

Forecasting Travel Time Reliability in Road Transport

A new Model for The Netherlands

02

Discussion Paper 2016 • 02

Marco Kouwenhoven

Significance, The Hague, Netherlands

Pim Warffemius

KiM Netherlands Institute for Transport Policy
Analysis, Ministry of Infrastructure and
the Environment, The Netherlands

**Forecasting Travel Time Reliability in Road Transport: a new
Model for The Netherlands**

Discussion Paper 2016-02

Prepared for the Roundtable:
Quantifying the Socio-Economic Benefits of Transport

9-10 November 2015,
International Energy Agency, Paris, France

Marco Kouwenhoven

Significance
The Hague
Netherlands

Pim Warffemius

KiM Netherlands Institute for Transport Policy Analysis
Ministry of Infrastructure and the Environment
The Netherlands

February 2016

The International Transport Forum

The International Transport Forum is an intergovernmental organisation with 57 member countries. It acts as a think tank for transport policy and organises the Annual Summit of transport ministers. ITF is the only global body that covers all transport modes. The ITF is politically autonomous and administratively integrated with the OECD.

The ITF works for transport policies that improve peoples' lives. Our mission is to foster a deeper understanding of the role of transport in economic growth, environmental sustainability and social inclusion and to raise the public profile of transport policy.

The ITF organises global dialogue for better transport. We act as a platform for discussion and pre-negotiation of policy issues across all transport modes. We analyse trends, share knowledge and promote exchange among transport decision-makers and civil society. The ITF's Annual Summit is the world's largest gathering of transport ministers and the leading global platform for dialogue on transport policy.

The Members of the Forum are: Albania, Armenia, Argentina, Australia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Canada, Chile, China (People's Republic of), Croatia, Czech Republic, Denmark, Estonia, Finland, France, Former Yugoslav Republic of Macedonia, Georgia, Germany, Greece, Hungary, Iceland, India, Ireland, Israel, Italy, Japan, Korea, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Mexico, Republic of Moldova, Montenegro, Morocco, the Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Russian Federation, Serbia, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, the United Kingdom and the United States.

International Transport Forum
2 rue André Pascal
F-75775 Paris Cedex 16
contact@itf-oecd.org
www.itf-oecd.org

ITF Discussion Papers

The ITF Discussion Paper series makes economic research, commissioned or carried out at its Research Centre, available to researchers and practitioners. They describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the ITF works. Any findings, interpretations and conclusions expressed herein are those of the authors and do not necessarily reflect the views of the International Transport Forum or the OECD. Neither the OECD, ITF nor the authors guarantee the accuracy of any data or other information contained in this publication and accept no responsibility whatsoever for any consequence of their use. This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Comments on Discussion Papers are welcome.

Table of Contents

Introduction	5
Methodology	7
Data	10
Determining raw travel times	11
Excluding extreme events.....	11
Correcting for variations in travel-time expectation.....	12
Calculation of the reliability indicator.....	12
Testing alternative empirical relations for travel time reliability	13
Linear model (The Netherlands – Hellinga).....	14
Length-standardised linear model (US – SHRP2).....	14
Length-standardised cubic model (United Kingdom – Mott MacDonald).....	15
Power-law relation between coefficient of variation and congestion index (UK–ARUP / WebTAG)	16
Exponential function between coefficient of variation and congestion index (Sweden – Eliasson) ..	17
Power-law relation between standard deviation and mean delay (Germany – Geistefeldt et al.)	18
Polynomial of mean delay and length (The Netherlands – Peer et al.).....	18
A new empirical relation for the Netherlands	19
Best functional form	19
Dependence on trip length.....	20
Differences between time-of-day periods.....	21
Impact of excluding outliers	22
Impact of correcting for travel-time expectation	22
Results for other roads.....	22
Policy implications	24
Current treatment of reliability in CBA.....	24
Better capturing the effects of policies that affect travel time variability.....	24

Impact on CBA results.....	25
Conclusions and future steps.....	26
Future steps.....	26
Acknowledgements.....	28
References	29
Notes	31

In this paper we describe how we included travel time variability in the national Dutch transport forecasting model and what the policy impacts of this new forecasting tool are. Until now, travel time reliability improvements for road projects were included in Dutch cost-benefit analysis (CBA) by multiplying the travel time benefits from reduced congestion by a factor 1.25. This proportionality is based on the linkage between congestion reduction and reliability improvements. However, this treatment of reliability is not useful to evaluate policies that especially affect travel time variability. From the start, this method was provisional and meant to be replaced by a better method capturing travel time variability. For this, we derived an empirical relation between the standard deviation of travel time, mean delay of travel time and length of route. This has been implemented in the national Dutch model as a post processing module. The new travel time reliability forecasting model will be incorporated in the Dutch guidelines for CBA.

Introduction

Absence of travel time reliability is what the travellers notice: “that frustrating characteristic of the transportation system that prompts motorists to allow an hour to make a trip that normally takes 30 minutes because the actual trip time is so unpredictable” (TRB, 2000, p. 4-1). The OECD (2010, p.31) defines reliability as “the ability of the transport system to provide the expected level of service quality upon which users have organised their activity”. The key aspect of this definition is the assumption that network users have an expectation of a particular level of service and that reliability is a measure of the extent to which the traveller’s experience matches their expectation (Hellings, 2011). In other words, reliability is equivalent to the predictability of travel times and, from the perspective of a traveller, associated with the statistical concept of variability. If one takes the perspective of a system or infrastructure network manager, then reliability measures are focused on the network and its performance, i.e. the fraction of time during which the system performs below a certain quality standard. In this paper we take the perspective of the road user (passenger as well as freight transport) whereby reliability is focused on trip characteristics.

Travellers and firms may account for the variability in their trips and transport of goods by building in time-buffers as insurance against late arrival. This implies that the consequences of late arrivals can be costly. Not only the efficiency and productivity lost in these buffers represent a cost that travellers and firms absorb due to unreliability, but also stress, late arrivals, missed connections, missed appointments and early arrivals can be costly. Reliable travel times are intrinsically valuable and network users place a significant value on reliability. Therefore, reliability can be formulated in terms of societal costs. This has the advantage that the investment costs of an infrastructure project to improve reliability can be traded-off against its benefits for society.

In the Netherlands, transport infrastructure projects and other transport policies are *ex-ante* evaluated using cost-benefit analysis (CBA). To incorporate reliability improvements in project and policy evaluation, the Dutch Ministry of Infrastructure and the Environment included the societal benefits of reliable and predictable travel times into CBA. Since 2004, in Dutch CBA practice, an

extra benefit of 25% of the travel time benefits due to reduced road congestion is added to account for reliability benefits (Besseling et al., 2004). This approach is only used for road projects and is based on the linkage between congestion reduction and improved reliability. However, it does not evaluate consequences of policies that especially affect travel time variability.

From the start, this method was provisional and meant to be replaced by a better method capturing travel time variability. To include travel time reliability in CBA, three types of information are needed, namely (De Jong and Bliemer, 2015):

- Monetary values to convert reliability benefits into money units.
- A model to predict how much an infrastructure improvement project will change travel time variability.
- A model to predict whether network users will change their route choice, mode choice or departure time choice due to changes in travel time variability.

The monetary value of changes in average travel time has been long included in CBA by making use of the so called value of travel time savings (VTTS). The VTTS refers to the monetary value travellers place on reducing their average travel time by one hour. In contrast, the value of travel time reliability savings (VTTRS) to convert changes in travel time variability in monetary units is relatively new. A recent study for the Dutch Ministry of Infrastructure and the Environment, delivered updated VTTS and VTTRS based on primary data (KiM, 2013; De Jong et. al, 2014; Kouwenhoven et. al, 2014). Based on earlier work (Hamer et. al, 2005; HEATCO, 2006) it was decided that the variability of travel time should be measured by standard deviation of the travel time distribution. The main reason behind this choice was the assessment that including travel time variability in transport forecasting models would be quite difficult and that using the standard deviation would be the easiest option. Any formulation that would go beyond the standard deviation or the variance of travel time¹ would be asking too much of the Dutch national and regional transport models (LMS and NRM) that are regularly used in CBA in the Netherlands. The standard deviation is also used as a reliability indicator in the US, UK and Scandinavian countries. The disadvantage of this definition is that it does not capture skewness of the travel time distribution. It is well known that travel-time distributions are skewed and their long tails towards extreme travel times are an important aspect of travel-time reliability (see also Hellinga, 2011). Studies on network robustness and vulnerability focus on these extreme travel times and their causes.

The inclusion of travel time variability in transport forecasting models is challenging because transportation planning models that are used to evaluate and prioritise transport policies have been developed to capture average travel time and not travel time variability. To adapt the Dutch national and regional models to capture reliability in terms of standard deviation, a project was started in 2013. The objective was to find a (new) empirical relation between the standard deviation of car travel time and other variables available in LMS and NRM. This paper reports the main results of this study. The improved modelling to forecast travel time variability will be implemented in Dutch policy making. Incorporating the consequences of policies affecting travel time variability into infrastructure CBA encourages proper consideration of options.

The next section of this paper discusses the methodology. The database which was used to derive the empirical relation is then described in the following section. Subsequent sections show how the best empirical relation was fitted. The policy impacts of the new reliability forecasting tool are then discussed before the paper ends with our conclusions and future steps.

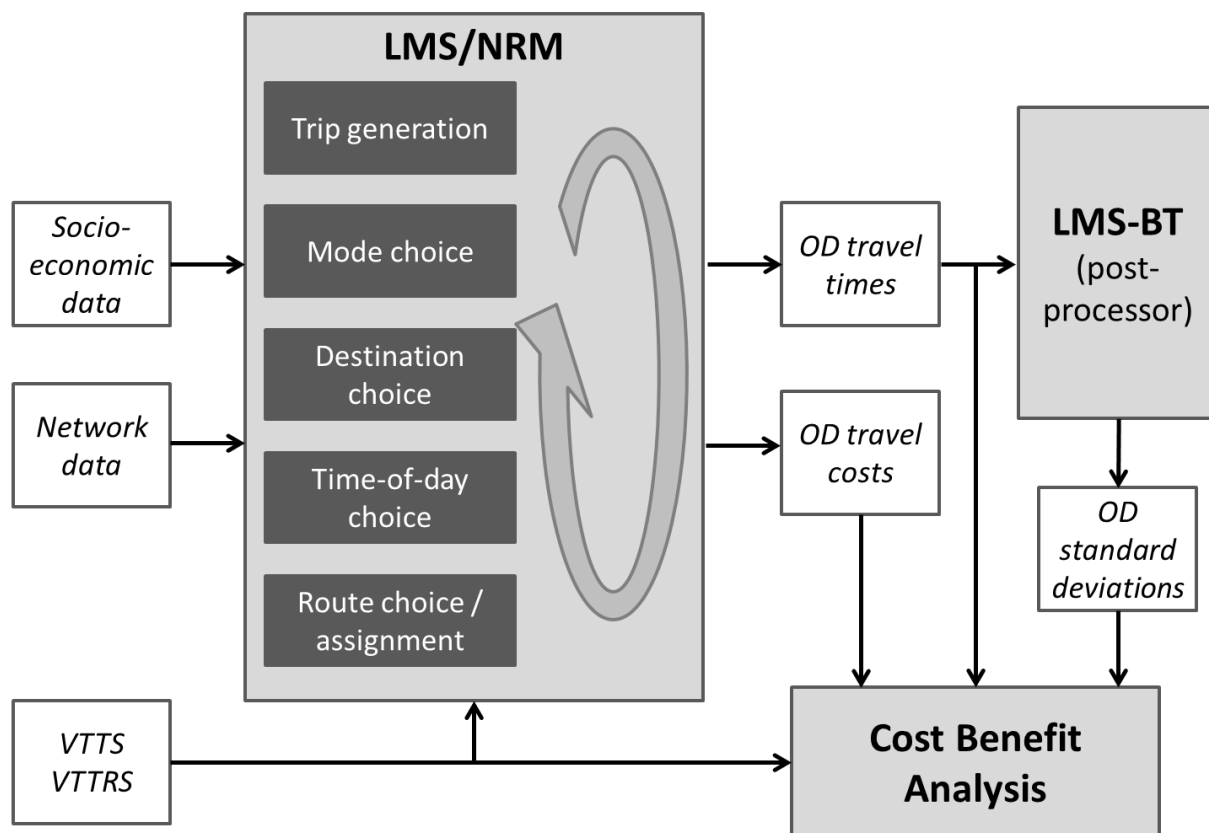
Methodology

In the Netherlands, traffic forecasts for CBAs are usually made using the LMS (the national transport model) or one of the NRMs (the regional transport models). These are similar tour-based models containing the four steps of a classical transport model (tour generation, destination choice, mode choice and route choice) plus a departure time choice module (Willigers & de Bok 2009; Significance 2011). The tour generation, destination, mode and departure time choices in the LMS and NRM are based on aggregate models, i.e. these choice models are estimated at the level of an individual traveller. Route choice (i.e. the assignment) for road transport is modelled at the level of origin-destination flows. In an iterative process, the resulting travel times are fed back into the earlier modules so that the travellers can adapt their choices as a result of possible congestion.

The assignment module generates the mean travel time on an average working day for each origin-destination pair and for each time-of-day period. In this process, flows are assigned to routes through the network and for each link in the route a travel time is calculated based on speed-flow curves and on a queuing model. The speed-flow curves are determined using empirical data for each road type. Note that the calculated mean travel time includes delays due to congestion.

Next to travel time and cost, reliability can be an important driver for mode choice, route choice and time-of-day choice (de Jong & Bliemer, 2015). Ideally, this variable would be included in these choice modules as an explanatory variable. However, this would require extensive data collection, modelling and adaptation of the current models. As a second best and much quicker solution, we have developed a post-processing module that calculates the reliability of road travel times for each origin-destination pair and for each time-of-day period. This allows us to calculate travel time reliability levels for any future scenario and to include the costs or benefits of possible changes in reliability in the CBA (Figure 1). Each policy measure that can be simulated with the LMS/NRM (adding road capacity, road pricing, etc.) can also be studied for its effects on travel time reliability.

Figure 1. The role of the transport models LMS / NRM and the post-processor LMS-BT in CBA



This post-processing module requires an empirical relation between travel time reliability and any of the output variables that are available in the LMS and NRM such as travel time, congestion, flow etc. but also the road characteristics that are available in the model (maximum allowed speed, number of lanes etc.). To derive such a relation, we need to have a database with observed levels of reliability and of all the other variables. This database is described in the next section.

When compiling such a database a number of decisions need to be made. These decisions can have a profound impact on the results. To prevent any inconsistencies, it is paramount that consistency with both the LMS/NRM and CBA procedure drives these decisions:

- As explained in the introduction, we use “the standard deviation of the travel-time distribution” as our reliability indicator. However, we still need to define *which* travel-time distribution. It is common to compile a travel-time distribution by measuring the mean travel times on a number of days within a certain period (e.g. one month or one year) when departing at the same time (e.g. between 08:00 and 08:15). In this way, reliability is interpreted as day-to-day travel-time variability. In this paper we follow this approach. We have measured mean travel times over a period of one year (2012) for vehicles departing in the same 15-minute interval. This means that for each 15-minute period, we have a different travel-time distribution and hence a different value of the reliability indicator.
- Note that by taking the mean travel time when departing at the same time, we already exclude the vehicle-to-vehicle variation from the reliability indicator. To a certain extent,

this makes sense: some vehicle-to-vehicle variation is caused by driver characteristics: some prefer to drive more slowly than others, and these drivers may have other mean travel times, but still have the same travel-time reliability. However, some of the vehicle-to-vehicle variation is caused by infrastructure: departing a fraction of a second later may result in just having to stop for a red traffic light and having a delay of one minute. Ideally, this is the type of variation that is included in the reliability indicator. However, in our project it is excluded because of the method we use to measure reliability.

- The LMS/NRM forecasts mean travel times and vehicle volumes for an average hour during the morning peak (lasting from 07:00 until 09:00), during the evening peak (from 16:00 to 18:00) and for a typical hour during the rest of the day (defined as an average hour between the mid-day period, i.e. between 10:00 and 15:00). In order to find the average reliability indicator for these three periods, we average the reliability indicators over all the 15-minute periods during these periods (eight 15-minute periods for each peak and twenty-eight 15-minute periods during the mid-day period). The reliability indicator for each 15-minute period is weighted with the average volume during that period.
- Note that this is different from first averaging the travel time over the full morning peak (for example) for every day of the year and then producing the travel-time distribution and determining the standard deviation. Travel times will not be constant over a two-hour period and this variability must be included. In order to calculate the pure reliability, the day-to-day variability must be determined with a small departure time resolution (15 minutes or less, depending on the time period over which the travel time on a day can be considered more or less constant), then the reliability indicator should be determined for each (small) time interval. Only in the final step, is the reliability indicator averaged over a longer period (e.g. over the whole peak).
- In Dutch policy, reliability and predictability are considered from the viewpoint of the traveller, (Ministry of Infrastructure and the Environment, 2012). Therefore, reliability should be defined in terms of the deviation of the real travel time from the predicted travel time. As a consequence, we should not look at the mere day-to-day variability of travel times, but correct for the variation in *expected* travel time. As an example: suppose that the travel time on a certain route is always 70 minutes on Mondays, and always 65 minutes on other days, then the travel time is perfectly predictable, and hence, perfectly reliable – given that the travellers know this.
- Extremely long travel times have a severe impact on the computed standard deviation. The long travel times can be caused by malfunctions in the detectors, but they can also be real events (e.g. a breakdown of the traffic system because of a severe accident or due to extreme weather conditions. Should these be included in the travel-time distribution? To decide on this, it is crucial to understand how results will be used. It is important to be consistent with the method used to determine the mean travel times in the demand model: did they include these long travel times? Also, consistency is needed with the method used to determine the VTTRS.
- In the Dutch situation, the speed-flow curves that are the base for the travel time calculations in the national model do not include extreme events. Additionally, the value of travel time reliability savings was determined using a stated preference experiment in which a travel-time distribution was shown without any extreme events (Significance et al., 2007). Finally,

we believe that these extreme travel times have different causes than normal day-to-day variation. From a policy point-of-view it is better to treat them separately.² Hence, in this project we excluded these data points from the analysis. In the next two sections of this paper, we analyse the impact of this on the outcomes.

Data

We compiled two databases with travel times: one for highway trips and one for trips on other roads. Since the Dutch national model produces forecasts for an average working day, we selected data from all 251 days in 2012 that fell within this definition.

Most Dutch highways are equipped with detector loops that measure average vehicle speed and volume at one-minute intervals. Fifteen minute averages of these variables were available for this project. We defined 250 routes on the Dutch highway network. Each realistic and logical route started at (or near) a highway entrance and ended at a highway exit, meaning that each route can comprise multiple links. The routes covered the network as completely as possible and overlapped each other as little as possible. Characteristics can be found in Table 1.

Table 1. **Characteristics of selected routes**

	Highways			Other roads		
	(250 routes)			(40 route)		
	Average	Minimum - Maximum		Average	Minimum - Maximum	
Length (km)	41.5	1.9 -	224.8	5.3	1.7 -	13.8
Average speed (km/h)	93.6	28.3 -	116.7	48.6	13.8 -	97.8
Av. max. speed (km/h)	112.7	92.4 -	123.6	66.0	23.5 -	109.0
Number of links	34.1	1 -	200	<i>n.a</i>	<i>n.a.</i>	

Major urban and regional roads are equipped with video cameras. With the use of Automatic Number Plate Recognition (ANPR) techniques, passage times of individual vehicles are recorded. Combining information from multiple cameras average travel times over fifteen minute can be obtained. Since this project focussed on highway travel, only 40 urban and regional routes on non-highways roads were defined to be used in this project.

The full data set was compiled in four steps, as described below.

Determining raw travel times

The travel time data from video cameras for other roads are already for the full route. However, the highway detectors only provided local speeds and these needed to be converted to mean travel times between two detector loops.

Vehicle speeds at each detector were estimated by adding these mean travel times, the total travel time for each route when departing within a certain 15-minute interval was determined. We took into account the fact that a vehicle on a long route does not pass all detectors in the same 15-minute interval. Since it turned out to be technically complex to combine data from consecutive days, we only looked at the 92 fifteen-minute intervals between 0:00 and 23:00. The average volume over a route was calculated by averaging the volumes at the detector loops (weighted with the length between the loops).

In this way, a database with mean travel times and volumes for 250 highway routes, on each of the 251 working days, for a departure time in each of the 92 fifteen-minute intervals was compiled. These data were enriched with the free-flow travel times and with route characteristics such as the length and the capacity (i.e. for each route, the maximum volume over 251 working days and 92 15-minute periods).

Travel time delays will be correlated between adjacent links, so the standard deviation of the total route travel time cannot be derived directly from the standard deviations of link travel times. Therefore, only the total route travel time is stored and the standard deviation is calculated only in the final step.

Excluding extreme events

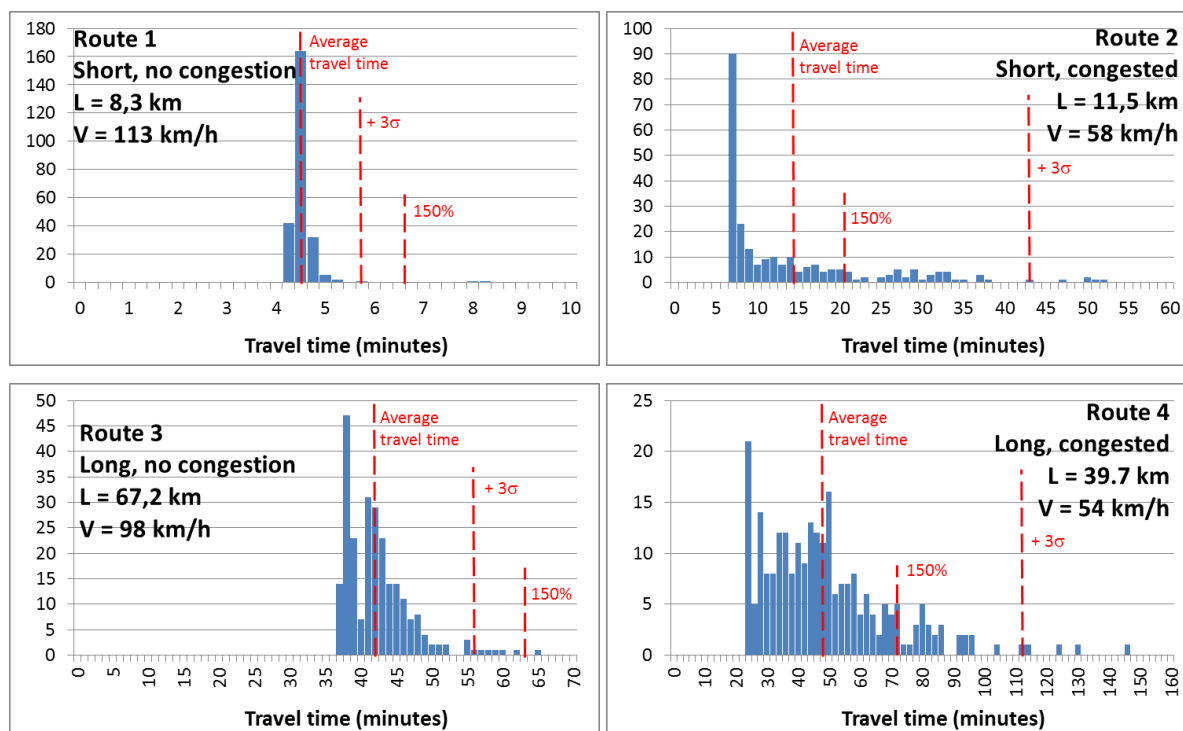
As explained in the previous section, we exclude extreme events from the travel-time distributions, in part because they have a strong impact on the standard deviation, but also for consistency in their application. However, it is not clear where to put the boundary for exclusion of an extreme travel time.

To determine this boundary, we visually inspected the raw travel-time distributions of all 250 highway routes (see Figure 2 for four examples). A boundary of three times the (raw) standard deviation above the mean travel time produced a good match with a visual classification of outliers for most routes. However, especially for routes with low congestion an additional criterion turned out to be necessary: the travel time of an outlier should be at least above 150% of the mean travel time. So, travel times are excluded with:

$$TT_{i,j,k} > \max\left(\overline{TT}_{j,k} + 3\sigma_{j,k}, 150\% \cdot \overline{TT}_{j,k}\right) \quad (1)$$

in which $TT_{i,k}$ is the travel time for day i , route j and 15-minute departure time period k , $\overline{TT}_{j,k}$ is the mean travel time for departure time period k , and $\sigma_{j,k}$ is the standard deviation of the travel-time distribution for time period k (before exclusion of extreme events). For each 15-minute period, equation (1) results in the exclusion of on average 4 out of the 251 working days. As a result, the average standard deviation over all 250 routes is reduced by 29 %. In other words, 1.6% of the days contribute to almost one third of the standard deviation.

Figure 2. Travel time distributions for four routes with different characteristics



Correcting for variations in travel-time expectation

As explained in the methodology section, we want to determine the deviation of the real travel time from the travel time as expected by the traveller. Unfortunately, no information is available about these expectations. Therefore, for each day, we approximate the expected travel time by taking the mean of the travel times on the same day-of-the-week in the four weeks before and after this day. For instance, our approximation of the expected travel time for Wednesday, 20th June is the mean of the travel times on the Wednesdays 23th and 30th of May, 6th, 13th, 27th of June, 4th, 11th and 18th of July. This running average reflects day-of-the-week and seasonal fluctuations in the mean travel time, but no incidental variations. No travel times that are marked as an outlier, are included in the calculation of the expected travel time. For holiday periods, when almost no congestion exists, the mean travel time of the four days directly before and after each day are taken.

As a result of this correction, the average standard deviation over all 250 routes is reduced by another 12 % on top of 29% reduction in the previous step.

Calculation of the reliability indicator

For each route, for each day of the year, for each 15-minute period, we calculate the deviation of the real travel time from the expected travel time. Next, for each route, for each 15-minute period, we determine the standard deviation of the deviations over all days (excluding the extreme events). For each period of the day (morning peak, evening peak, mid-day period), the mean travel time and the mean standard deviation is calculated by taking the average over the relevant 15-minute periods weighted by their mean flows.

Testing alternative empirical relations for travel time reliability

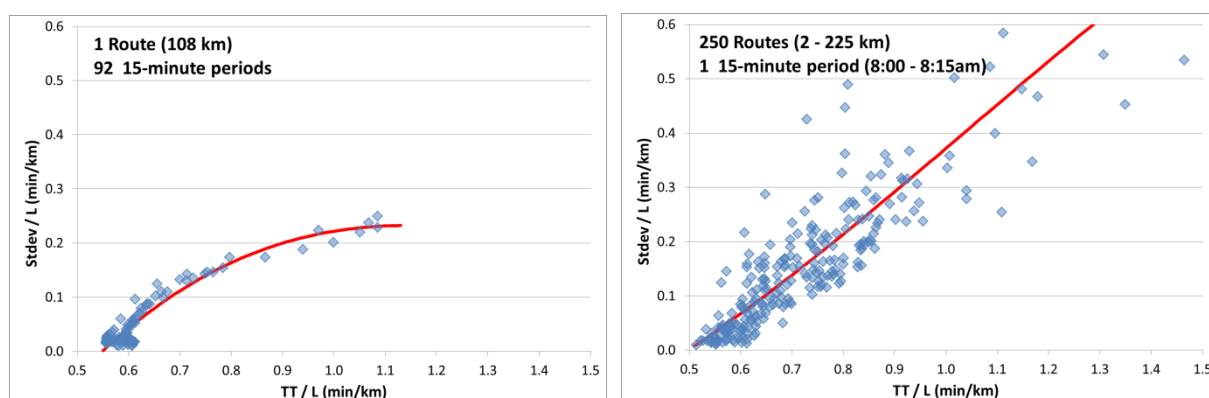
Our final databases contain 750 highway observations (250 routes times 3 time-of-day periods) and 120 observations non-highway observations (40 routes times 3 time-of-day periods). Each observation consists of an average travel time (averaged over all workdays in 2012), a standard deviation (of the distribution of deviations from the expected travel time), and other characteristics of the route (length, number of lanes, etc.). The non-highway routes are very diverse in nature (highway feeder routes in urban areas versus regional routes connecting two highways etc.) and are relatively short (see Table 1). This limited number of observations did not allow for extensive modelling efforts. However, the highway database was extensive and we were able to fit several functional forms described in the literature such that the results could be compared.

This analysis aims at finding the best functional forms of the empirical relations for our highway database. We do not try to compare the estimated coefficients with other studies as these values strongly depend on the outlier criterion used, on the way the variation of expected travel time is taken into account and on the period over which the data is averaged (also discussed in the next section). The conclusions on the best functional form are only valid for our own data set. It is well conceivable that other functional forms describe datasets in other countries better.

Many other researchers have used data sets that combined several routes and a range of 15-minute periods. The variation of standard deviation by routes may follow a different relation than the variation by 15-minute periods as can be seen from Figure 3. For the purpose of this project, we are only interested in the variation between routes for the morning peak, evening peak and mid-day period. In the evaluation of the results, we concentrate on the best functional form for the variation of the standard deviation by routes for the morning peak. The best function is one that describes the data best, and does not show any curvature that is not supported by the data (neither inside nor outside the data range).

In this analysis we tested functional forms that were suggested in earlier studies covering a wide range of possible functional forms that use different independent variables (travel time, travel time per kilometre, congestion index, mean delay). However, this list is not comprehensive. For a more complete overview of functional forms, see De Jong & Bliemer (2015).

Figure 3. **Travel time per km versus standard deviation per km for 92 15-minute periods**



Note: Values for a single route (left) and for 250 routes for the morning peak (right). Both datasets are fitted with cubic polynomials (red lines) which are clearly different.

Linear model (The Netherlands – Hellinga)

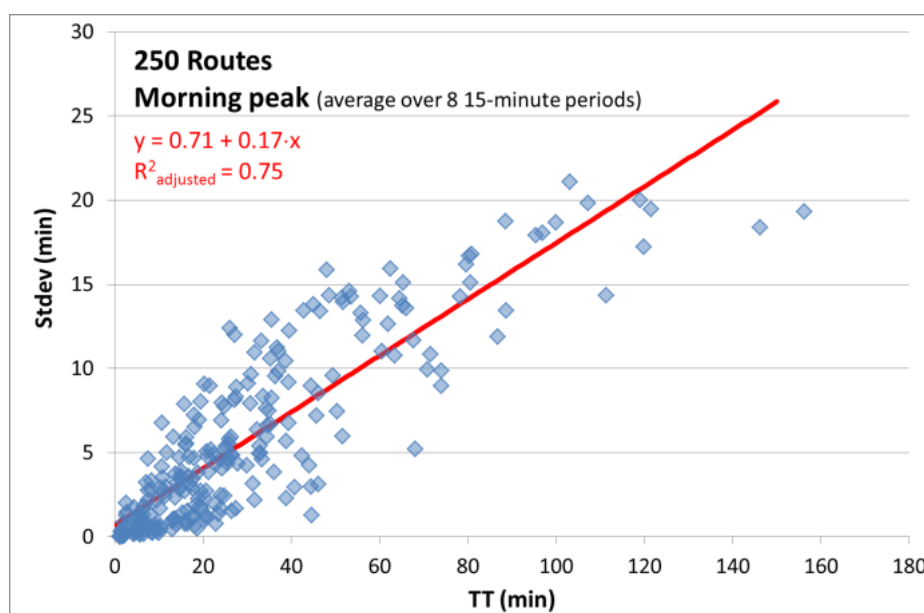
The first functional form under consideration is a simple linear relation between the standard deviation σ and the mean travel time TT:

$$\sigma = a[0] + a[1] \cdot TT \quad (2)$$

in which $a[0]$ and $a[1]$ are coefficients to be estimated. This functional form was used by Hellinga (2011) in his study of the variation of travel time on a single route of about 25 kilometres on the A12 highway in the Netherlands. Travel times were derived from detection loop data averaged over 15-minute periods. Each 15-minute period provided one data point, so his final data set consisted of 92 points, since he excluded trips after 23:00.

Hellinga only considered the variation between 15-minute periods and concluded that a linear relation was sufficient. If we analyse the variation between 15-minute periods for each of our 250 routes, we see that for short routes a linear relation is sufficient, though for longer routes with congestion, a decreasing slope can be observed (as is illustrated by the single route data in Figure 3-left). If we fit the linear relation (2) to our morning peak data for 250 routes (Figure 4) we see that our data can be nicely described by this relation, though a reasonable amount of spread around this relation remains. The adjusted R^2 is 0.75.³

Figure 4. **Travel time versus standard deviation for 250 routes for the morning peak fitted with a linear function**



Length-standardised linear model (US – SHRP2)

Several US researchers within the second Strategic Highway Research Program (SHRP2, see Mahmassani et al., 2014) use the standard deviation of standardized travel time. This network approach has the advantage that it can also be applied if multiple routes are used between A and B with different lengths. It is especially suitable for dense urban networks. The SHRP2 researchers have shown that there is an almost linear relation between the travel time per unit length and the standard deviation per unit length:

$$\frac{\sigma}{L} = a[0] + a[1] \cdot \frac{TT}{L} \quad (3)$$

in which L is the length of the route and $a[0]$ and $a[1]$ are coefficients to be estimated.

If the observations in our database are converted to this metric, a linear function fits very well (Figure 5, red line). However, some variation remains unexplained: the adjusted R^2 is 0.78, which is slightly better than for the linear relation above.

Length-standardised cubic model (United Kingdom – Mott MacDonald)

In the UK, Mott MacDonald estimated relations for the day-to-day variability which they describe as “what remains after accounting for all predictable variations (time of day effects, day type effects and seasonal effects) and variability due to incidents” (Sirivadidurage et al. , 2009). They used data from inductive loop sensors, automatic number plate recognition and matching and GPS tracking, averaged over 15-minute periods on several highway routes. Journey times that are more than 2 standard deviations above the mean are flagged as incidents and were excluded. Predictable variations were accounted for by allocating each day of the year to one of 21 day types and by determining average journey times and standard deviations for each of these day types.

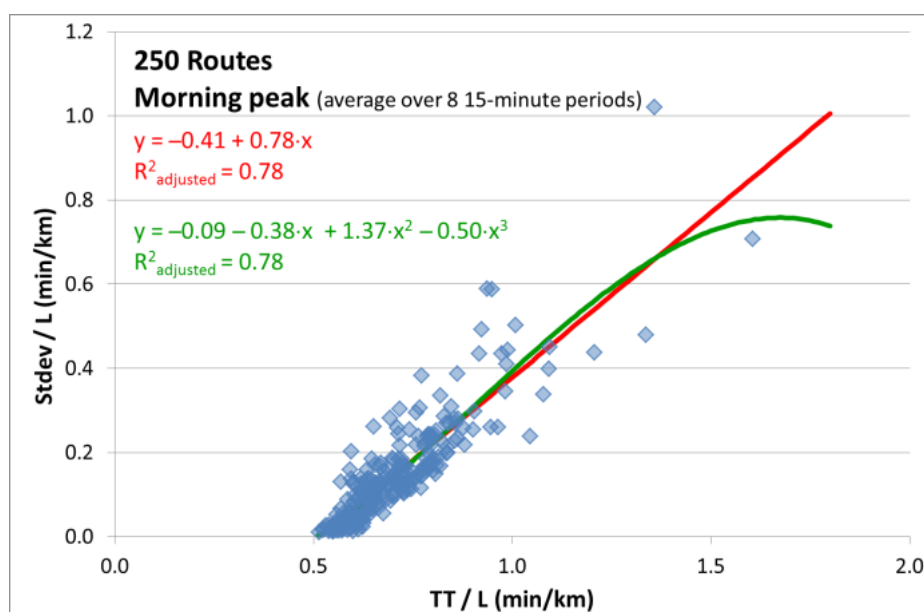
They present graphs with mean journey time per kilometre versus standard deviation per kilometre for several motorway types. The graphs for motorways with mandatory variable speed limits and with dynamic hard-shoulder running show indications for a relation that is slowly increasing for low congestion levels, increasing quickly for medium congestion levels and flattening for high congestion levels. They tried several functional forms and they obtained the best results when describing the standard deviation of travel time per kilometre as a cubic polynomial of the mean travel time per kilometre:

$$\frac{\sigma}{L} = a[0] + a[1] \cdot \frac{TT}{L} + a[2] \cdot \left(\frac{TT}{L}\right)^2 + a[3] \cdot \left(\frac{TT}{L}\right)^3 \quad (4)$$

They observed that mean *delay* per kilometre instead of mean *travel time* per kilometre gave a marginally better fit, however, the computation of free flow travel times had some difficulties.

The fit of function (4) on our morning peak data (Figure 5, green line) flattens above around 1.3 minutes/km which is not supported by the data. The adjusted R^2 is the same as for the fit of the linear function (3). We conclude that for the standard deviation per kilometre as a function of the travel time per kilometre a linear function is sufficient and that applying a cubic polygon does not improve the fit.

Figure 5. Travel time per km versus standard deviation per km for 250 routes for the morning peak fitted with a linear relation (red line) and a cubic polynomial (green line)



Power-law relation between coefficient of variation and congestion index (UK – ARUP / WebTAG)

Arup et al. (2003) analysed travel time variability on urban roads by estimating a model to travel times from a few probe vehicles in London and Leeds gathered over a period of about one month. They observed that variability is likely to be greater as flows reach capacity. Based on some theoretical considerations, they estimated a power-law relation between the coefficient of variation (CV, i.e. the ratio of standard deviation to the mean travel time), the congestion index (CI, i.e. ratio of the mean travel time to the free-flow travel time) and the length of the route L :

$$\frac{\sigma}{TT} = a[0] \cdot \left(\frac{TT}{TT_{ff}} \right)^{a[1]} \cdot L^{a[2]} \quad (5)$$

in which TT_{ff} is the travel time under free-flow conditions.

A consortium led by Hyder Consulting (Hyder Consulting et al. 2008a, 2008b; Gilliam et al. 2008) collected new data from GPS equipped vehicles on 34 routes (up to 12 km long) within the 10 largest urban areas in England for a period of three years. They estimated the same function on their data and found similar coefficients as Arup et al. Today, this functional form is recommended in the WebTAG guidance of the UK Ministry of Transport (2014).

Unfortunately, this functional form does not fit our data well (adjusted R^2 is only 0.57, which is much lower than for the functions above), as can be seen from Figure 6 (red line). This is probably due to the fact that our data is for highways and this functional form was derived for urban roads. Most notably, our data strongly supports a functional form that goes (closely) through the point $(CI, CV) = (1, 0)$ whereas this functional form does not. For highways routes it is understandable that in the

absence of any congestion, very little travel time variability is observed, whereas for urban roads variability will remain through differences in signalling or pedestrian crossings.

Exponential function between coefficient of variation and congestion index (Sweden – Eliasson)

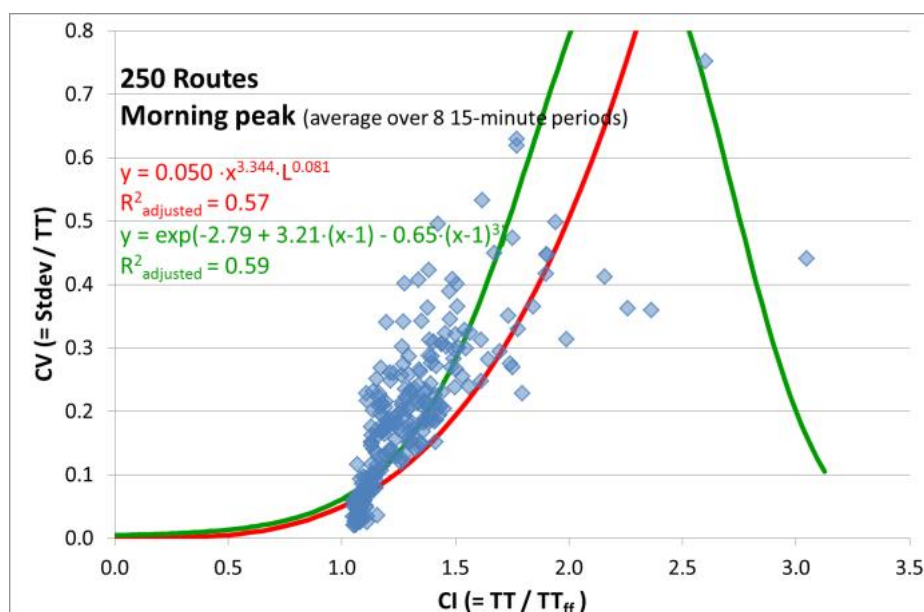
Eliasson (2006) fitted an exponential function to the coefficient of variation for 20 roads and for 96 15-minute periods in Stockholm, Sweden. Data was collected from automatic camera systems taking pictures of licence plates. These roads were characterised as “urban”, i.e. neither highways, nor small local streets. Lengths varied between 300 metres and 5 kilometres.

When inspecting the relation between the congestion index and the coefficient of variation for all 15-minute periods for each road, Eliasson noticed that the coefficient of variation remained roughly constant for low levels of congestion and increased for slightly higher levels. For high levels of congestion the coefficient of variation decreased again. Therefore, he used a cubic polynomial (excluding the second-order term) of the congestion index minus 1:

$$\frac{\sigma}{TT} = \exp \left(a[0] + a[1] \cdot \left(\frac{TT}{TT_{ff}} - 1 \right) + a[2] \cdot \left(\frac{TT}{TT_{ff}} - 1 \right)^3 \right) \quad (6)$$

Again, we tried to fit this functional form to our data, but this did not lead to a satisfactory result (Figure 6, green line), though the adjusted R^2 is slightly better than for the fit of function (5). In our data we observed neither a roughly constant coefficient of variation for low congestion levels, nor a decreasing coefficient of variation for high congestion levels. This different behaviour is likely due to the fact that our database consists of longer highway routes rather than short urban routes. Furthermore, we investigate the variation between routes whereas Eliasson also included the variation between 15-minute periods.

Figure 6. Congestion index (CI) versus coefficient of variation (CV) for 250 routes for the morning peak fitted with a power law (red lines) and with an exponential function (green lines)



Power-law relation between standard deviation and mean delay (Germany – Geistefeldt et al.)

Recently, the German Federal Ministry of Transport (BMVBI) funded a research project on the reliability of travel time on their highways. Geistefeldt et al. (2014) suggested using a power-law function between the standard deviation and the mean delay (i.e. the difference between the mean travel time and the free flow travel time):

$$\sigma = a[0] \cdot MD^{a[1]} \quad (7)$$

where MD is the mean delay. They estimated their coefficients on simulated data from a macroscopic traffic simulation model.

This functional form seems to describe our data well (Figure 7, red line). Note that the spread of the data points in Figure 7 is small compared to the spread when relating the standard deviation to the travel time (Figure 4) or when relating the travel time per kilometre to the standard deviation per kilometre (Figure 5). The adjusted R^2 of 0.82 is the better than for the fits previously discussed. So, using the mean delay as the explanatory variable seems to be a good idea.

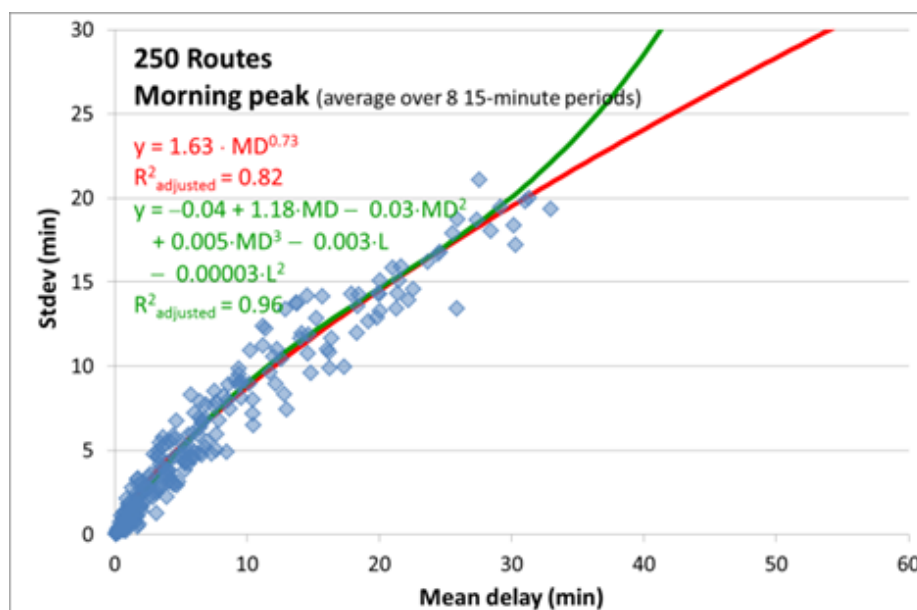
Polynomial of mean delay and length (The Netherlands – Peer et al.)

Peer (2012) estimated a relation between the standard deviation and the mean delay. In her PhD research she tried multiple functions on data from 145 highway routes and 57 15-minute periods. The best function contained (among other terms) a cubic polynomial in the mean delay and a quadratic polynomial in the length:

$$\sigma = a[0] + a[1] \cdot MD + a[2] \cdot MD^2 + a[3] \cdot MD^3 + a[4] \cdot L + a[5] \cdot L^2 + \text{other terms} \quad (8)$$

This function fits our data very well (note the adjusted R^2 of 0.96 in Figure 7, green line). However, the slope of the fitted function for the morning peak seems to steepen above a mean delay of about 30 minutes which is not supported by the data. So, the cubic polynomial may lead to unwanted behaviour outside the range on which it was fitted.

Figure 7. Travel time versus standard deviation for 250 routes for the morning peak (right) fitted with a power law (red line) and a cubic polynomial (green line)



A new empirical relation for the Netherlands

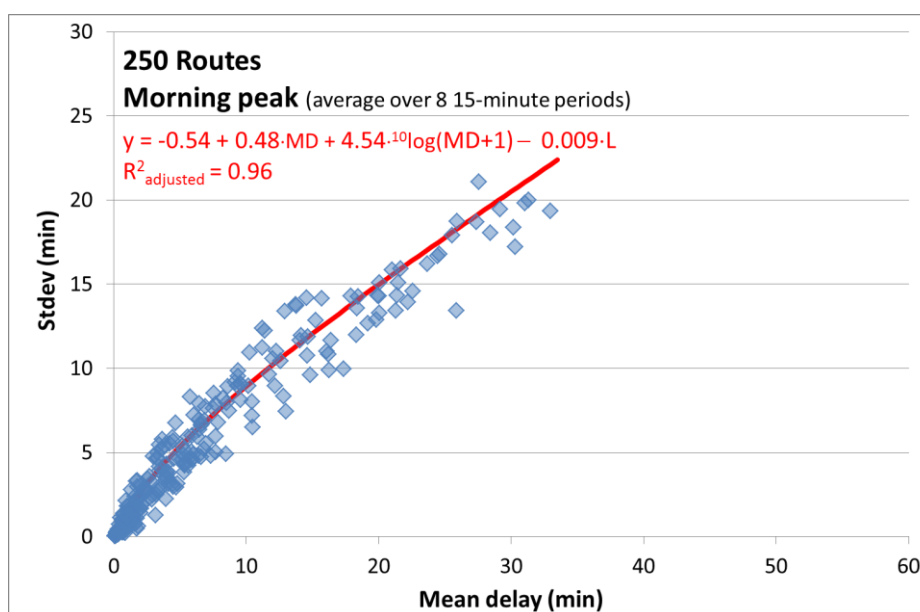
Best functional form

From Chapter 4 we conclude that the best results are obtained when relating the standard deviation to the mean delay. However, using a cubic polynomial may not be optimal. Therefore, we decided to test a combination of a linear and logarithmic function for the mean delay, and added a linear term in the length. Higher order terms and terms proportional with other parameters such as density, number of lanes, average weather conditions, and frequency of incidents, were not found significant.

$$\sigma = a[0] + a[1] \cdot MD + a[2] \cdot \log(MD + 1) + a[3] \cdot L \quad (9)$$

For the morning peak data, this fits the data very well (Figure 8) and does not lead to unwanted behaviour outside the range on which it was fitted. This function is selected as our final relation to forecast standard deviation based on the elements available from the traffic model.

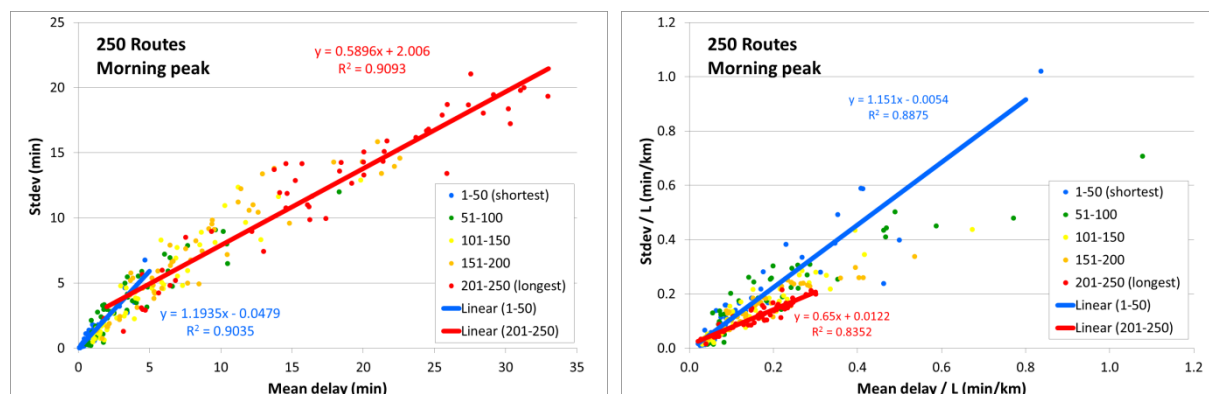
Figure 8. Mean delay versus standard deviation for the morning peak fitted with a combination of a linear and a logarithmic function



Dependence on trip length

When we look at the estimation result for our final function, we note that the coefficient on length is very small (only 0.009 as can be seen from Figure 8). One might conclude that the length is not important. However, length is also related to mean delay: the longer the route, the more likely it is that some congestion occurs. This becomes clear when we colour-code all data points based on their length. In Figure 9-left, the 50 shortest routes are displayed in blue, while the longest routes are shown in red. We see that all blue points are on the left of the diagram, while the red points are on the right. We have made a linear regression on both the red and blue points. We see that the slope decreases with length as well. This property is also clearly visible when we plot the standard deviation per kilometre as a function of mean delay per kilometre (Figure 9-right): the slope of their linear relation is correlated with length. This can be intuitively understood: if for a long route on a certain day, the congestion is worse than normal, traffic might flow better downstream, so any delay can be (somewhat) compensated later along the route. This will reduce the variation of the day-to-day travel time.

Figure 9. Variability and delay relations for 250 routes for the morning peak



Note: Mean delay versus standard deviation (left) and mean delay per unit length versus standard deviation per unit length (right). Blue dots indicate the 50 shortest routes (less than 12.6 km) and the blue line is the linear regression through these points. Red dots indicate the 50 longest routes (above 63 km) and the red line is the linear regression through these points.

Differences between time-of-day periods

We used the same functional form to analyse the evening peak (16:00 – 18:00) data and the mid-day data (10:00 – 15:00), see Table 2 for the estimates of the coefficients. Even though the three periods have significantly different coefficients (based on an F-test), the functional form fits each data set well.

For the mid-day period, we did not find $a[2]$ - and $a[3]$ - coefficients that were significantly different from zero. Therefore, we tried a fit with these coefficient constrained to zero, effectively turning equation (9) into a linear equation. This is understandable since the maximum mean delay for the mid-day period is only 10 minutes. Even in the morning peak the observations in Figure 8 below a mean delay of 10 minutes almost follow a straight line. Note that we kept the $a[0]$ constant though it is not significantly different from zero, since we did not want to force the function to go through the point with $(MD, \sigma) = (0, 0)$.

Table 2. Best fit coefficients for the empirical relation between the standard deviation and the mean delay (equation 9) for highway routes

	Morning peak			Mid-day period			Evening peak			
	Coefficient	(t-ratio)		Coefficient	(t-ratio)		Coefficient	(t-ratio)		
$a[0]$	-0.540	± 0.186	(-2.9)	-0.066	± 0.051	(-1.3)	-0.901	± 0.172	(-5.3)	<i>min.</i>
$a[1]$	0.476	± 0.026	(18.2)	1.034	± 0.019	(53.1)	0.268	± 0.017	(16.1)	
$a[2]$	4.538	± 0.415	(10.9)	-			5.555	± 0.351	(15.8)	<i>min.</i>
$a[3]$	-0.009	± 0.003	(-2.7)	-			0.011	± 0.003	(4.0)	<i>min. / km</i>
Adj. R^2	0.956			0.919			0.960			

Note: for the mid-day period no significant value for the $a[2]$ and $a[3]$ coefficients was found.

Impact of excluding outliers

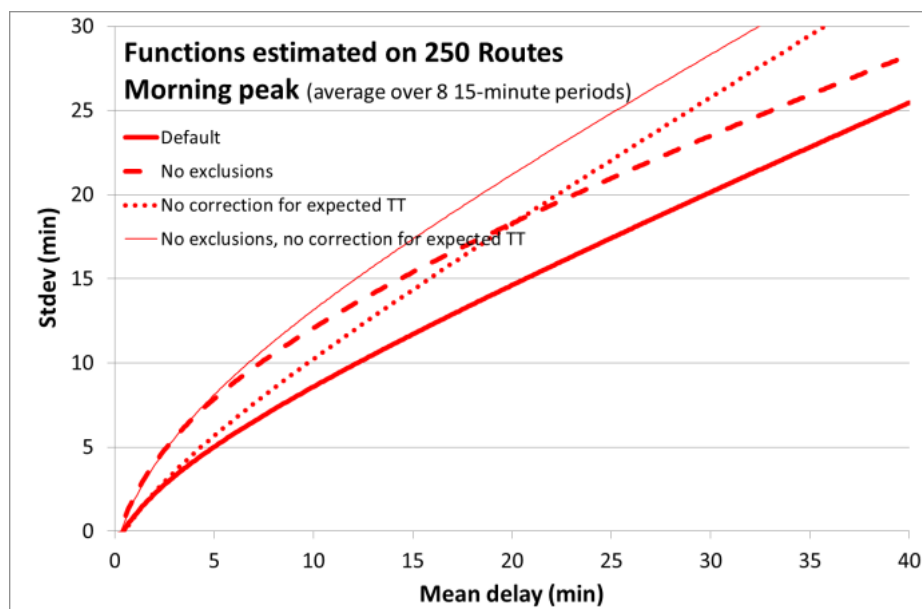
In Chapter 3, we noted that the exclusion of the outliers caused a decrease of the average standard deviation of 29%. However, excluding outliers also has an impact on the mean delay. So, it is theoretically possible that the data before and after exclusion fall on the same line: the exclusions may only cause a shift along the line. To test this, we fitted a function on the full data set (no exclusions), but with the correction for expected travel time as discussed. The resulting best fit is the dashed line in Figure 10, which lies roughly 3 minutes above the default (solid) line. From this, we conclude that outliers have more impact on the standard deviation than on the mean delay and excluding the outliers has a strong impact on the coefficients of the relationship.

Impact of correcting for travel-time expectation

The correction for the expected travel time (see the third step in the Data section) influences the standard deviation but not the mean delay. So, we expect that without this correction, the empirical relation between mean delay and standard deviation would shift upwards. This can indeed be seen from Figure 10. If no correction is made for expected travel time (dotted line), the curve is located up to 30% above the default line.

We conclude that both the outlier criterion and the travel-time expectation correction both have a clear impact on the coefficients (though our best functional form still describes the data well). As such, comparisons of coefficient values between different studies are not very useful unless these studies use exactly the same outlier criterion and travel-time expectation correction.

Figure 10. Results of fits for mean delay versus standard deviation for 250 routes for several choices of the data analysis



Results for other roads

We also fitted the same functional form to our database of 40 routes on other (non-highway) roads. Since these routes are small compared to the highway routes (see Table 1), we also have relatively small mean delays and standard deviations. As a result, only the linear term in equation (9)

was found to be significant. Figure 11 shows the data and the fit for the morning peak. Table 3 shows the best fit coefficients for all time-of-day periods. Note that insignificant constants were kept in the models. The coefficients for the evening peak are significantly different from those for the morning peak. The coefficients for the mid-day period are significantly different to those of the evening peak, but not from those of the morning peak. Also note that the slope for the morning peak (0.468) is much lower than the slope for the same period for short highway routes (1.1935, see Figure 9-right). So, the reliability relation for other roads is clearly different from that for highways.

Figure 11. Mean delay versus standard deviation for the morning peak for 40 non-highway routes fitted with a linear function

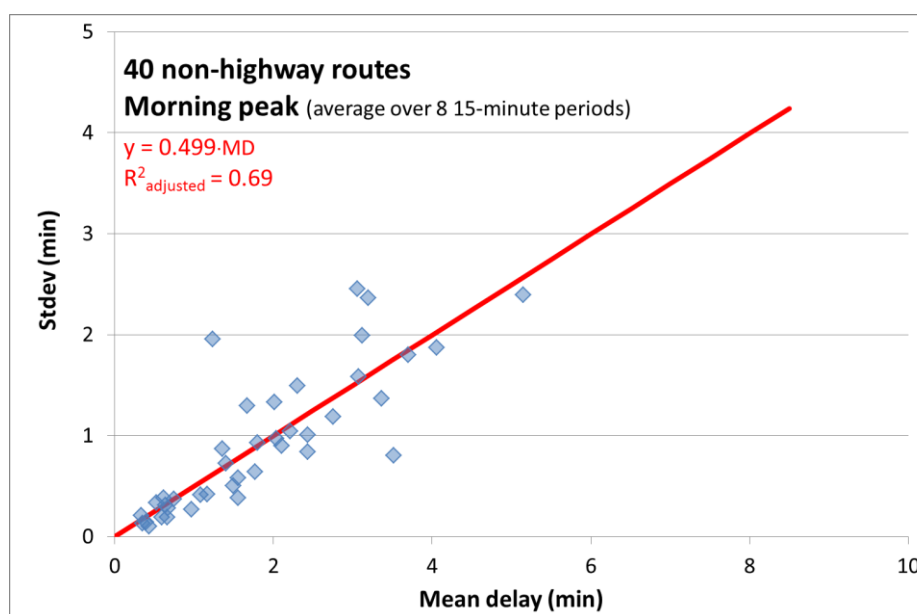


Table 3. Best fit coefficients for the empirical relation between the standard deviation and the mean delay (equation 9) for other routes

	Morning peak			Mid-day period			Evening peak			
	Coefficient	(t-ratio)		Coefficient	(t-ratio)		Coefficient	(t-ratio)		
a[0]	0.049	± 0.120	(0.4)	-0.074	± 0.049	(-1.5)	-0.079	± 0.106	(-0.7)	<i>min.</i>
a[1]	0.468	± 0.054	(8.7)	0.534	± 0.030	(17.6)	0.637	± 0.044	(14.6)	
a[2]	-			-			-			<i>min.</i>
a[3]	-			-			-			<i>min. / km</i>
Adj. R ²	0.662			0.891			0.848			

Note: “-” indicates that no significant value was found.

Policy implications

Current treatment of reliability in CBA

Until now, reliability is included in Dutch CBAs using the practical and provisional way as developed by Besseling et al (2004). That means that reliability benefits are included by multiplying the travel time benefits from reduced congestion by a factor of 1.25. This proportionality is based on the linkage between congestion reduction and reliability improvements.

However, this current treatment of reliability in CBA is not useful to evaluate policies that especially affect travel time variability. An approach to reflect the effects of policies that affect travel time variability in CBA will encourage proper considerations of options. Project appraisal will then not only offer incentives for policies that reduce the average travel time but also for policies that improve travel time variability.

Better capturing the effects of policies that affect travel time variability

The new travel time reliability forecasting model does not require any adjustments with respect to the transport model. It is a separate module that uses outputs from the transport model to forecast the impact of infrastructure projects on travel time variability. It is a so called post-processing module. Its outputs will not feed back into the transport model. That means that the reactions of network users to changes in reliability are not incorporated in the predicted levels of reliability.

The empirical relations presented in the previous section were built into this post-processing module. Based on a LMS/NRM scenario, this module calculates the value of the reliability indicator for each origin-destination pair. However, due to the iterative assignment process in the LMS/NRM, multiple routes can be assigned to people travelling between an origin and a destination. Our post-processing module repeats this route assignment and stores all routes in each iteration step. Once the final link travel times have been calculated, our module loops back to all these routes and calculates the reliability for each of them using equation (9) and the coefficients from Tables 2 and 3. The final value of the reliability indicator for an origin-destination pair is an average of the reliability indicators in each iteration step weighted by the flow assigned in that step.

If a route travels over both highways and other roads, the reliability indicator is calculated for both road types separately. The total reliability for this route is the root of the squared sums of these two reliability indicators. Implicitly, we have assumed here that travel time delays on highways are not correlated with those on other routes. A (limited) analysis of our data has shown that this correlation is indeed small, so this is a reasonable assumption.

The module also calculates a value for the national (or regional) reliability indicator by adding the standard deviations of all origin-destination pairs weighted by their traffic flows. These totals are calculated for each time-of-day period and can be added to get a reliability indicator for a whole day, using a weight of 2 for both peak values. We analysed 24-hour data to derive that the mid-day period should get a weight of 9.5 to get a correct daily total.

A test run with this new module revealed that for 2004 the reliability indicator (i.e. the summed standard deviations over all origin-destination pairs) was 48,400 hours for one hour in the morning peak. 60% of this originated from highways and 40% from other roads. The corresponding LMS-run

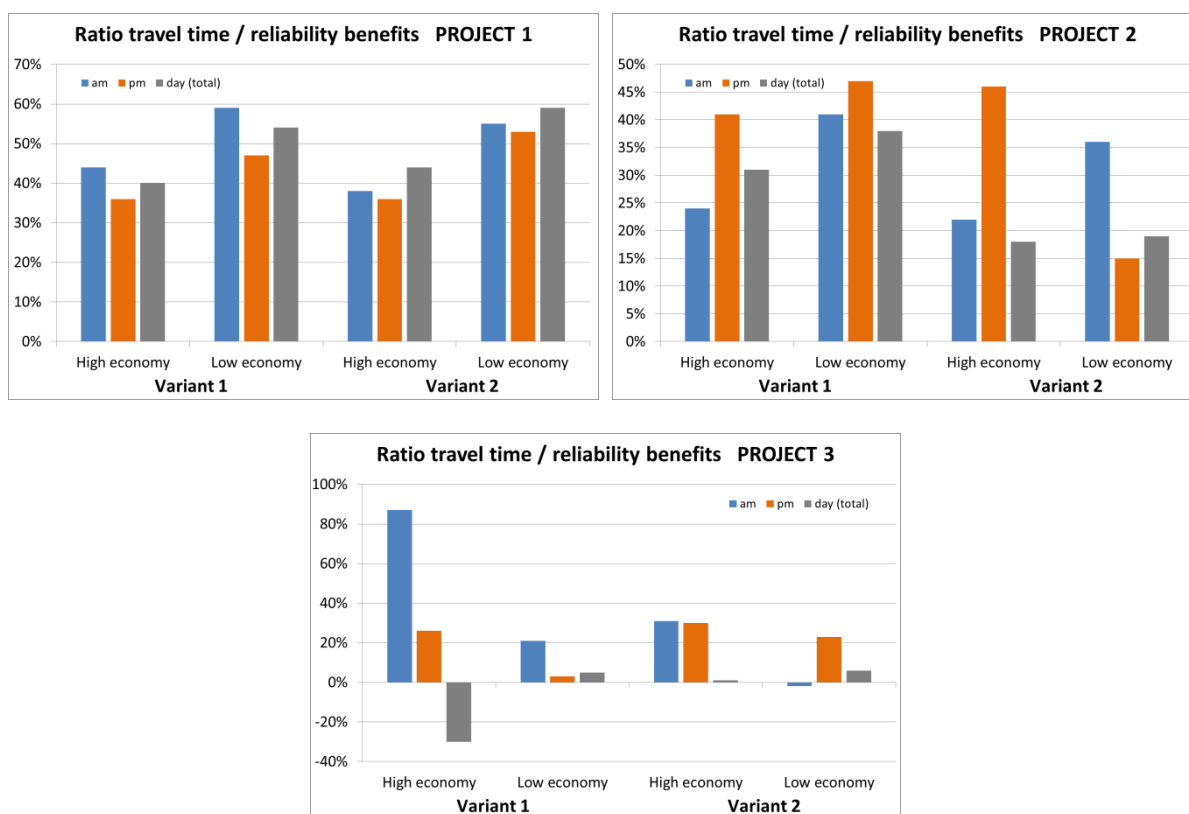
showed that the total delay for all travellers in one hour in the morning peak was 77,000 hours, so the ratio for the (national) reliability to the travel time delays is 63%.

The new module clearly makes a distinction between travel time gains due to shorter routes (e.g. a bypass) or due to reduction of congestion. The former does not lead to reliability gains, whereas the latter does. If some mild congestion occurs on the bypass, reliability may even deteriorate, while overall travel times go down. The new module also takes the exchange of traffic between highways and other roads into account. When congestion is reduced on a highway, this may cause diversion of traffic from secondary roads which can lead to non-standard amounts of reliability gains and losses due to the different relations to mean delay on both types of roads. Also, a lower maximum speed will lead to a reduction of mean delays and hence better reliability which is indicative for the more uniform traffic flow (though we have not yet tested the size of the effect of this policy with observed data).

Impact on CBA results

As a test, a research team from 4Cast simulated the reliability effects of several future infrastructure projects with the new reliability forecasting model described in this paper. The reliability benefits (in terms of euro, i.e. the changes in reliability and multiplied by the VTTRS) appeared to be between 15% and 60% of the travel time benefits, though higher and lower values also occurred (depending on the project, the time-of-day period and the economic scenario, see Figure 12). These numbers are of the same order-of-magnitude as the initial rule-of-thumb of adding 25% to the travel time benefits (CPB 2004). The range between the projects is mainly caused by the differences in trip length and the amount of traffic.

Figure 12. Ratio of travel time benefits and reliability benefits for three projects, each for two variants and two economic scenarios (high and low, figure adapted from 4Cast)



Conclusions and future steps

The most important findings of this paper are:

- When fitting functions between reliability and parameters that are available in demand models, distinction should be made between variation between 15-minute periods and variation between routes as each may be described by a different function.
- For our data set consisting of 250 routes on Dutch highways for three periods of the day (morning peak, evening peak, mid-day), the best empirical relation to describe reliability was an expression of the standard deviation as a function of the mean delay, the logarithm of the mean delay and the length of the route. Other functional forms that have been described in the literature either had much lower adjusted R^2 or showed behaviour that was not supported by the data.
- If observed travel times over multiple days are used to compile a travel-time distribution, a decision needs to be made whether to exclude outliers and whether to correct for variation in travel-time expectation. Consistency with the congestion functions in the demand model and with the method used to determine the valuation of travel-time reliability should be leading in this decision.
- Excluding outliers can have a profound impact on the standard deviation and on the coefficients of the empirical relation. In our project, a criterion of three times the standard deviation above the mean travel time with a minimum of 150% of the mean separated the visually clear outliers from the tail of the standard travel-time distribution. On average, the travel time on 4 out of 251 working days exceeded this criterion. Excluding them reduced the standard deviation by 29%.
- Correcting for variation in travel-time expectation also reduced the standard deviation by another 12%. This depends on the method used to calculate the expected travel times. Very little research is available to support the selection of a method for this.
- This unreliability model is built into a post-processing module for the national and regional transport models. These transport models have not been altered but their outputs are used to calculate the changes in the standard deviation of the travel time distribution due to an infrastructure project.
- The post-processing module calculates the reliability benefits (i.e. the changes in the standard deviation expressed in hours and multiplied by the VTTRS) which can be used in a CBA.
- Better capturing the effects of policies that affect travel time variability in CBA will encourage proper consideration of options. Project appraisal will then not only offer incentives for policies that reduce the average travel time but also for policies that disproportionately improve travel time variability.

Future steps

Reliability will be better embedded in the transport policy making process by the following concrete policy actions. First, the new travel time reliability forecasting model will be incorporated in the Dutch guidelines for CBA. Second, the consequences of policies that especially affect travel time variability will be part of CBA. Attribution of an economic value to travel time variability recognises that transport projects can create more value than they have traditionally realised when they invest to

reduce congestion if an improvement in reliability is produced independent of a reduction in travel time. Third, in order to properly consider such investments in the resource allocation decision process they will be included in the investment tradeoff analysis to prioritise, rank and select infrastructure improvement projects. And finally, a guideline on also including the consequences of extreme travel times, network robustness and vulnerability into the decision making process will be developed. However, a special VTTRS to value extreme changes of travel time variability in the CBA does not exist.

Better integration of reliability into transport policy making is synthesized into a short and mid-term strategy as discussed below to improve the post-processing reliability model. However, it is recommended that these future steps are embedded in a long term strategy (10+ years) to be developed for the national and regional models to assess unreliability in CBA. The basis for such a strategy can be to identify the set of policy measures for which evaluations are or likely will be required. These policies should be matched against the capabilities of the set of modeling tools available.

Short term improvements of the post-processing reliability model

- The reliability model only deals with road transport. However, the national Dutch transport model is also capable of forecasting the effects of changes in the average travel times for public transport (train and bus/tram/metro). It should be possible to estimate equations explaining the standard deviation of travel time for public transport from explanatory variables available in LMS or NRM. At the moment of writing this paper, KiM works on a project to measure how different policies affect travel time reliability in public transport chains.
- Dutch highways are well equipped with detector loops providing inputs for the transport model. However, network users make trips on other roads as well. The regression line is fitted on 250 highway routes and 40 routes on other roads. Collecting extra data and expanding the database can improve the regression analysis for non-highway routes.

Longer term improvements of the post-processing reliability model

- Build a specific database for policies which will increase the travel time but may decrease unreliability. These are policies such as changing the maximum speed or ramp metering. Based on this database a specific regression line can be fitted.
- In reality, mode choice, departure time choice and route choice are sensitive to reliability. The post-processing reliability model can be extended with a feed-back loop into the transport model so that the decisions of the network users are impacted explicitly by changes in reliability.
- The standard deviation contains several sources of unreliability, namely due to recurrent congestion, road works, accidents, unexpected weather conditions, and a random component of day-to-day variation in travel times. Extreme events are removed from the data before fitting the function. Therefore the model predicts reliability changes without considering extremes. Analysing the extreme events, will provide insight in the robustness and vulnerability of the network. However, a special VTTRS to value extreme changes of travel time variability in the CBA would need to be developed through additional primary research.

Acknowledgements

This research was sponsored by the Dutch Ministry of Infrastructure and the Environment. We would like to thank Jasper Willigers (Significance) and Nick Bel (currently ProRail) for their help with the data analysis. We thank 4Cast for their tests with the newly developed tool.

We thank Arjen 't Hoen, Jan van der Waard (both KiM Netherlands Institute for Transport Policy Analysis), Henk van Mourik (Ministry of Infrastructure and the Environment), Marcel Mulder (Rijkswaterstaat – Water Traffic and Environment, Ministry of Infrastructure and the Environment) and Gerard de Jong (Significance) for their comments and suggestions.

References

- Arup, Bates, J., Fearon, J., Black, I. (2003), “Frameworks for Modelling the Variability of Journey Times on the Highway Network”, London: Arup.
- Besseling, P., W. Groot, A. Verrips, (2004), “Economische toets op de Nota Mobiliteit”, CPB Document 65, CPB, The Hague. (In Dutch).
- Eliasson, J. (2006), “Forecasting travel time variability”, paper presented at ETC 2006, Strasbourg.
- Fosgerau, M. (2015), “The valuation of travel time variability”, draft OECD/ITF discussion paper for the roundtable on Quantifying the Socio-Economic Benefits of Transport, Paris.
- Geistefeldt, J., Hohmann, S., Wu, N. (2014), “Ermittlung des Zusammenhangs von Infrastruktur und Zuverlässigkeit des Verkehrsablaufs für den Verkehrsträger Straße”, Schlussbericht für Bundesministerium für Verkehr und digitale Infrastruktur, März 2014.
- Gilliam, C., Kean Chin, T., Black, I., Fearon, J. (2008), “Forecasting and appraising travel time variability in urban areas”, paper presented at ETC 2008, Noordwijkerhout.
- Hamer, R., G.C. de Jong, E.P. Kroes (2005), “The value of reliability in Transport – Provisional values for the Netherlands based on expert opinion”, RAND Technical Report Series, TR-240-AVV, Netherlands.
- HEATCO (2006), “Developing Harmonised European Approaches for Transport Costing and Project Assessment”, Deliverable 5, Proposal for harmonized guidelines. IER, University of Stuttgart.
- Hellinga, B.(2011), “Defining, measuring, and modeling transportation network reliability”, report written for the Dutch Ministry of Transport, Rijkswaterstaat, Delft.
- Hyder Consulting, Black, I., Fearon, J. (2008a), “Forecasting Travel Time Variability in Urban Areas”, Deliverable 1: Data Analysis and Model Development, Department for Transport, United Kingdom.
- Hyder Consulting, Black, I., Fearon, J. (2008b), “Forecasting Travel Time Variability in Urban Areas”, Deliverable 2: Model Application, Department for Transport, United Kingdom.
- Jong, G.C. de and M.C.J. Bliemer (2015), “On including travel time reliability of road traffic in appraisal”, *Transportation Research A*, 73, 80-95.
- Jong, G.C. de, Kouwenhoven, M., Bates, J., Koster, P., Verhoef, E.T., Tavasszy, L., Warffemius, P.M.J. (2014), “New SP-values of time and reliability in freight transport in the Netherlands”, *Transport Research Part E*, 64, pp 71-87.
- KiM Netherlands Institute for Transport Policy Analysis (2013), “The social value of shorter and more reliable travel times”, <http://www.kimnet.nl/en/publication/social-value-shorter-and-more-reliable-travel-times>.

- Kouwenhoven, M., de Jong, G.C., Koster, P., van den Berg, V.A.C., Verhoef, E.T., Bates, J., Warffemius, P.M.J (2014), “New values of time and reliability in passenger transport in the Netherlands”, *Research in Transportation Economics*, 47, pp. 37-49.
- Mahmassani, H.S., Kim, J., Chen, Y., Stogios, Y., Brijmohan, A., Vovsha, P. (2014), “Incorporating reliability performance measures in planning and operations modelling tools”, SHRP2 report S2-L04-RR-1, Transportation Research Board, Washington DC.
- Ministry of Infrastructure and the Environment (2012), “Structuurvisie Infrastructuur en Ruimte (SVIR)”, The Hague, <http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2012/03/13/structuurvisie-infrastructuur-en-ruimte.html> (In Dutch, accessed on 8 December 2015).
- Ministry of Transport (2014), “Transport analysis guidance: WebTAG”, <https://www.gov.uk/guidance/transport-analysis-guidance-webtag> (accessed on 8 December 2015).
- OECD/ ITF (2010), *Improving reliability on surface transport networks*, OECD Publishing, Paris. DOI: <http://dx.doi.org/10.1787/9789282102428-en>
- Peer, S., Koopmans, C. and Verhoef, E.T. (2012), “Prediction of travel time variability for cost-benefit analysis”, *Transportation Research Part A*, pp. 79-90.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. (2007), *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., New York: Cambridge University Press.
- Significance, VU University Amsterdam, & John Bates. (2007), “The value of travel time and travel time reliability, survey design”, Final Report prepared for The Netherlands Ministry of Transport, Public Works and Water Management, Significance, Leiden.
- Significance (2011), “Schattingen van keuzemodellen voor het LMS 2010”, Report prepared for RWS, Significance, Den Haag.
- Sirivadidurage, S., Gordon, A., White, C. and Watling, D. (2009), “Forecasting day to day variability on the UK motorway network”, presented at ETC 2009, Noordwijkerhout.
- TRB (2003), “Interim Planning for a Future Strategic Highway Research Program”, NCHRP national Cooperative Highway Research Program, Report 510, Washington DC. ISBN 0-309-08777-5.
- Willigers, J. and M. de Bok (2009), “Updating and extending the disaggregate choice models in the Dutch national model”, paper presented at the ETC 2009, Noordwijkerhout.

Notes

- 1 Fosgerau (2015) prefers to use the variance of travel time rather than the standard deviation. In his paper, he shows that the variance is theoretically more appropriate for commuters with flexible work times. Furthermore, using the variance has the advantage that it is additive over links (provided that the travel times on the links are independent). However, for this study we prefer to use the standard deviation since (a) it is consistent with the Dutch valuation study, (b) most travellers in the peaks are commuters with inflexible work times and (c) the typical link lengths in our study are so short that the travel times on adjacent links are certainly correlated.
- 2 Extreme travel times are removed from the data before fitting the function. Therefore, our model predicts reliability changes without considering these extremes. Analysing extreme events (separately) will provide insight in the robustness and vulnerability of the network.

Policies may affect the reliability, but also may affect the robustness and the vulnerability of the network. Both should be included in a cost-benefit analysis. The tool that we describe in this paper only looks at the reliability component without the extreme events. Including the extreme events in this or in a separate tool is one of the longer term improvements of the post-processing reliability model.
- 3 All estimations in this paper were made using the LFIT algorithm of Numerical Recipes (Press et al. 1992)

