

# THE BEST OF TWO WORLDS - A ROBUST ROAD NETWORK DESIGN METHOD BASED ON AN OPTIMIZATION MODEL AND EXPERT JUDGEMENT

Maaïke SNELDER<sup>a,b,c</sup>, Ben IMMERS<sup>a,c,d</sup> and Henk Van ZUYLEN<sup>a, c, e</sup>

<sup>a</sup> *Delft University of Technology, Faculty of Civil Engineering and Geosciences,  
Transport and Planning, The Netherlands*

*Email: m.snelder@tudelft.nl*

<sup>b</sup> *TNO, the Netherlands*

<sup>c</sup> *TRAIL research school*

<sup>d</sup> *Catholic University of Leuven, CIB, Traffic & Infrastructure, Belgium*

<sup>e</sup> *Hunan University, Faculty of Civil Engineering, China*

## ABSTRACT

In many road networks all over the world unexpectedly large delays occur as a result of unforeseen disturbances like incidents. Robustness measures can be taken to reduce these costs. In this paper we focus on network structure related robustness measures like adding spare capacity, creating parallel routes that are spread in a balanced way over the network, adding buffers and unbundling. A network design method for designing road networks that are robust against incidents is proposed which is applicable to real sized networks and can be used in practice. The method combines optimization and evaluation models with expert input. An application to a small test network shows the quality and the sensitivity of the method. The design method is also applied to a large realistic network of Amsterdam and surroundings which shows that large improvements in the network performance can be realised with a positive benefit-cost balance.

## 1. INTRODUCTION

Road networks all over the world are vulnerable (= not robust) for disturbances with a temporary or permanent impact. This is especially the case for road networks in urban areas and between closely related urban areas, because of the fact that these road networks are often heavily used. Small or large disturbances can have a major impact on the accessibility of urban areas.

In this paper, robustness is defined as the extent to which, under pre-specified circumstances (disturbances), 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 (Snelder et al., 2012). A link or road section is vulnerable when there is a probability of disturbances on that link having a large impact for instance in terms of travel time delays. With pre-specified circumstances we refer to the fact that, in principle, a network can be made robust against all kinds of disturbances. However, in practice, choices have to be made by policy makers about the disturbances on which they want to focus. There are many ways to classify disturbances. In the literature, a distinction is often made between regular and non-regular disturbances and between predictable and non-predictable disturbances. Furthermore, a distinction can be made between 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. In this paper we will focus on incidents since they occur regularly and can have a large effect which results in unexpectedly high travel times and travel costs. Therefore it is likely that robustness measures for incidents will be beneficial.

The question that remains is where in the network which robustness measures can best be taken. This so-called robust road network design problem (NDP) is a complex problem from a practical and theoretical point of view. It has been recognized as one of the most difficult and challenging problems in transport (Yang and Bell, 1998). Most often, the objective is to minimize the total travel time or the total travel cost under regular conditions subject to a budget restriction on infrastructure costs. When reliability or robustness is added to the problem, it becomes even more complex. Several advances in this direction have been made (e.g. Chootinan et al., 2005; Sumalee et al., 2006; Ukkusuri et al., 2007). However, due to the complexity of the problem and the required computational efforts, it remains a problem to apply the methods to large networks, to take into account different possible robustness measures and to evaluate the design in the lower level by taking the relevant traffic dynamics into account that occur as a result of incidents (e.g. spillback effects, alternative route choice). Finally, many different stakeholders are involved that all have different partly overlapping and partly contradicting objectives. In this paper we present a design method that overcomes many of the above mentioned shortcomings. A network design method for designing road networks that are robust against incidents is proposed which is applicable to real sized networks and can be used in practice.

In Section 2 a formulation is presented of the robust network design problem. Section 3 presents the network design method. In section 4, two case studies are presented. Finally, the conclusions and recommendations are presented in section 5.

## 2. ROBUST NETWORK DESIGN PROBLEM FORMULATION

This section starts with an explanation of the characteristics of a robust road network and measures that can be taken to make a road network more robust. In section 2.2 the robust network design problem is formulated. Solving this problem shows which robustness measures can best be taken at which location in the network.

### 2.1. Robustness Measures

The vulnerability of networks can be analysed by using complex network theory or graph theory as is explained by Boccaletti et al. (2006). 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). However they did not consider the dynamics of traffic. In (Boccaletti et al., 2006), (Nagurney and Qiang, 2009), and (Bornholdt and Schuster, 2003), the vulnerability of different networks is described not only from a static but also from a dynamic point of view. Bornholdt and Schuster conclude for transportation networks that these networks are particularly interesting and complex, since the one-dimensional dynamics on the links (the occurrence and solving of congestion) interacts with the network aspects (e.g. spillback effects) and since car drivers are intelligent and learning in their route choice, mode choice, or activity generation. In order to be able to decide which robustness measures can best be taken a deeper understanding is needed of what makes road networks robust. We showed in (Snelder et al., 2012) that a road 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 vehicles 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.

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. In Klem et al. (2004) a few examples can be found. Below, we present a, not necessarily complete, list of measures that influence the robustness of the road network (including the measures from Klem et al. (2004)) classified in terms of the elements of

robustness on which they have the most influence ( prevention, redundancy, compartmentalization, resilience, and flexibility):

#### *Prevention*

1. Construct a road high enough to make sure that traffic can continue to move in times of *floods*.
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.
3. Use more porous asphalt (ZOAB) to reduce the capacity reduction during *rainfall*.
4. Spread salt on roads so that glazed *frost* and *snow* lead less quickly to problems.

#### *Redundancy*

5. *Spreading out or concentrating activities*. Both contrasting strategies can have a positive or a negative effect on robustness.
6. *Better alignment, abolition, or liberalization of time windows*. The liberalization or abolition of time windows, leads to a more balanced distribution of trips 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.
7. *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.
8. *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.
9. *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.
10. *Introduce more robustness in the logistic chain of goods*, for example through extra stocking of (bottleneck) products. 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.
11. *Adding (spare) capacity*.
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.

### *Compartmentalization*

13. *Introduce buffers to prevent congestion spillback and to regulate the inflow of traffic.*

### *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.

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).

### *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.

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.

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.

Some of the above-mentioned measures have both advantages and disadvantages. This means that customized measures must be formulated. In the remainder of this paper we focus on the following structure related robustness measures:

- Adding spare capacity on a road.
- Create equivalent routes whereby back-up options have the same capacity as major routes.
- Separate through traffic and urban/regional traffic through ‘unbundling’
- Introduce buffers to prevent congestion spillback and to regulate the inflow of traffic.

The effectiveness of these measures has already been shown in (Immers et al. (2004); Schrijver et al. (2008); Snelder (2010a) and Snelder et al. (2010b)).

## **2.2. 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 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. The NDP is not always formulated as a bi-level program. Meng et al. (2001) transferred the bi-level program of the

continuous NDP into a single level, continuously differentiable, but still non-convex, optimization problem. In line with the mainstream of literature, we formulate the robust road network design problem as a bi-level problem. The formulation we present below differs from formulations of the NDP in the literature, because it combines destination choice and mode choice (trip choice), and dynamic route choice in the lower-level problem, and it includes short-term variations in supply caused by incidents in the NDP.

(Top level)

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

s.t.  $G_1(\mathbf{f}, \mathbf{I}^{\text{new}}, \mathbf{D}) \leq 0$

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

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

(Lower level : route choice)

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

s.t.  $G_2(\mathbf{f}, \mathbf{I}^{\text{new}}, \mathbf{D}) \leq 0$

(Lower level : trip choice)

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

s.t.  $G_3(\mathbf{f}, \mathbf{I}^{\text{new}}, \mathbf{D}) \leq 0$

In the above formulation,  $Z_1$  is the objective function of the top level problem and  $\mathbf{I}^{\text{new}}$  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. Hereafter, the problem formulation is discussed in more detail.

## Top level

### *Decision variables*

We chose to use a discrete formulation of the NDP. This implies that the number of lanes that have to be added to each link,  $\mathbf{I}^{\text{new}}$ , 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  $\mathbf{I}^{\text{new}}$  is larger than 0, a new link is constructed. By using this discrete decision variable, the following robustness measures can be modelled:

- Adding spare capacity on a road. Spare capacity is added by adding additional lanes.
- Create equivalent routes whereby back-up options have the same capacity as major routes. Spare capacity is added by adding additional lanes to existing alternative routes, or new additional routes can be created.
- Introducing buffers. Buffers that prevent spillback can be modelled by including very short links at locations that are logical for buffers.

In section 3 and 4, 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. ‘Unbundling’ can also be done in the design process by expert judgement.

### Objective function

Below, the objective function  $Z_l$  is shown. The objective is to maximize the generalized travel costs benefits under regular conditions (= total travel time cost benefits  $TTCB$  + total distance related cost benefits  $TDCB$ ) plus the reliability/robustness benefits ( $TCVB$ ) minus the infrastructure costs, including maintenance costs ( $TCI$ ).

$$\max_{\mathbf{I}^{\text{new}}} Z_l(\mathbf{f}, \mathbf{I}^{\text{new}}, \mathbf{D}) = TTCB(\mathbf{f}, \mathbf{I}^{\text{new}}, \mathbf{D}) + TDCB(\mathbf{f}, \mathbf{I}^{\text{new}}, \mathbf{D}) + TCVB(\mathbf{f}, \mathbf{I}^{\text{new}}, \mathbf{D}) - TCI(\mathbf{f}, \mathbf{I}^{\text{new}}, \mathbf{D}) \quad (1)$$

The total costs of vulnerability  $TCV$  are a summation of the travel time losses caused by the expected number of incidents that happen on a link. They are multiplied by an average value of robustness. The benefits are computed by using the rule of half on the OD-level, in which the travel time, distance related and vulnerability costs in the adjusted network are compared to the costs in the original network by taking into account the changes in demand and by taking the net present value of the costs.

### Constraints

Several constraints can be added to the NDP. One often added constraint is a budget constraint. 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 section we present a solution algorithm in which the budget constraint can easily be added. 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 should 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.

### Lower Level: Route Choice

At the lower level, the traffic flows are determined given the number of lanes and the number of new lanes that result from the top level NDP, and given the demand that results from the lower level NDP (trip choice), which is presented below. We choose to use a stochastic dynamic formulation of the traffic assignment problem instead of the static formulation, because traffic flow dynamics are required when robustness issues are considered. The formulation can be found in (Bliemer, 2005).

### Lower Level: OD Demand – Trip Choice

There is a strong interaction between the demand pattern and the quality of the infrastructure. On 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 depend on the quality of the infrastructure and the services offered. The interaction between future supply and future demand is not often considered 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. However, to the best of our knowledge, there are no examples

of elastic demand in the design of large networks. We chose to include demand elasticity in the design problem by using a gravity model formulation in which trip distribution and modal split are carried out 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. The demand is determined given the number of lanes and the number of new lanes that result from the top level NDP, and given the travel costs that results from the lower level NDP with respect to route choice, as described in the previous section.

### 3. DESIGN METHOD

In the previous section, the robust road network design problem is formulated as a bi-level problem in which at 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. In this section a design method/solution algorithm is presented by which the formulated robust road network design problem can be solved.

It is generally known that the NDP 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 (e.g. 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. Figure 1 shows the design dilemma of optimizing the spare capacity of the network in the simplified 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.

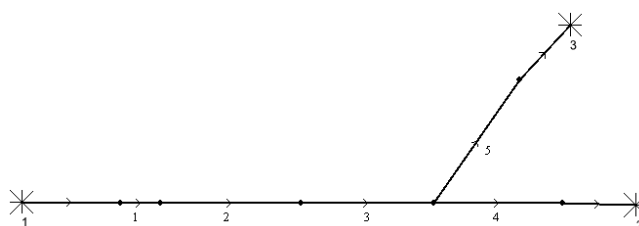


Figure 1. Example network

In order to solve this problem a design method is presented in Figure 2. The method 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 were involved in the design process. The fact that experts get the opportunity to interfere in the optimization process can be considered as an advantage since this makes it possible to enrich the model input and outcomes with knowledge of stakeholders that cannot be included in a model. Furthermore, as it is very difficult to include many design variables in a full model approach, this makes it possible to include all kinds of robustness measures in the optimization process. Finally, it creates support for the network design during the design phase and might thus 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 models show how vulnerable the existing network and new designs are and which locations of the network are vulnerable. The optimization models give suggestions about how much spare capacity is to be added on which locations. Finally, this gives quantitative support to the decisions that are made.

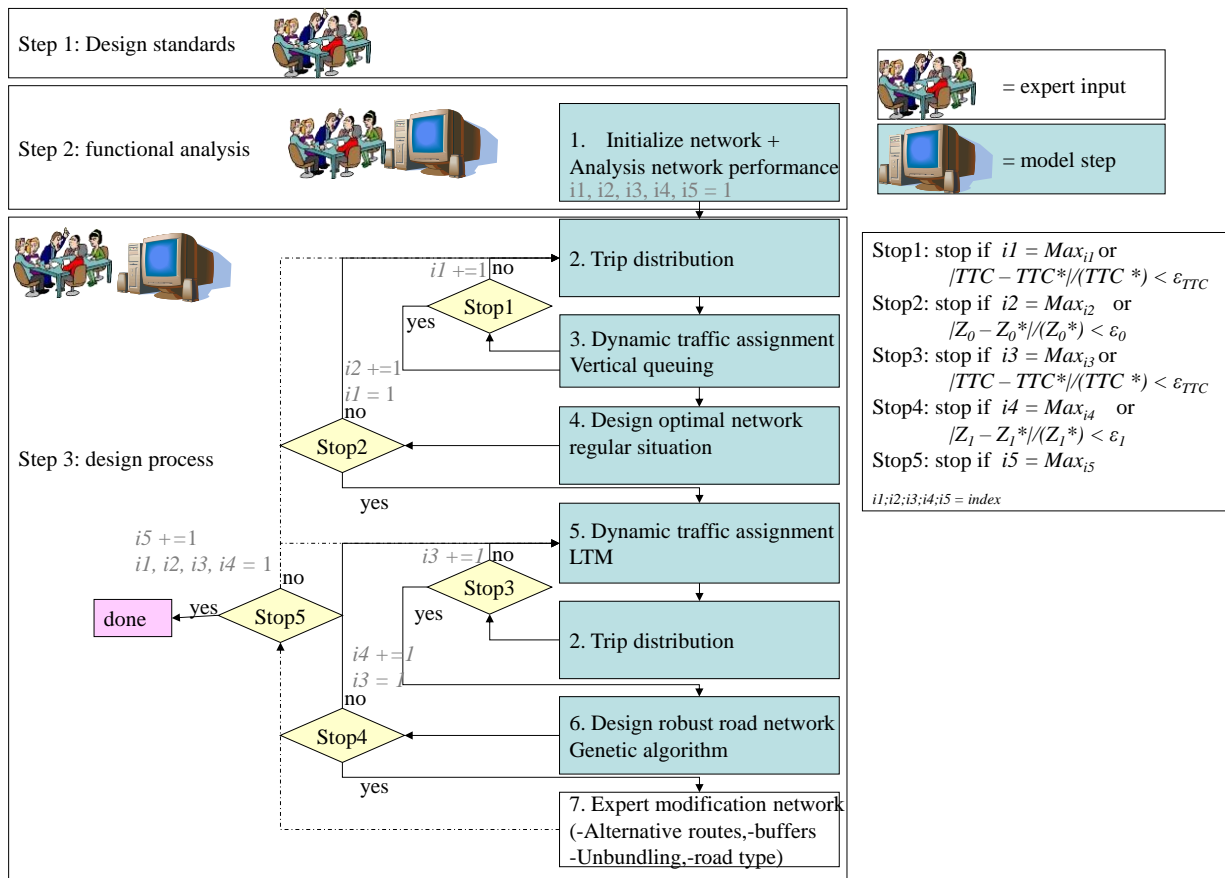


Figure 2. Design method

The solution algorithm consists of three steps:

1. In the first step, the design standards are set. This implies that the objective of designing/improving the network, the design variables, and the restrictions are specified as well as some parameter settings of the models. This can be done in half a day by experts in a workshop setting.
2. In the second step a functional analysis of the reference network is carried out. 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. The modelling part of this step is carried out with the macroscopic dynamic traffic assignment model Indy (Bliemer, 2005; Bliemer, 2007; Yperman, 2007) in combination with a marginal incident computation model (Corthout, 2009). 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 (link transmission model). Modelling spillback is important for robustness analyses. Given the equilibrium and traffic flows under regular circumstances, the marginal incident computation model is used to get an estimate of the impact of incidents very quickly. Four incident types are 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 evaluation method is described in detail in (Snelder et al., 2012).



3. In the third step the robust road network is designed in two sub-steps. At first the network is optimized for regular conditions by using a simple heuristic. Thereafter, the network is optimized for irregular conditions (incidents) by combining a genetic algorithm which optimizes the spare capacity with an expert adjustment of the resulting network.

The ‘optimal’ number of extra lanes for the regular situation is determined by a simple optimization process (model step 2, 3 and 4 in figure 2) in which a lane is added for each link until the objective function no longer improves. This simple procedure works because the objective function per link is unimodal. During the optimization the demand and the link flows are assumed to be fixed. The travel times are recomputed for each network configuration by using the adapted Smulders function (Smulders, 1988). After the number of extra lanes for each link is optimized, the trip distribution and dynamic traffic assignment model are run. This process continues until a maximum number of iterations is reached or the objective function does not change more than a certain  $\epsilon$ . It is not proven that this process converges. Furthermore, this process is not guaranteed to find the global optimum. However, in the section 4 it is shown that for an example network, the algorithm converges and the global optimum is found. During the optimization for the regular situation the dynamic traffic assignment model *Indy* is run with a simplified network loading model. Vertical queues are chosen (in which travel times are calculate with the same Smulders function), because the locations 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 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.

Based on the results of the optimization of the network under regular conditions, the ‘optimal’ number of extra lanes of spare capacity for improving the robustness of the network is indicated (model step 5, 2 and 6 in figure 2). 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. 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. A detailed explanation of the working of genetic algorithms can be found in Goldberg (1989) and Deb (2002). In our case, an initial population of chromosome is generated. A chromosome consists of a string of integer values that contains the number of additional lanes (design variable) for each existing link. The fitness of the chromosomes is determined by using the marginal incident computation model (Corthout et al., 2009), in combination with the alternative route approximation (see evaluation method step 2). 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. 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 by means of genetic operators (denoted as reproduction and mutation). The genetic algorithm stops when it reaches a predefined number of generations. After the number of extra lanes for each link is optimized, the trip distribution and dynamic traffic assignment model (with the link transmission model) are run. This process continues until a maximum number of iterations is reached or the objective function does not change more than a certain  $\epsilon$ .

Given the ‘optimal’ spare capacity, the experts can indicate how this capacity should be spread over alternative routes. Furthermore, they can add all kinds of other measures to improve the robustness of the road network. 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 (step 2). 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 figure 2, 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.

The way in which the design method works is explained in more detail in section 4 on the basis of an example for a small test network and a large case study for Amsterdam.

## 4. CASES

In this section the working of the design method is demonstrated for two cases. The first is a test network (section 4.1) and the second is the network of Amsterdam and surroundings in the Netherlands (section 4.2)

### 4.1. Test Network

In this section, the functioning of the algorithm is shown for a small test network as shown in Figure 1.

#### *Step 1: Design Standards*

We chose to optimize the network according to the objective function that is specified in section 3, without computing the net present value of the costs (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 do 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 1), 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 1.

Table 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 and from home to a place other than work. The number of trips produced for each purposes is respectively computed by 0,09 and 0,10 times the number of residents. The number of trips attracted for each purpose is respectively computed by 0,20 and 0,30 times the number of 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 the dynamic traffic assignment model. 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 2.

Table 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, 0.9, 0.7, 0.5, 0.4, 0.3, and 0.1. Table 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 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]
1	1.84	4	9664	120	5298.8	1.70
2	6.36	4	9664	120	5298.8	1.70
3	6.01	3	7248	120	5298.8	1.70
4	5.80	2	4832	120	5298.8	1.70
5	7.34	2	4832	120	5298.8	1.70

In this initial equilibrium state, congestion occurs on link 2, because the capacity of link 3 is smaller than the capacity of link 2. The simulation of four incident types on each link shows that the expected costs of vehicle loss hours caused by incidents is the highest on link 3 followed by link 5, 1, 2 and 4.

### 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 1 and 2 get two extra lanes, link 3 gets three extra lanes, and links 4 and 5 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%.

Figure 3 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 is shown in a string. For instance, [00050] means that links 1, 2, 4, 3, and 5 (in this order) get 0, 0, 0, 5, and 0 extra lanes, respectively. In the first iteration, 5 lanes are added to link 3. 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 1, 2, and 3. 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 3. Again, this appeared too much, so the number of extra lanes is reduced to 2. Also, links 1 and 2 get an extra lane as a result of the increased demand. Again this creates a bottleneck, so link 3 gets an extra lane. After this step, convergence is reached and the calculations end.

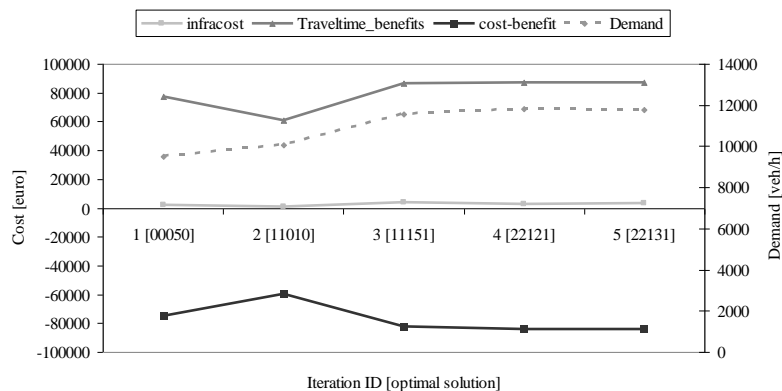


Figure 3. Optimization under regular circumstances

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 the dynamic traffic assignment model 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.

The optimal spare capacity, expressed by the optimal number of extra lanes, is shown in Figure 4. 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.

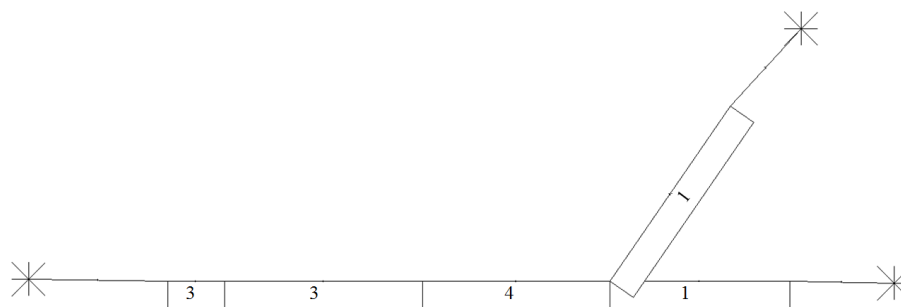


Figure 4. Number of extra lanes

In Figure 5, 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.

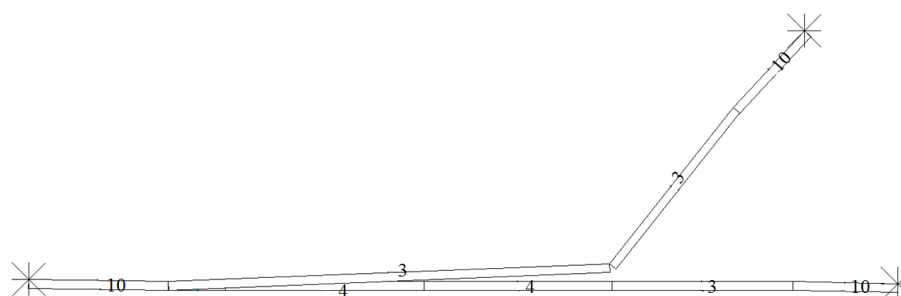


Figure 5. Number of lanes in variant structure

In table 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 4. Network indicators

	Initial network	After regular 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

### *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 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. 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.

### *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. It appears 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 (default €1.70 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. 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. A similar sensitivity analysis can be done by changing the value of time. 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. 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. If fixed costs of 5,298.80 €/km are considered, the optimal solution changes to [32131] in the case of elastic demand, and to [00010] in the case of fixed demand. This implies that 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.

### *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 3 shows that the algorithm that is used for optimizing the capacity under regular situations converges after four iterations. 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 6.

In the first iterations 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 is 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.

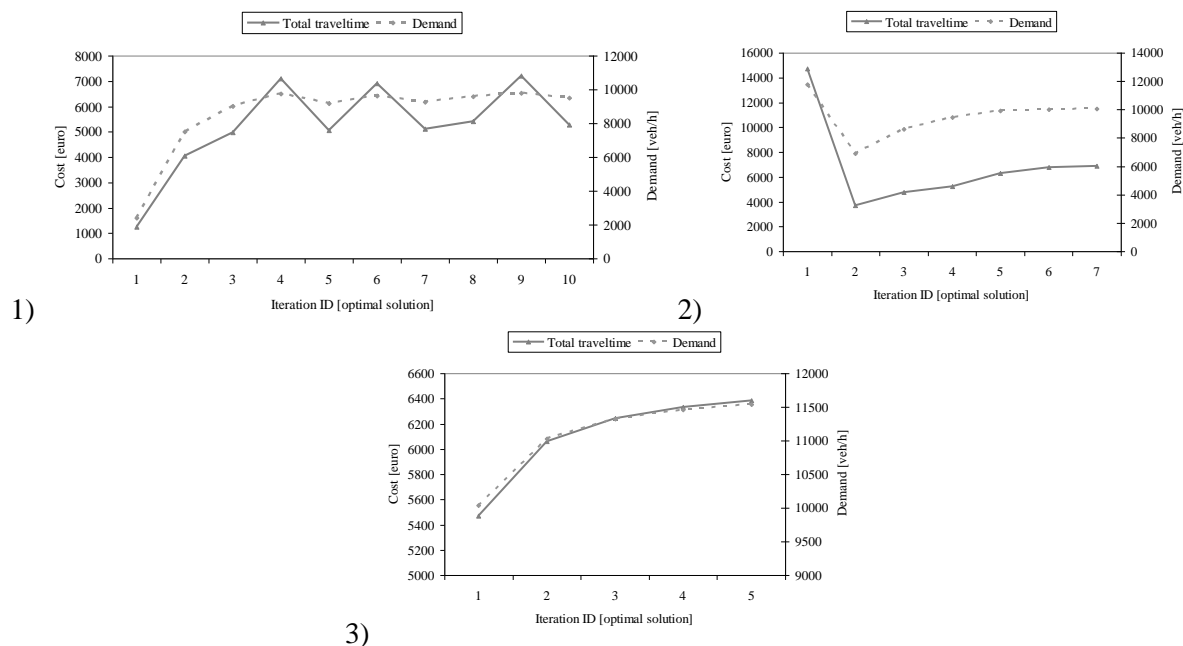


Figure 6. Convergences demand-assignment model in the first three iterations 1), 2) and 3) of the optimization model under regular conditions

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™ Core™2 Duo CPU T7500 @ 2.2 GHz 2.00 GB of RAM laptop.

#### 4.2. Amsterdam and Surroundings

In (Egeter and Snelder, 2010) a robust road network design is made for Amsterdam and surroundings in the Netherlands. However, models were not used. The design resulted from several workshops in which experts were involved. In this section 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 section 3 is applied in practice, an integrated approach should be followed. Therefore, the results presented in this section 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.

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 7 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. The initial network and matrices are derived from a static model and calibrated using a dynamic OD-estimator. 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.



Figure 7. Modelled network of Amsterdam (black: motorways)

#### *Step 1: Design Standards for Amsterdam and Surroundings*

We chose to optimize the network according to the objective function that is specified in section 3. 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.

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. 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. We did not include a budget constraint.

We chose to optimize the capacity under regular conditions and to optimize the spare capacity. 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:

- We calibrated the network based on an adjusted OD-matrix and network from a static model. 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 year 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. In practice, this percentage can be varied for the different incident types.

#### *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. The main conclusions of the functional analysis for the through traffic and external traffic are (the names of the roads can be found in figure 8):

- 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. The results of the functional analysis are an input for the design process.

*Step 3: Design process for Amsterdam and surroundings*

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.

Table 5 summarizes these results. 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.

Table 5. Summary network indicators for Amsterdam and surroundings

Indicator	Initial network	After regular 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

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 section, in this example no expert modification took place of the model optimization. However, based on Egeter and Snelder (2010), we do indicate what kinds of changes could be made. In figure 8 a robust road network design 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, 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 is 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. 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.

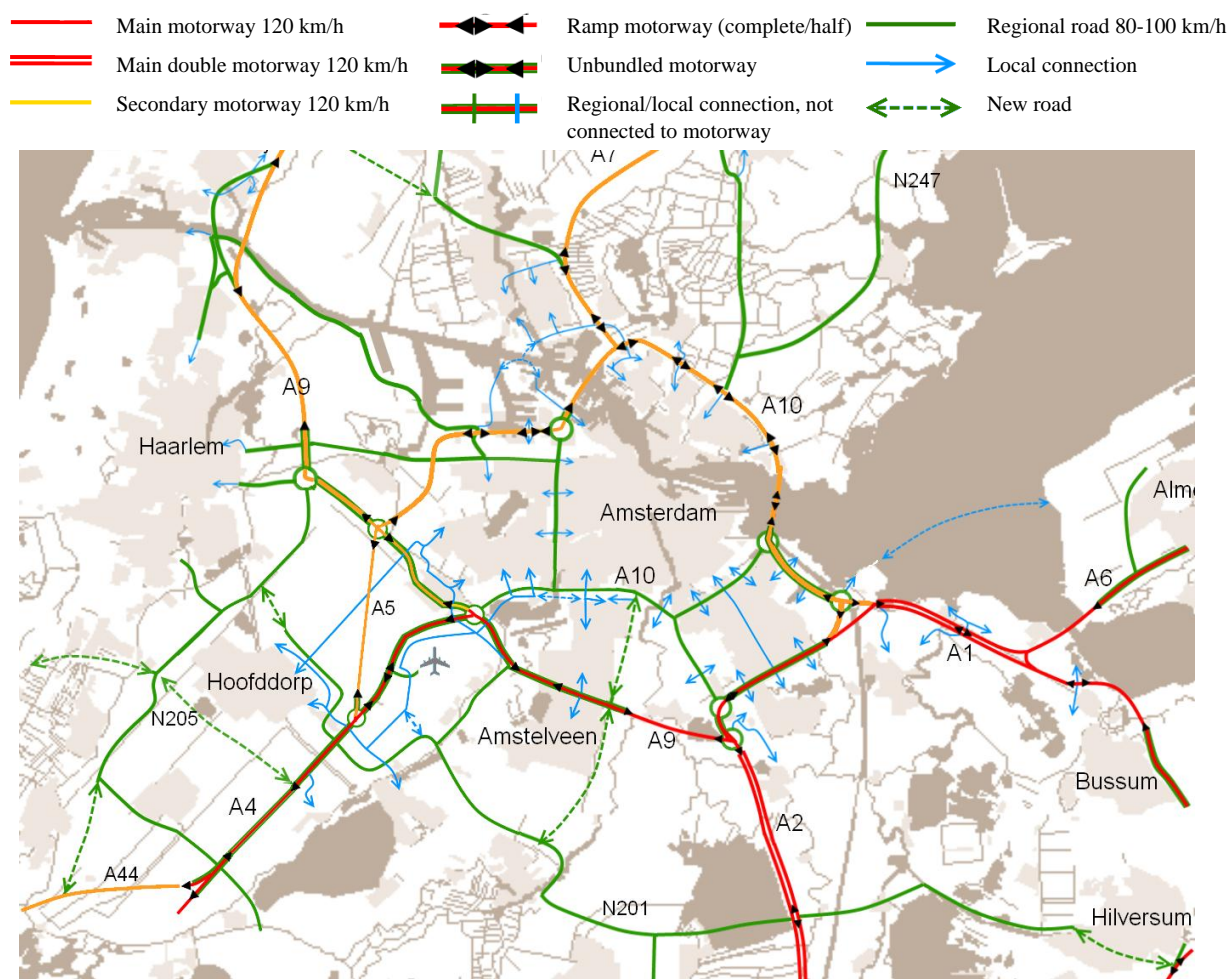


Figure 8. Function robust design after expert modification

It turned out the design method can be applied to large networks. This process takes several months, 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 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 a dynamic traffic assignment run with the link transmission model on the network of Amsterdam and surroundings takes about 12 hours on a normal PC. A dynamic traffic assignment 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 for the network of Amsterdam was 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 for instance by implementing the algorithm more efficiently, by using parallel computing techniques and by using better convergence criteria.

## 5. DISCUSSION AND RECOMMENDATIONS

In this paper a design method is proposed which combines expert knowledge and advanced modelling techniques. The fact that experts get the opportunity to interfere in the optimization process can be considered as an advantage since this makes it possible to enrich the model input and outcomes with knowledge of stakeholders that cannot be included in a model. Furthermore, as it is very difficult to include many design variables in a full model approach, this makes it possible to include all kinds of robustness measures in the optimization process. Finally, it creates support for the network design during the design phase and might thus 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 and new designs are and which locations of the network are vulnerable. 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 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 following 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 using a marginal incident computation model. A full dynamic traffic assignment with a link transmission model is only done 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. Nevertheless, we would recommend reducing the computation time of the models, because this makes application of the models easier and better solutions can be found in the same or less computation time. There are many possibilities for reducing the computation time for instance by implementing the algorithm more efficiently, by using parallel computing techniques and by using better convergence criteria.

The design method is also applied to a large realistic network of Amsterdam and surroundings. The application of the modelling steps of the design method shows that large improvements in the network performance can be realised with a positive benefit-cost balance. Finally, an impression is given on how expert judgements can further improve the robustness of the design by applying all kinds of robustness measures other than adding spare capacity to the vulnerable roads.

The design method can be improved and extended in many ways. It is for instances important to better understand the behavioural responses of drivers in case of disturbances and take them into account in the network design. Furthermore, it is important to consider other disturbances than incidents as well

and to consider other robustness measures besides structure related measures in the design process. Also, methods could be implemented that make it easier to learn from the intermediate results. 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.

## ACKNOWLEDGEMENTS

This research is supported by the research programmes Transumo and Next Generation Infrastructures, the Netherlands Organisation of Applied Scientific Research TNO, the Delft University of Technology and the Netherlands Research School for Transport, Infrastructure and Logistics TRAIL. Furthermore, we would like to thank the reviewers for their useful comments.

## REFERENCES

- Abdulaal, M., L.J. LeBlanc (1979) Continuous Equilibrium Network Design Models. *Transportation Research Part B*, 13, 1, 19-32.
- Boccaletti, S., V. Latorab, Y. Morenod, M. Chavezf, D.U. Hwanga (2006) Complex networks: Structure and dynamics. *Physics Reports* 424, 175 – 308.
- Bornholdt, S., H. G. Schuster (2003) *Handbook of Graphs and Networks, From the Genome to the Internet*. WILEY-VCH, Weinheim, Germany.
- 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. *Transportation Research Record: Journal of the Transportation Research Board*, 2029, 14-21.
- Boyce, D.E., B.N. Janson (1980) A discrete transportation network design problem with combined trip distribution and assignment. *Transportation Research Part B*, 14, 147-154.
- Chiou, S (2005) Bilevel programming for the continuous transport network design problem. *Transportation Research part B*, 39 361-383.
- Chootinan, P. S.C. Wong, A. Chen (2005) A Reliability-Based Network Design Problem. *Journal of Advanced Transportation*, 39, 3, 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. *Transportation Research Record: Journal of the Transportation Research Board*, 2099, 22-29.
- Davis, G. A. (1994) Exact local solution of the continuous network design problem via stochastic user equilibrium assignment. *Transportation Research part B*, 28, 1, 61–75.
- Deb, K. (2002) *Multi-Objective Optimization Using Evolutionary Algorithms*. Wiley, Chichester, United Kingdom.
- 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.
- 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. *European Journal of Operational Research*, 65, 44–57.

- Goldberg, D. E. (1989) *Genetic Algorithm in Search, Optimization and Machine Learning*. Addison-Wesley, Boston.
- Immers, B., J. Stada, I. Yperman, A. Bleukx (2004) Towards robust road network structures. *Slovak Journal of civil engineering*, XII (4), 10-17.
- Jamakovic, A., H. Wang, P. Van Mieghem (2006) Topological Characteristics of the Dutch Road Infrastructure. *Proceedings Seminar Infrastructure Reliability*, Delft, 2006.
- Klem E., M. Barten, M. Bijlsma (2004) Inspiratieboek Robuustheid, Verkenning van robuustheidvergrotenende maatregelen, Traffic Test bv, Veenendaal, TT04-060.
- LeBlanc, L.J. (1975) An Algorithm for the Discrete Network Design Problem. *Transportation Science*, 183-199.
- Li, H. (2009) Reliability-based Dynamic Network Design with Stochastic Networks, T2009/3, May 2009, TRAIL Research School, the Netherlands.
- Magnanti, T.L., R.T. Wong (1984) Network Design and Transportation Planning: Models and Algorithms. *Transportation Science*, Vol. 18, No. 1, 1984, pp. 1-48.
- 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. *Transportation Research Part B*, 25, 83-105.
- Nagurney, A., Q. Qiang (2009) *Fragile Networks, Identifying Vulnerabilities and Synergies in an Uncertain World*. John Wiley & Sons, Inc., Hoboken, New Jersey.
- 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. (2010a) Designing Robust Road Networks A general design method applied to the Netherlands, thesis, ISBN 978-90-79787-29-6.
- Steenbrink, P. A. (1974) *Optimization of Transportation Networks*, John Wiley & Sons, London.
- Snelder, M., B. Egeter, T. van Rooijen, L.H. Immers (2010b) P+R-plus, 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., H.J. van Zuylen, L.H. Immers (2012) A framework for robustness analysis of road networks for short term variations in supply. *Transportation research part A*, Volume 46, Issue 5, June 2012, 828–842.
- Sumalee, A., D.P. Watling, S. Nakayama (2006) Reliable Network Design Problem: the case with uncertain demand and total travel time reliability. *Proceedings of the Transportation Research Board of the National Academies*, Washington D.C., United States of America.
- Ukkusuri, S.V., T.V. Mathew, S.T. Waller (2007) Robust transportation network design under demand uncertainty. *Computer-Aided Civil and Infrastructure Engineering*, 22, 6–18.
- Yang, H., M.G.H. Bell (1998) Models and algorithms for road network design, a review and some developments. *Transport Reviews*, Vol. 18, No. 3, 257-278.
- Yperman, I. (2007) The Link Transmission Model for Dynamic Network Loading, Katholieke Universiteit Leuven, Faculteit Ingenieurs Wetenschappen, Afdeling Verkeer en Infrastructuur, Leuven.