondanks is de kans op een verstoring wel van belang bij het bepalen van een investeringsstrategie voor robuustheidsmaatregelen. In dit onderzoek lag daarom de nadruk op onvoorspelbare verstoringen die regelmatig voorkomen: incidenten.

Wegennetwerken kunnen in principe robuust worden gemaakt tegen allerlei verstoringen. Het is aan de netwerkmanagers en/of beleidsmakers om te bepalen hoe robuust het wegennetwerk moet worden gemaakt tegen welke verstoringen en wat de kosten daarvan mogen zijn. In dit onderzoek lag de nadruk op incidenten.

Om een robuust netwerk te kunnen ontwerpen is een definitie voor robuustheid onvoldoende. De vraag wat een netwerk robuust maakt moet hiervoor ook worden beantwoord. Een vergelijking tussen het wegennetwerk, het spoor- en binnenvaartnetwerk, het bloedvatenstelsel en het zenuwstelsel en het internet heeft aangetoond dat netwerken robuuster worden als er alternatieve routes beschikbaar zijn met voldoende reservecapaciteit. Het beschikken over deze redundantie in het netwerk is niet voldoende. De getransporteerde elementen moeten zo snel mogelijk kunnen worden geherrouteerd naar deze routes. Dit vraagt om flexibiliteit in het netwerk en om informatie-, verkeersmanagement- en incidentmanagementstrategieën. Verder moeten fileterugslageffecten tot een minimum worden beperkt. Tot slot kunnen preventieve maatregelen netwerken robuuster maken. Een aanvullende analyse van de ongelukken die in Zuid-Holland zijn gebeurd onderstreept dat de volgende vijf elementen een netwerk robuust maken: preventie, redundantie, compartimentering, veerkracht en flexibiliteit. Deze vijf elementen geven richting aan de keuze voor een robuustheidindicator en robuustheidsmaatregelen.

Een robuust wegennetwerk is een netwerk waarin preventieve maatregelen worden genomen en dat redundant, gecompartimenteerd, veerkrachtig en flexibel is.

Voor de reiziger is het begrip robuustheid niet relevant. Uiteindelijk is het van belang dat reizigers betrouwbare reistijden hebben en dus geen onverwacht grote vertragingen oplopen als gevolg van incidenten. In het opgestelde begrippenkader hebben we daarom aangegeven hoe robuustheid, reistijden en betrouwbaarheid van reistijd met elkaar samenhangen. We hebben laten zien dat robuustheid en betrouwbaarheid twee concepten zijn die aan elkaar zijn gerelateerd, maar een andere focus hebben. Robuustheid richt zich op de netwerkstructuur en de gevolgen van verstoringen met een onverwacht effect op een specifiek moment in de tijd. Betrouwbaarheid richt zich daarentegen meer op de reiziger en de kans dat bepaalde verwachte verstoringen zich voordoen over een langere periode. In een robuust wegennetwerk leiden verstoringen gemiddeld tot kleinere vertragingen en dus betrouwbaardere reistijden. Dit is echter geen één op één relatie. Robuustheidsmaatregelen hoeven niet tot betrouwbaardere reistijden voor alle reizigers te leiden. Als robuustheidsmaatregelen worden genomen, moet bij de evaluatie van de betrouwbaarheid van de reistijd met veel verschillende verstoringen rekening worden gehouden. Speciale aandacht moet echter gegeven worden aan verstoringen met een onverwacht groot effect, omdat de effecten hiervan in traditionele betrouwbaarheidsanalyses vaak worden onderschat.

Robuustheid en betrouwbaarheid zijn gerelateerde concepten die in samenhang moeten worden geanalyseerd.

Om robuustheid mee te kunnen nemen in het netwerkontwerp is een indicator voor robuustheid nodig die al dan niet een relatie heeft met de betrouwbaarheid van reistijd. We hebben daarom een lijst met indicatoren opgesteld die allemaal naar één of meerder elementen van robuustheid (preventie, redundantie, compartimentering, veerkracht en flexibiliteit) of naar reistijd of reistijdbetrouwbaarheid verwijzen. Uiteindelijk is een keuze gemaakt voor de extra reistijd als gevolg van incidenten, omdat deze indicator het concept robuustheid met bijbehorende elementen het meest volledig beschrijft. Afhankelijk van de wijze waarop de indicator wordt berekend, kan deze indicator rekening houden met fileterugslageffecten (compartimentering), routealternatieven (redundantie) en veerkracht. Als flexibiliteit in het netwerk wordt ingebouwd kan hier ook rekening mee worden gehouden. Het element preventie verwijst naar de capaciteitsreductie als gevolg van incidenten. Hier kan rekening mee worden gehouden in de input van het te gebruiken model. Verder is de indicator goed uitlegbaar en kan de indicator worden gewaardeerd in kosten-batenanalyses. We hebben ook laten zien dat het monitoren van de indicator (de indicator berekenen op basis van data) tot op zekere hoogte mogelijk is, maar wel complex is en nog extra onderzoek vergt. De grootste uitdaging van deze indicator is het opstellen van een model waarmee de indicator binnen korte rekentijd kan worden berekend.

Als indicator voor robuustheid kan het best de extra reistijd als gevolg van incidenten worden gehanteerd, omdat deze indicator het concept robuustheid het meest volledig beschrijft.

Het dilemma bij de modelkeuze voor de evaluatie van robuustheid is dat een keuze moet worden gemaakt tussen nauwkeurigheid (gebruik van een dynamisch toedelingsmodel dat rekening houdt met fileterugslageffecten en met meerdere soorten van routekeuzegedrag) en rekentijd. De meest nauwkeurige modellen hebben over het algemeen een lange rekentijd. Vuistregels nemen nauwelijks rekentijd in beslag, maar zijn niet nauwkeurig. Voor sommige doeleinden kunnen vuistregels goed genoeg zijn om een snelle indruk te krijgen van de robuustheid van een netwerk, maar om een gedegen beslissing te kunnen nemen over robuustheidsmaatregelen is het beter om een model te kiezen met bovengenoemde eigenschappen en een zo kort mogelijke rekentijd, omdat heel veel robuustheidevaluaties benodigd zijn bij netwerkontwerp.

Wij hebben er voor gekozen om het macroscopische dynamische toedelingsmodel Indy te gebruiken om een evenwichtstoedeling onder reguliere omstandigheden uit te voeren. Dit model heeft een nauwkeurige vorm van filemodellering (eerste orde kinematische schokgolf theorie van Newell) waarmee fileterugslageffecten goed kunnen worden gemodelleerd. Het model kan een evenwicht berekenen en kan ook met vaste routes rekenen. En-route routekeuze is niet mogelijk. Dit is echter geen probleem, omdat Indy alleen gebruikt wordt om de 'basisroutekeuze' onder reguliere omstandigheden (zonder verstoringen) te bepalen. Hiervoor is een evenwichtstoedeling het meest geschikt. Aan Indy is een model gekoppeld waarmee de marginale effecten van incidenten kunnen worden berekend (MIC) (Corthout et al., 2009). Binnen enkele minuten kan hiermee het effect van tientallen incidenten worden ingeschat. Vier incidenttypes zijn gedefinieerd waarvan de effecten zijn bepaald door incidenten te simuleren op alle wegen uit netwerk. Het MIC-model veronderstelt dat niemand zijn/haar route wijzigt bij incidenten. Omdat dit in praktijk wel gebeurd is hier een algoritme aan toegevoegd waarmee via een benadering rekening kan worden gehouden met alternatieve routes. De kwaliteit van de evaluatiemethode is aangetoond door een vergelijking te maken met lusdata van vijf incidenten. Tot slot, is aangegeven hoe de betrouwbaarheidsbaten van robuustheidsmaatregelen kunnen worden gekwantificeerd met deze methode ten behoeve van kosten-batenanalyse.

Om de robuustheid van een wegennetwerk te bepalen is een nauwkeurige en snelle evaluatiemethode nodig die rekening kan houden met fileterugslageffecten en alternatieve routes. Hiervoor is een macroscopisch dynamisch toedelingsmodel gecombineerd met een model waarmee de marginale effecten van incidenten kunnen worden bepaald en een algoritme dat bij benadering rekening houdt met alternatieve routes.

Als robuustheidsmaatregelen worden voorgesteld, moet bij de voorspelling van de effecten van die maatregelen rekening worden gehouden met de robuustheids-/betrouwbaarheidsbaten, omdat deze maatregelen bedoeld zijn om de robuustheid van het netwerk en daarmee de betrouwbaarheid van de reistijd te verbeteren. Alleen rekening houden met de reguliere reistijdbaten, zoals vaak wordt gedaan, is dus onvoldoende. Verschillende voorbeelden hebben aangetoond dat de robuustheids-/betrouwbaarheidsbaten significant kunnen zijn: mogelijk van gelijke grootte als de reistijdbaten. Dit betekent dat het kosten-batensaldo van robuustheidsmaatregelen veel positiever kan zijn, dan wanneer de robuustheids/betrouwbaarheidsbaten niet worden meegenomen.

De betrouwbaarheidsbaten van robuustheidsmaatregelen moeten worden meegenomen in kosten-batenanalyses om goed afgewogen investeringsbeslissingen te kunnen nemen.

De vraag is nu welke robuustheidsmaatregelen kunnen worden genomen. Een lijst met maatregelen is gepresenteerd die onderverdeeld zijn in maatregelen die op de verplaatsingenmarkt (maak ik wel of geen verplaatsing en, zo ja, waarheen?), de vervoermarkt (met welke vervoerwijze reis ik?) en de verkeersmarkt (welke route kies ik en wat is de verkeersafwikkeling?). We hebben uitgelegd op welke elementen van robuustheid de maatregelen aangrijpen en voor welke verstoringen ze relevant zijn. Voor een selectie van maatregelen die het best passen bij de focus van dit onderzoek op de netwerkstructuur (alternatieve routes creëren, ontvlechten en het aanleggen van buffers en reservecapaciteit), zijn de mogelijke effecten in kaart gebracht aan de hand van verschillende voorbeelden.

We hebben aangetoond dat in het voorbeeldnetwerk van Rotterdam en omgeving het hebben van één alternatieve route de extra reistijd als gevolg van incidenten met 3,1% tot 77,1% kan verminderen afhankelijk van de informatie die wordt geboden. Extra alternatieve routes geven slechts kleine verbeteringen. Uiteraard is dit afhankelijk van de kwaliteit (reservecapaciteit en reistijd) van de alternatieve routes. Het ontvlechten van wegen kan zowel positief als negatief zijn voor de robuustheid van een netwerk afhankelijk van de wijze en locatie van ontvlechten. Tot slot hebben we laten zien dat buffers de reistijd in potentie kunnen verlagen met 12% in het netwerk van Den Haag en omgeving onder reguliere omstandigheden. In andere netwerken kan dit effect uiteraard anders zijn. Hoe meer congestie er is, hoe hoger de baten waarschijnlijk zijn. Voor incidentsituaties kunnen vergelijkbare conclusies worden getrokken en dus maken buffers een netwerk robuuster. Immers et al. (2004d) hebben eerder al aangetoond dat de optimale reservecapaciteit voor een snelweg met drie rijstroken ongeveer 800 (pcu/uur) is. Hierbij is, in tegenstelling tot de methode die in dit proefschrift is gebruikt, nog geen rekening gehouden met fileterugslageffecten.

Veel verschillende maatregelen kunnen worden genomen om de robuustheid van een wegennetwerk te verbeteren. Met betrekking tot de structuur van een netwerk hebben het toevoegen van reservecapaciteit, het creëren routealternatieven die gelijkmatig zijn verdeeld over het netwerk, het ontvlechten van wegen en het toevoegen van buffers de hoogste potentie. De baten van deze maatregelen hangen echter af van locatiespecifieke omstandigheden en kunnen dus niet worden benaderd door vuistregels of regressieanalyses die geen rekening houden met deze locatiespecifieke factoren.

De volgende vraag die opkomt, is op welke plek het beste welke maatregelen kunnen worden genomen. Dit zogenaamde robuuste Netwerk-Ontwerp-Probleem is zowel vanuit theoretisch als praktisch oogpunt complex, omdat er veel belanghebbenden zijn die allemaal andere deels overlappende en deels tegenstrijdige doelstellingen hebben. Bovendien kunnen veel verschillende soorten verstoringen op veel verschillende plekken optreden onder verschillende omstandigheden waar verschillende groepen gebruikers anders op reageren. Tot slot kunnen ook nog eens veel verschillende maatregelen worden genomen. Wij hebben het robuuste Netwerk-Ontwerp-Probleem geformuleerd als een 'bi-level'-probleem waarbij op het bovenste niveau het netwerk wordt ontworpen en op het onderste niveau de routekeuze, vervoerwijzekeuze en bestemmingskeuze worden gemaakt afhankelijk van de netwerkstructuur. De keuzes op het onderste niveau zijn uiteraard weer van invloed op de keuzes op het bovenste niveau. De gepresenteerde formulering verschilt van de formuleringen van het Netwerk-Ontwerp-Probleem in de literatuur, doordat het bestemmingskeuze, vervoerwijzekeuze en dynamische routekeuze op het onderste niveau combineert en rekening houdt met korttermijnverstoringen (incidenten) aan de aanbodzijde bij het ontwerp.

De ontwikkelde methode combineert expertkennis met geavanceerde modelleringstechnieken waardoor het mogelijk is om grote realistisch robuuste wegennetwerken te ontwerpen waarin veel ontwerpvariabelen (=robuustheidsmaatregelen) zijn opgenomen en dat wordt gedragen door alle partijen die betrokken zijn bij het ontwerpproces. Het feit dat experts (onderzoekers, belanghebbenden etc.) betrokken zijn in het ontwerpproces is een voordeel omdat hierdoor de invoer en uitvoer van het model wordt verrijkt met expertkennis. Bovendien kunnen hierdoor veel verschillende robuustheidsmaatregelen in het optimalisatieproces worden meegenomen. In een optimalisatieproces waarin alleen een model gebruikt wordt, is dit heel moeilijk zo niet onmogelijk. Tot slot, kan het nieuwe ontwerp waarschijnlijk op meer draagvlak rekenen bij alle betrokkenen. Het feit dat gebruik wordt gemaakt van modellen voor evaluatie en optimalisatie is eveneens een voordeel omdat hierdoor de kwetsbaarheid van verschillende wegennetwerkontwerpen kwantitatief kan worden bepaald. De optimalisatiemodellen geven suggesties over hoeveel reservecapaciteit waar nodig is. De modeluitkomsten helpen vervolgens de experts bij het ontwerpen van een robuust wegennetwerk. Tot slot geven de modellen een kwantitatieve onderbouwing bij het ontwerp.

De ontwerpmethode bestaat uit drie stappen die hieronder kort worden toegelicht:
- Stap 1: ontwerpstandaarden. Het ontwerpen van een wegennetwerk begint met het specificeren van de ontwerpstandaarden door de experts. De doelstelling van het ontwerp, de ontwerpvariabelen en de restricties worden bepaald.
- Stap 2: functionele analyse. Een functionele analyse is een analyse waarbij de prestatie van het referentienetwerk (=uitgangspunt) wordt geëvalueerd. Dit gebeurt op basis van expertkennis van de verschillende wegen in combinatie met het evaluatiemodel. Het evaluatiemodel bepaalt eerst de reistijden en congestie onder reguliere omstandigheden en vervolgens voor incidentsituaties. Bij de functionele analyse wordt zoveel mogelijk rekening gehouden met verschillende wegtypes en verschillende soorten verplaatsingen.
- Stap 3: ontwerpproces. In deze stap vindt het ontwerp plaats. De ontwerpstandaarden en de resultaten van de functionele analyse vormen hiervoor het uitgangspunt. Een optimalisatie model wordt gebruikt om te bepalen op welke locaties het best reservecapaciteit kan worden toegevoegd. Hierbij wordt het eerst het ‘optimale’ aantal extra rijstroken bepaald voor de reguliere situatie. Daarna wordt het ‘optimale’ aantal stroken bepaald die in de vorm van reservecapaciteit het netwerk robuuster maken. Na deze modelexercitie krijgen experts de gelegenheid om op basis van de modeluitkomsten het ontwerp aan te passen. Zij kunnen bijvoorbeeld aangeven waar alternatieve routes moeten worden gecreëerd en waar allerlei andere robuustheidsmaatregelen genomen moeten worden. Om de experts hierbij te helpen zijn enkele ontwerpprincipes gepresenteerd.

Het ontworpen netwerk wordt tot slot geëvalueerd met het eerder genoemde evaluatiemodel. Als blijkt dat er daarna nog problemen in het netwerk zijn, kan stap 3 worden herhaald.

Door expertkennis te combineren met geavanceerde evaluatie- en optimalisatiemodellen kunnen grote realistische robuuste wegennetwerken worden ontworpen met veel verschillende ontwerpvariabelen (= robuustheidsmaatregelen). Dit ontwerp kan rekenen op draagvlak bij de betrokken experts/belanghebbenden.

De ontwerpmethode is uitgebreid getest op een klein testnetwerk om de kwaliteit en de gevoeligheid te toetsen. Deze test heeft uitgewezen dat het globale optimum werd gevonden voor het kleine testnetwerk. Dankzij enkele vereenvoudigingen in het algoritme blijft het aantal keren dat een volledige dynamische modelrun uitgevoerd hoeft te worden beperkt waardoor het mogelijk is om de ontwerpmethode ook toe te passen op grotere netwerken.

De ontwerpmethode is ook toegepast op een realistisch netwerk van Amsterdam en omgeving. Voor zover bekend bij de auteur is dit de eerste praktische toepassing waarbij een combinatie van expertkennis (o.a. van belanghebbenden) en optimalisatiemodellen is gebruikt om een robuust wegennetwerk te ontwerpen. Bovendien is dit de eerste toepassing waarbij in het optimalisatiemodel een dynamisch toedelingsmodel in combinatie met een model dat de marginale effecten van incidenten bepaalt, is gebruikt. De toepassing van het evaluatiemodel waarbij met fileterugslag en alternatieve routes rekening wordt gehouden is op zich al een stap voorwaarts bij de bepaling van de robuustheid van een netwerk. De toepassing op een groot netwerk heeft verschillende inzichten geboden in de wijze waarop robuustheidevaluatie en netwerkontwerp in praktijk kunnen worden gedaan. Deze inzichten hebben betrekking op de keuzes van het detailniveau van het netwerk, de kalibratie, de simulatie van incidenten, de parameterkeuze van de modellen en de te verwachten resultaten. Uit de toepassing van de modelstappen uit de ontwerpmethode is gebleken dat de netwerkprestatie veel kan worden verbeterd met een positieve kosten-batenverhouding. De totale reiskosten onder reguliere omstandigheden zijn gereduceerd met 12% en de totale reiskosten onder irreguliere omstandigheden (incidenten) zijn gereduceerd met 97% en dus bijna gereduceerd tot nul. Deze resultaten zijn behaald met slechts enkele iteraties van het optimalisatiemodel. Dit geeft aan dat waarschijnlijk nog betere oplossingen gevonden kunnen worden.

Tot slot is een beeld geschetst van hoe expertinbreng het netwerkontwerp nog robuuster kan maken door naast optimalisatie van de reservecapaciteit aanvullende robuustheidsmaatregelen te nemen.

De robuustheid van veel delen van het wegennetwerk kan worden verbeterd met een positieve baten-kostenverhouding. Voor een groot netwerk van Amsterdam en omgeving hebben voorbeeldberekeningen aangetoond dat door optimalisatie van de capaciteit de reiskosten onder reguliere omstandigheden met 12% kunnen worden verbeterd en de reiskosten onder irreguliere omstandigheden bijna kunnen worden gereduceerd tot nul (reductie van 97%).

Robuustheidsmaatregelen kunnen tot een toename van de vraag leiden als de reistijd onder reguliere omstandigheden verbetert. Deze toename in de vraag was niet beschouwd in het netwerk van Amsterdam, maar wel in het testnetwerk. Uit de analyses voor het testnetwerk is gebleken dat door de toename in de vraag de verwachte robuustheidverbeteringen als gevolg van de robuustheidsmaatregelen niet volledig werd behaald. Een verbeterde robuustheid en een toename van de vraag kunnen beide tot positieve welvaartseffecten leiden en dus moet een goede afweging tussen beide gemaakt worden. De gepresenteerde ontwerpmethode maakt dit mogelijk.

Het is nodig om een goede afweging te maken tussen het verbeteren van de reistijden onder reguliere omstandigheden en de bijbehorende mobiliteitsgenererende effecten en robuustheidsmaatregelen en benuttingsmaatregelen.

De resultaten van dit onderzoek hebben verschillende implicaties voor beleidsmakers, netwerkmanagers, weggebruikers en andere belanghebbenden. Voor beleidsmakers geeft dit onderzoek het belang van robuustheid aan en het legitimeert investeringen in robuustheidsmaatregelen. Het onderzoek toont aan dat het nodig is om bij de evaluatie van maatregelen niet alleen naar de gebruikelijke posten te kijken (reistijden onder reguliere omstandigheden, externe kosten, etc.), maar dat het ook nodig is om de betrouwbaarheidsbaten van robuustheidsmaatregelen mee te nemen. We hebben ook laten zien dat verschillende maatregelen tegenstrijdige effecten kunnen hebben. Het maximaal benutten van het netwerk onder reguliere omstandigheden resulteert bijvoorbeeld in minder tot geen reservecapaciteit die ingezet kan worden bij verstoringen. Er moet dus een afweging worden gemaakt tussen benutten en robuustheid. Daarnaast moet structuurgerelateerde robuustheidsmaatregelen worden aangevuld met adequate informatieverschaffing en de juiste DVM- en incidentmanagementstrategiëen om zo goed mogelijk gebruik te maken van een robuustere structuur bij verstoringen. Industrieën kunnen hier ook een rol bij spelen via intelligente transportsystemen, zoals navigatiesystemen.

Voor onderzoekers en consultants zijn in dit onderzoek veel handvatten geboden waarmee robuustheidsanalyses kunnen worden uitgevoerd en waarmee robuuste netwerken kunnen worden ontworpen. Er is een rol voor onderzoekers en consultants om er voor te zorgen dat robuustheidsanalyse een standaardonderdeel wordt bij de evaluatie van maatregelen en dat ook met robuustheid rekening wordt gehouden bij het ontwerp. Verder zijn enkele suggesties voor vervolgonderzoeken opgenomen die door hen kunnen worden opgepakt. Voor andere belanghebbenden, zoals belangenverenigingen, is er een taak om zich in het onderwerp robuustheid te verdiepen (voor zover dat nog niet is gedaan), daar een mening over te vormen en daar waar mogelijk te helpen bij het realiseren van een robuust wegennetwerk.

Tot slot: dit onderzoek heeft de intentie bij te dragen aan de realisatie van een robuuster wegennetwerk. Het zijn uiteindelijk de weggebruikers die daarvan profiteren, doordat zij minder vaak met oververwacht grote vertragingen worden geconfronteerd.

- Maaike Snelder, 1 december 2010

About the author

Maaike Snelder was born on 28 August 1980 in Rotterdam, the Netherlands. In 1998 Maaike graduated from her high school "The Caland Lyceum" in Rotterdam. Thereafter, she started her study econometrics at the Erasmus University in Rotterdam. What fascinates her in econometrics is that mathematics can be used to improve and optimize real-life situations. Econometrics can be applied in all kind of fields. During her study transport problems interested her most, because we are all daily confronted with transport issues. Maaike did a so-called "Duale Master" in which she had the opportunity to work two times six months at a company in order to enrich her professional skills before she graduated. She spent these six months in 2002 and 2003 at TNO at the department of Traffic and Transport (later called Mobility and Logistics). This was also the place where she worked on her master thesis in which she made an optimal redesign of the Dutch road network which made it a few years later to almost all newspapers and several journals in the Netherlands. In October 2003 she graduated.

After her graduation, Maaike stayed at TNO where she worked half time at the department of Traffic and Transport and half time at the department of Logistics. She specialized in traffic and transport models for both passenger and freight transport. She (co-) developed and/or applied the Optimization Algorithm for designing Road Networks (ROADNet), the Strategic Model for Analyzing the Reliability of Accessibility (SMARA), the strategic traffic and transport model SMART, the European transport model TRANS-TOOLS, the dynamic traffic assignment model Indy and the general spatial equilibrium model RAEM. Furthermore, she was involved as a project member and as a project leader in many projects about robust road network design and the evaluation of robustness and robustness/reliability benefits for many different clients. In 2009 Maaike was awarded with the price of most excellent young researcher of TNO.

In 2005 Maaike was asked by her two promoters Henk van Zuylen and Ben Immers to start a PhD research at the faculty of Civil Engineering, department of Transport and Planning of the Delft University about robust road network design. In August 2005 she actually started with this research in a part-time construction in which she had contract at the Delft University for 1 day a week and a 36 hour contract at TNO. Maaike is also a member of the TRAIL research school. During her PhD-study Maaike wrote many papers about robust road network design and she participated in many conferences as well.

In 2009 Maaike visited Professor Mike Florian at the Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) in Montréal, Canada, where she made a comparison between two dynamic traffic assignment models: Indy and Dynameq.

Maaike will continue to work on traffic and transport problems at TNO, where she hopes to contribute to a mobility system that has a high quality, is robust and therewith offers short and reliable travel times in such a way that we all can go to the places where we want to be.

Author's publications

Book chapters

Immers, L.H., B. Egeter, M. Snelder, C. Tampere (to be published) Reliability of travel times and robustness of transport networks, in: Editor M. Kutz (eds.) *Transportation Engineers' Handbook*, McGraw-Hill, United States of America.

Snelder, M., Dat gaat zo echt niet langer met het Nederlandse wegennetwerk!, in: Editors K. Ruijgrok, B. Jansen, L.A. Tavasszay (eds.) *Ik Ben Mister … Immers, Liber amicorum voor Ben Immers*

Papers

Snelder, M. L.A. Tavasszy (2010) Quantifying the full reliability benefits of road network improvements, in: *Proceedings of WCTR 2010*, Lisbon, Portugal.

Immers, L.H., M. Snelder, K. Ruijgrok (2010) Design principles for improving the robustness of networks, in: *Proceedings of the 4th International Symposium on Transportation Network Reliability*, Minnesota, United States of America.

Yperman, I., Snelder M. (2010) Evaluating network performance using different dynamic traffic models within a DTA framework, in: *Proceedings of the Third International Symposium on Dynamic Traffic Assignment*, Takayama, Japan.

Knoop, V.L., M. Snelder, H.J. van Zuylen, S.P. Hoogendoorn (2010) Link-level Vulnerability Indicators for Real-World Networks, in: *Proceedings of the 89th Annual Meeting of the Transportation Research Board*, Washington D.C., United States of America, Accepted for publication.

Egeter, B., T. Hendriks, M. Snelder (2010), ANWB-visie Robuust P+R, in: *Proceedings Colloquium Vervoersplanalogisch Speurwerk*, Roermond, The Netherlands.

Snelder, M., J.M. Schrijver, L.H. Immers (2009) Nietsdoen is geen optie - methodes voor het bepalen van de robuustheid van wegennetwerken, in: *Proceedings Colloquium Vervoersplanalogisch Speurwerk*, Antwerpen, Belgium.

Immers, L.H., M. Snelder, J.M. Schrijver, B. Egeter (2009), Architecture for Robust Network Design, in: *Proceedings of the 1st NECTAR Transatlantic Conference*, Arlington, United States of America.

Snelder, M., J.M. Schrijver, L.H. Immers, B. Egeter (2009) Architecture for Designing Robust Road Networks, *Proceedings of the 88th Annual Meeting of the Transportation Research Board*, Washington D.C., United States of America.

Snelder, M., J.M. Schrijver (2008) De toekomst voorspeld in de toekomst: files realistisch gemodelleerd, in: *Proceedings Colloquium Vervoersplanalogisch Speurwerk*, Santpoort, The Netherlands.

Snelder, M., L.H. Immers, H.J. van Zuylen (2008) The Importance of Alternative Routes for the Robustness of a Road Network, in: *Proceedings of IEEE Conference on Infrastructure Systems - Building Networks for a Better Future*, Rotterdam, the Netherlands.

Snelder, M., L.A. Tavasszy, L.H. Immers, H.J. van Zuylen (2008) The vulnerability of road networks, Now and in the future, in: *Proceedings of the TRAIL-congres*, Delft, the Netherlands.

Snelder, M., A. Wagelmans, J.M. Schrijver, H.J. van Zuylen, L.H. Immers (2007), Optimal Redesign Of The Dutch Road Network, in: *Transportation Research Record: Journal of the Transportation Research Board*, 2029, pp. 72-79.

Knoop, V.L., M. Snelder (2007) Comparison of Link-level robustness indicators, *Proceedings of the 3rd International Symposium on Transportation Network Reliability*, Delft, The Netherlands.

Snelder, M. (2007) Een eigen blik op de weg: een vergelijking tussen objectief wetenschappelijk onderzoek en de weergave van dit onderzoek in de media, in: *Proceedings Colloquium Vervoersplanalogisch Speurwerk*, Antwerp, Belgium.

Snelder, M., A. Wagelmans, J.M. Schrijver, H.J. van Zuylen, L.H. Immers (2006) Optimal Redesign of the Dutch Road Network, *Proceedings of the TRAIL-congres*, Delft, the Netherlands.

Verweij, C.A., M. Snelder (2006) Scherper schatten met fijnere kengetallen, in: *Verkeerskunde*, 3, pp. 58-62.

Snelder, M., A. Burgess (2005) TRANS-TOOLS: TOOLS for TRansport forecasting And Scenario testing: A focus on the logistic module, in: *Proceedings of the Vervoerslogistieke Werkdagen*, Hoeven, the Netherlands, pp. 156-163.

Immers, L.H., A. Bleukx, M. Snelder, J. Stada (2004) Betrouwbaarheid van reistijden: het belang van netwerkstructuur en optimale wegcapaciteit, in: *Tijdschrift Vervoerswetenschap*, 40(4), pp. 18-28.

Snelder, M., L.H. Immers, I.R. Wilmink (2004) De begrippen betrouwbaarheid en robuustheid nader verklaard, *Proceedings Colloquium Vervoersplanalogisch Speurwerk*, Zeist, the Netherlands, pp. 741-760.

Reports

Duijnisveld, M., M. Snelder, M. Minderhoud, S. Calvert (to be published) *Kalibratieproces verkeersmodel verkeershinder A10-A1,* TNO-rapport, Delft, the Netherlands.

Egeter, B., M. Snelder (2010) ARKO-werkboek, *Een ontwikkelingsbeeld voor het wegennet in de Amsterdamse region op basis van de ARKO-methodiek*, Bart Egeter Advies, Rotterdam, the Netherlands.

Snelder, M., B. Egeter, T. van Rooijen, L.H. Immers (2010) *✱P+R, ANWB-visie op een nieuw type P+R, Uitgewerkt voor de regio Rotterdam - Den Haag*, TNO-034-DTM-2010-03924, Delft, the Netherlands.

Snelder, M., J. Mak, L.A. Tavasszy (2010) *Kwetsbaarheidsanalyse NMCA*, TNO, TNO-034-DTM-2010-01984, Delft, the Netherlands.

Van Meijeren, J., M. Snelder, T. van Rooijen, B. Egeter (2010) *Bereikbaarheidsvisie Mainports 2040, hoofdrapport en achtergrondrapport*, TNO, Delft, the Netherlands.

Snelder, M., J.M. Schrijver, R. Smokers, I. Wilmink, N. Ligterink, R. Sharpe, V. Swinkels, D. van de Wall Bake (2010) *Effecten van de MobiliteitsAanpak op de uitstoot van CO2, hoofdrapport en bijlagenrapport*, TNO-034-DTM-2009-02026, Delft, the Netherlands.

Snelder, M., M. Muller, J.M. Schrijver (2010) *Begrippenkader en indicatoren voor robuustheid*, TNO-034-DTM-2009-05026, Delft, The Netherlands.

Snelder, M. (2009) *A Comparison between Dynameq and Indy*, Cirrelt-2009-48, Montréal, Canada.

Snelder, M. L.A. Tavasszy, T. van Rooijen, W. Spit, E. Devillers, E. Buckman (2009) *Kansrijke combinaties van de concepten uit de MobiliteitsAanpak: Een economische toets en een uitwerking van de concepten naar gebiedstype*, TNO and ECORYS, TNO-034-DTM-2009-02313, Delft, the Netherlands.

Snelder, M. (2009) *Kwetsbaarheidsanalyse wegennetwerk Almelo*, TNO, TNO-034-DTM-2009-03938, Delft, the Netherlands.

Snelder, M., M. de Kievit, J.M. Schrijver (2009) *Internationale vergelijking van het gebruik van modellen in het infrastructuurplanningsproces*, TNO, 2008-D-R1301/A, Delft, the Netherlands.

Snelder, M, J.M. Schrijver, R. Landman, J. Mak, M. Minderhoud (2008) *De kwetsbaarheid van Randstedelijke vervoernetwerken uit verkeerskundig perspectief*, TNO, 2008-D-R0882/A, Delft, the Netherlands.

Schrijver, J.M., B. Egeter, L.H. Immers, M. Snelder (2008) *Visie robuust wegennet ANWB*, TNO, 2008-D-R0661/C, Delft, the Netherlands.

Schrijver, J.M., M. Snelder (2008) *Gemeenschappelijke conclusies aanpak A1*, TNO, 200808.34.15/N118/034.20058/JS/LK, Delft, the Netherlands.

Snelder, M., J.M. Schrijver, T. van Rooijen (2008) *Naar een klimaatbestendig wegennetwerk. De robuustheid van het wegennetwerk voor verstoringen die door klimaatveranderingen vaker voor zullen komen,* TNO, Delft, the Netherlands.

Snelder, M., O. Koops, O. Ivanova, L.A. Tavasszy (2008) *Indirecte effecten KBA Schiphol-Amsterdam-Almere*, TNO, 2008-D-R0248/A, Delft, the Netherlands.

Tavasszy, L.A., M. Snelder (2008) *Naar een robuust wegennet: probleemdefinitie en oplossingen*, essay for the Ministry of Transport, Public Works and Water Management, TNO, Delft, the Netherlands.

Klunder, G., J. Mak, M. Snelder, L.H. Immers (2007) *Verkeerssimulaties van grootschalige evacuatie van de Randstad*, TNO, 2007-D-R1323/B, Delft, the Netherlands.

Snelder, M., O. Koops (2007) *RAEM-berekeningen case personenvervoer. Case studie indirecte effecten van investeringen in infrastructuur*, TNO, Delft, the Netherlands.

Snelder, M. (2006) *Quick-scan van de omvang van het vermeden wegverkeer door de omlegging van de Zuid-Willemsvaart*, TNO, 06.34.15/N185/034.65129/MS/MM, Delft, the Netherlands.

Schrijver, J.M., L.H. Immers, M. Snelder, R. de Jong (2006) *Effecten van de landelijke invoering van incidentmanagementmaatregelen op de voertuigverliestijd in het netwerk*, TNO, Delft, the Netherlands.

Bovenkerk, M., M. Snelder, M. Jordans, A. Burgess (2006) *Validatie Smile+ 2.0*, TNO, 2006-75, Delft, the Netherlands.

Kuipers, B., M. Snelder, A. van der Ham (2006) *De betekenis van betrouwbaarheid van de reistijd voor de logistieke organisatie van verladers en vervoerders*, TNO, 2006-D-R-0325, Delft, the Netherlands.

Louter, P.J., W. van Eikeren, M. Bovenkerk, Y. de Kluizenaar, W.P.N. van der Laan, J.M. Schrijver, M. Snelder (2005) *Maatschappelijke waarde analyse Mainport Schiphol*, Bureau Louter en TNO, 04026, Delft, the Netherlands.

Kuipers, B., A. Burgess, W.J.J. Manshanden, C.A. Verweij, A.C. Muskens, M. Rustenburg, M. Snelder (2005) *De economische betekenis van het goederenwegvervoer*, TNO, 2004-49 and 2004-61, Delft, the Netherlands.

Henstra, D., M. Bovenkerk, J. Muskens, J. Schrijver, M. Snelder, M. Thissen (2004) *Integrale effectbepaling van de hervorming van beprijzingssystemen in het verkeer,* TNO, TNO Inro rapport 2004-62, Delft, the Netherlands.

Burgess, A., K. Uil, M. Bovenkerk, M. Snelder (2004) *Calibratie Transportblok SMILE+,* TNO, 2004-36, Delft, the Netherlands.

Meeuwissen, A.M.H., M. Snelder, J.M. Schrijver (2004) *SMARA: Statistische analyse variabiliteit reistijden*, TNO, Inro-V+V/2004-31, Delft, the Netherlands.

Snelder, M. (2003) *ROADnet: Optimalisatiealgoritme voor het ontwerpen van wegennetwerken: methode en toepassing*, TNO, Delft, the Netherlands.

Designing Robust Road Networks

A general design method applied to the Netherlands

The Dutch road network is, like many other road networks in the world, congested in the morning and evening peaks. The locations of congestion are quite often the same; this makes it relatively easy to take the delay of this regular congestion into account when planning a trip. However, as a result of disturbances, also unexpectedly large delays occur. If no measures are taken, the Dutch road network, especially in major urban areas, is becoming more and more vulnerable to disturbances like incidents.

This PhD study proposes a framework for robustness analysis that includes definitions, indicators and a set of measures that can be applied to make the road network robust against incidents. Furthermore, a method is developed by which robust road networks can be designed given these measures. The method combines expert knowledge with advanced modelling techniques. The quality of the method is proven by applying it to a small test network. Finally, this thesis shows how the method can be applied to a large realistic network of Amsterdam and surroundings.

TNO
Schoemakerstraat 97
NL-2628 VK Delft
The Netherlands

T +31 15 269 69 69
F +31 15 261 24 03
E infodesk@tno.nl

tno.nl

ISBN 978-90-5986-370-5

TNO | Innovation for Life