DEVELOPMENT OF A TOOL TO ASSESS THE RELIABILITY OF DUTCH ROAD NETWORKS

Marco Kouwenhoven
RAND Europe
Arnout Schoemakers
Transport Research Centre
Rik van Grol, Eric Kroes
RAND Europe

1. INTRODUCTION

The Netherlands has about 2500 km of highways (or A-roads, since they are named A1, A2, etc.). Given the size of the Netherlands this results in a very dense network, especially in the Randstad area (the main economic area connecting the cities of Amsterdam, Rotterdam, The Hague and Utrecht).

Despite the high road network density, it is still not enough to satisfy traffic demand, resulting in long traffic jams during the peak hours. Traffic jams with a summed length of 200 to 300 kilometres are quite normal, though this number is highly variable. This results in longer and very unpredictable travel times.

The objective of the study presented in this paper is to develop a simple and pragmatic tool to forecast future levels for the reliability of journey times. This tool is based on an empirical model, fitted on road induction loop data and is coupled to other models predicting future travel demand.

2. POLICY BACKGROUND

In September 2004, the Dutch government published its new “Dutch Mobility Policy Document” (Ministry of Transport, Waterworks and Infrastructure, 2004). This document contains a global outline of national transport and traffic policy for the coming decades. At the policy’s core is the realisation that mobility is an essential prerequisite for economic and social development. A properly functioning transport system for people and goods with reliable access times is necessary to strengthen the Dutch economy and the international competitive position of the Netherlands. The Mobility Policy document explains how the Ministry plans to achieve these aims.

2.1 Reliability in the Dutch Mobility Policy Document

A high quality road network, as carrier of the spatial main structure, is crucial for an internationally competitive climate for the establishment of businesses and for efficient road transport. The central government strives to achieve reliable and acceptable journey times from door to door. The national road network and the regional road network are approached as a single coherent network. Road administrators and private parties at local levels need to find solutions for traffic jams together. The state will introduce measures to increase reliability and limit delays, including maintenance, incident management and topical travel information and utilisation measures, as well as extra construction measures. Priority will be given to the main links, particularly the triple A links (The A2, A4 and A12 that form the main national and international connections)

2.2 Road transport: reliable and fast

Increasing urbanisation and commercialisation are resulting in extra mobility in the Netherlands. In 2020 road traffic is expected to have grown by more than 40% compared to 2000 (both on the national and the regional road network). The growth is most prominent for short distance journeys (up to thirty kilometres) and for journeys in urban areas. In all regions, national road network accounts for the most kilometres.

Road capacity is reaching its limits. The existing national road network is unable to handle this growth in mobility during peak hours. The intensity of road usage is increasing steadily, with particularly high traffic levels in the Randstad. Without extra measures, this will cause increases in journey time and decreases in predictability.

Journey times will become less predictable. Actual journey times often differ from the journey time estimates that citizens and companies make in advance. This means that road users do not know what to expect. Without extra policy, the reliability on the national road network will decrease strongly towards 2020.

Decreasing reliability is a direct result of more intensive use of the road. Incidental events such as accidents, weather conditions and road works are more apt to cause traffic jams when roads are used more intensively.

2.3 Social costs

Traffic jams have a direct financial impact on companies. The social costs of traffic jams will amount to 1.7 milliard euros in 2020. These costs include costs related to direct journey time loss, economically indexed according to the journey purpose. These costs increase even further due to unreliable journey times and negative driving behaviour motivated by a desire to avoid traffic jams. The total social costs without changes in policy will therefore rise to approximately 2.5 milliard euros in 2020.

2.4 Ambition

One of the ambitions in the Mobility Policy Document is to increase reliability and decrease journey times from door to door. The cabinet's ambition is increase reliability on the national road network so that 95% of all movements in 2020 during rush hour are on time (see section 3 for a precise definition of this target).

2.5 Tool to assess the reliability of Dutch road networks

Currently the AVV Transport Research Centre (part of the Rijkswaterstaat organisation) is working on defining and quantifying terms as predictability, reliability and robustness. In 2003, a quantitative study was started about the size and origin of reliability on road and rail (Goudappel Coffeng, 2003). This study concluded that the traffic demand is the main cause of variety in travel speed and travel times. Besides this, the weather conditions, road works and accidents have a substantial effect on reliability.

The objective of the study presented in this paper is to develop a tool to predict future travel time reliability of networks. This tool needs to function as a post processor to the National Model (LMS) and the New Regional Model (NRM, regional versions of the LMS). These models are used to calculate future traffic demand and congestion under different policies and scenarios. The new tool (Kouwenhoven et al., 2004), which will be based on an empirical model, can be used for national policy documents and highway infrastructure studies.

3. RELIABILITY INDICATORS

3.1 Definitions

For this study we used four indicators that are related to reliability. The first reliability indicator was defined in the Dutch Mobility Policy Document as the probability that a trip is “on time”. “On time” means:

- for a short trip (less than or equal to 50 minutes): deviation from the expected travel time is less than 10 minutes;
- for a long trip (longer than 50 minutes): deviation from the expected travel time is less than 20% of the expected travel time.

The expected travel time is defined as the median of all travel times for the same trip as observed on all working days in a year at the same period of day.

One can argue that having a journey time that is shorter than expected (“arriving early”) will be valued differently from having a journey time that is longer than expected (“arriving late”), though in both cases a traveller might not be able to use his time in the way that he would have preferred. To be able to differentiate between unreliability caused by journeys that are longer than expected and journeys that are shorter than expected, we have taken the probability that a journey is “not too long” (instead of “on time”) as a second reliability indicator.

The last two indicators are the percentile-10 and the percentile-90 of the speed distribution. These are related to the width of the distribution and are therefore also related to reliability. In our final tool, they are used to estimate the value of unreliability.

3.2 Illustration

To illustrate the standard policy indicator for reliability we look at two highway tracks of about 25 kilometres during the morning peak. The first track is between Waddinxveen and The Hague (A12) within the Randstad (the main economic area in the Netherlands and suffering from heavy congestion), the second track is between Gorichem and Breda (A27), well outside the Randstad where the traffic situation is much less severe.

Figure 1 shows the distribution of the average travelling speed during the morning peak over 154 working days in 2002. For the A12 track (left figure) the distribution is very wide. The median travel time corresponds to a median travelling speed of about 50 km/h. When travelling at a speed below about 38 km/h or when travelling at a speed above about 75 km/h, the journey time is more than 10 minutes longer or shorter than expected and a traveller is said to be not “on time”. This occurs in about 42% of all trips in the morning peak. On the A27 track (right figure) the travel time is much more predictable: only 4% of all journeys is not “on time”. From Figure 1 we can conclude that the travel times on the A27 track are much more reliable than on the A12 track.

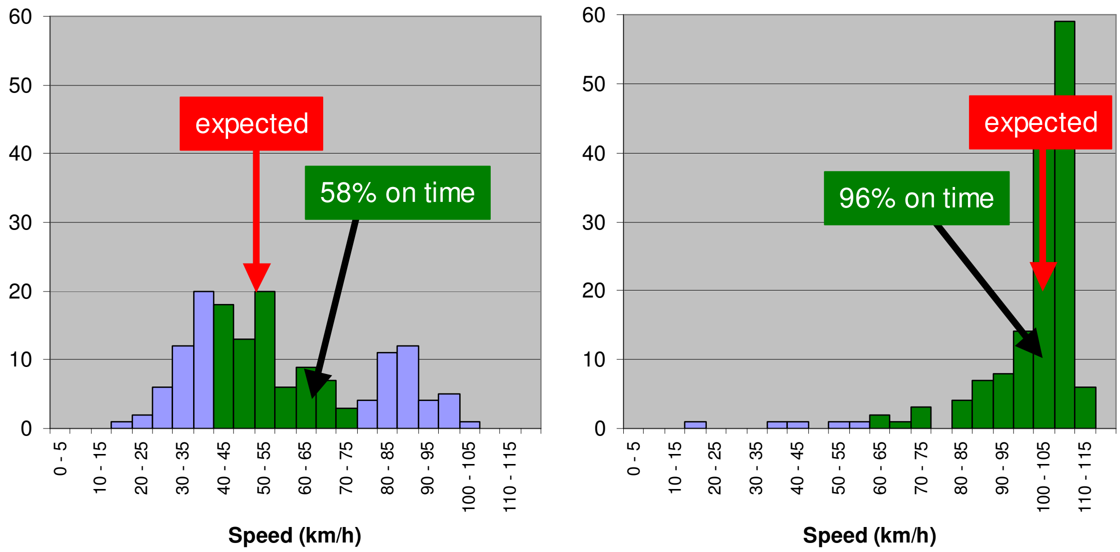

Figure 1: Speed distribution during the morning peak for two tracks in the Netherlands

3.3 Data

A considerable part of the Dutch road network is equipped with induction loops. From these loops we have speed and intensity averaged over 15-minute periods. This information can be used for empirical relations that we need to estimate, since similar variables are available in the national and regional models.

On the selected network some 212 routes were identified and used for the estimation. These routes had lengths between 2 and 120 kilometre and the distribution over all lengths was similar to the observed distribution of journey lengths. The selected network consists of only highways, as empirical information for the secondary road network is very limited.

From an entire year of data (2002) 154 days of valid data were selected. Although official public holidays and weekends have been omitted, the data does cover both normal workdays as well as the summer holiday period. Since traffic delays (and hence, unreliable journey times) outside the peak hours are uncommon, we selected data from both peaks only (7:00 –9:00 and 16:00 – 18:00).

For each route and for each peak, we calculated (over all selected days)[1]:

- median speed V_{median}
- median travel time T_{median}
- mean speed V_{mean}
- mean travel time T_{mean}

Furthermore, for each route the length L and the mean speed limit $V_{max,mean}$ are available. Note that the standard speed limit on Dutch highways is 120 km/h, which is reduced to 100 km/h (sometimes even 80 km/h) around the major cities and road interchanges.

Using the definitions from Section 3.1, the four indicators for each route and for each peak were calculated[2]

- reliability indicator = probability of a trip being “on time”: P_{OnTime}
- alternative reliability indicator = probability of a trip being not “too long”: $P_{NotTooLong}$
- percentile-10 of the speed distribution: V_{10}
- percentile-90 of the speed distribution: V_{90}

These four indicators are shown in Figure 2 as a function of the length of the route divided by the mean journey time (which is similar, but not equal to the mean velocity, see footnote 2). From this figure it can be seen that the standard reliability indicator P_{OnTime} is almost 100% for routes with high average speeds (i.e. high values of L / T_{mean}). For medium values of L / T_{mean} the reliability indicator is very widely distributed. For low values of L / T_{mean} the reliability indicator seems to increase again, however, this might be only due to a lack of data points. The alternative reliability indicator $P_{NotTooLong}$ shows similar behaviour.

Both percentiles V_{10} and V_{90} are close to the value of L / T_{mean} for high speeds. Below a value of 100 km/h the percentiles start to deviate clearly. The percentile-90 seems to remain constant at a value of about 100 km/h, whilst the percentile-10 seems to remain at a constant difference of about 20 km/h below the value of L / T_{mean}.

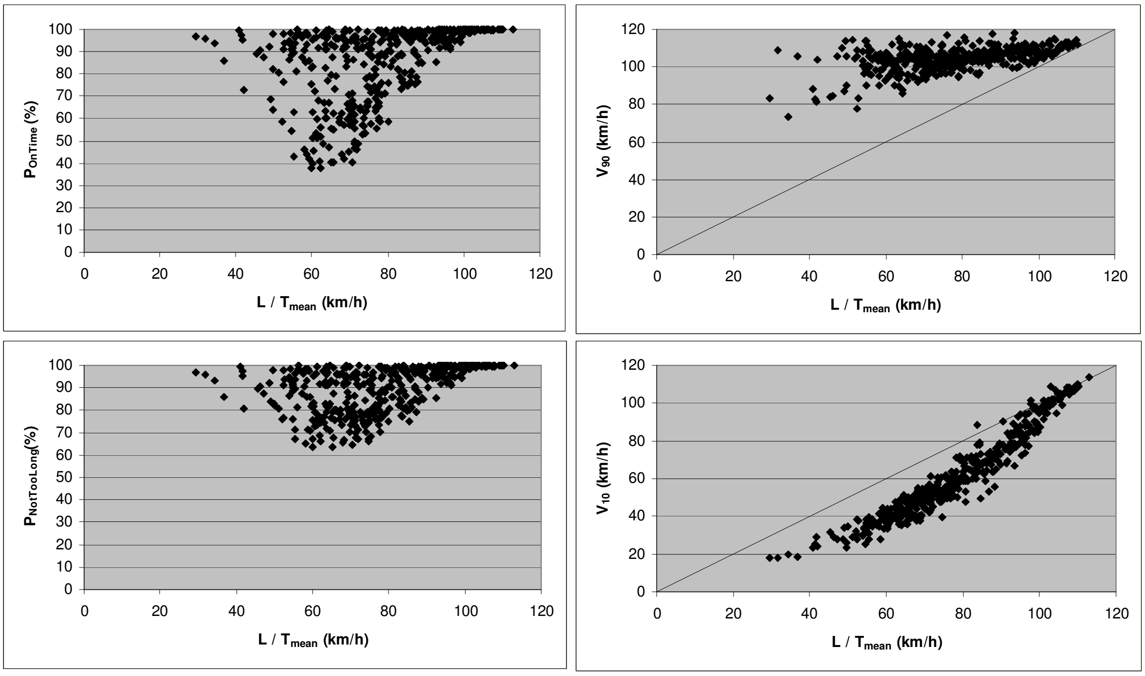

Figure 2: Four reliability indicators in the dataset as a function of the journey length divided by the mean journey time.

4. MODEL ESTIMATION

4.1 Approach

The national and regional models to which the new tool needs to be connected (see Section 2.5), calculate future traffic demand and congestion under different policies and scenarios. These models predict:

- intensity (number of cars per hour) and
- travel speed

on each major road in the Netherlands (characterised a.o. by route length and speed limit). Therefore, the underlying empirical model for the new reliability tool is restricted to use only available variables (e.g. intensity, travel speed, route length, speed limit) as its input. The objective of the modelling effort in this study is to find empirical functions that relate the four reliability indicators to these national and regional model variables (travel speed, route length, speed limit), i.e. finding a function

$$\text{Indicator} = f\,(V_{median}, T_{median}, V_{mean}, T_{mean}, L, V_{max,mean})$$

A least square methods has been used to find the best functional form and the best coefficients.

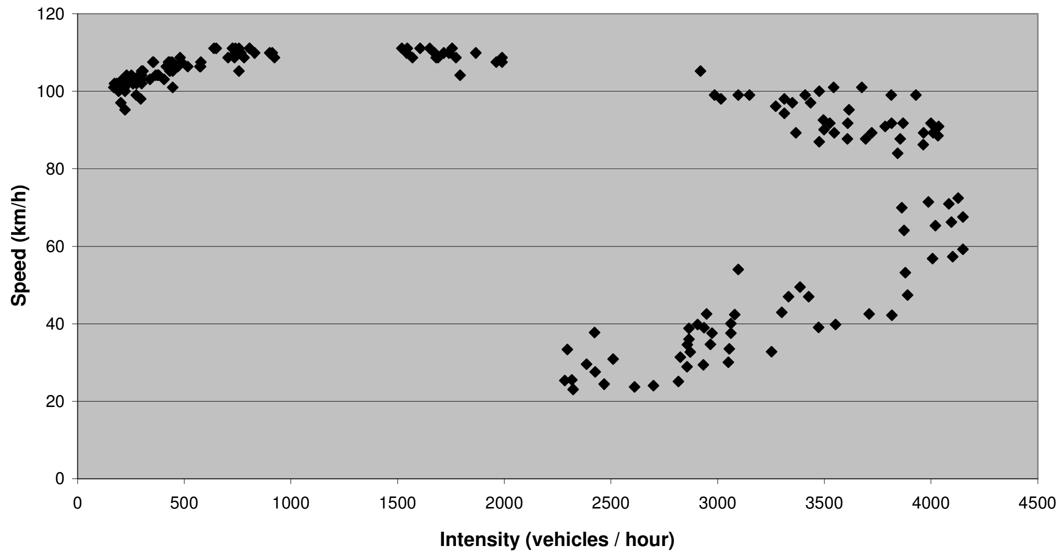

Figure 3: Speed-flow diagram as observed on a two-lane highway (A12 between Waddinxveen and The Hague)

We decided not to use the intensity as an explanatory variable in the modelling process, since this variable has the well-known “dual” relation with the real traffic situation, as can be seen from a speed-flow diagram (Figure 3). In quiet traffic situations (e.g. at night), the intensities are low and the speeds are close to the speed limit (upper left corner in Figure 3). When the traffic starts to increase, the intensity increases but the speed remains about constant. As the intensity approaches the capacity of the highway (for a two-lane highway the capacity is about 4000 – 4500 vehicles per hour), the speed starts to drop. If the traffic becomes even denser, a traffic jam will start: speed will go down even further, but the intensity will drop as well. So, for a given intensity there are two possible traffic situations: a quiet traffic situation with high speeds and a congestion situation with low speeds. Therefore, intensity is not a good explanatory variable for use in our modelling exercise.

4.2 Results

It turned out that the route length divided by the mean journey time L / T_{mean} was the best speed variable to use for the modelling. Model fits improved when this variable was corrected for the speed limit. This correction was to be expected for the following reason: suppose that on two different routes a mean speed of 100 km/h is observed; on the first route the local speed limit is 100 km/h, so there are no delays and a high reliability is expected; on the second route the local speed limit is 120 km/h, so there is some small delay and a lower reliability is expected. From this example we conclude that it is not the observed speed alone that can explain reliability, but the difference between the observed speed and the speed limit that should be used as the explanatory variable. For this reason, we used a corrected speed V* in our fits:

$$V^* = L / (T_{mean} - T_{V=Vmax,mean} + T_{V=120km/h})$$

The observed mean travel time is corrected using the travel time when travelling at the local speed limit and the travel time when travelling at a standard speed of 120 km/h.

The remaining vertical spreads in P_{OnTime} and $P_{NotTooLong}$ after this speed correction (these spreads are similar to the spreads in the left diagrams in Figure 2) are strongly related to the route length: longer routes have lower reliability indicators. The difference between the observed probability and 100% turned out to be proportional to the journey length L, up to some maximum. This maximum above which the reliability was independent from the journey length was determined by the fitting routine (about 64 km for P_{OnTime} and about 49 km for $P_{NotTooLong}$.

The remaining vertical spreads in the percentiles V_{10} and V_{90} after the speed correction mentioned above could be reduced by an extra correction for the local speed limit.

Several functional forms were tried: linear, bi-linear, logistic, Gaussian, etc. For all indicators the bi-linear form produced the best fits, except for the percentile-10 for which a logistic function was better.

The functional forms that produced the best fits are (all greek symbols are coefficients that have been fitted):

$$P_{OnTime} = \min(100, \alpha \cdot \min(L, \gamma) \cdot V^* + \beta)$$

$$P_{NotTooLong} = \min(100, \alpha \cdot \min(L, \gamma) \cdot V^* + \beta)$$

$$V_{10} = (L / T_{mean}) - \alpha / [1 + \exp\{\beta \cdot (V^* - \gamma)\} + \delta \cdot (120 - V_{max,mean})]$$

$$V_{90} = \max(L / T_{mean}, \alpha \cdot V^* + \beta + \gamma \cdot (120 - V_{max,mean}))$$

$$\textit{with} \quad V^* = L / (T_{mean} - T_{V=Vmax,mean} + T_{V=120km/h})$$

with the coefficients shown in Table 1. The last line in the table gives a measure for the goodness of the fit (the square root of the mean value of the squared residuals).

	P_{OnTime}	$P_{NotTooLong}$	V_{10}	V_{90}
α	-0.02314	-0.01358	21.76825	0.13045
β	2.365363	1.511752	0.193074	97.86791
γ	63.97739	49.08467	98.39749	-0.60988
δ			0.11347	
$\sqrt{[(1/N) \cdot \Sigma (X - f(x))^2]}$	*4.30*	*3.33*	*4.72*	*4.43*

Table 1: Best fit coefficients and a measure for the goodness of the fit

Figure 4 shows the best fitting models. Since these models depend on the length and the mean speed limit of each route, all data points have been scaled to a length of 20 km and a speed limit of 120 km/h.

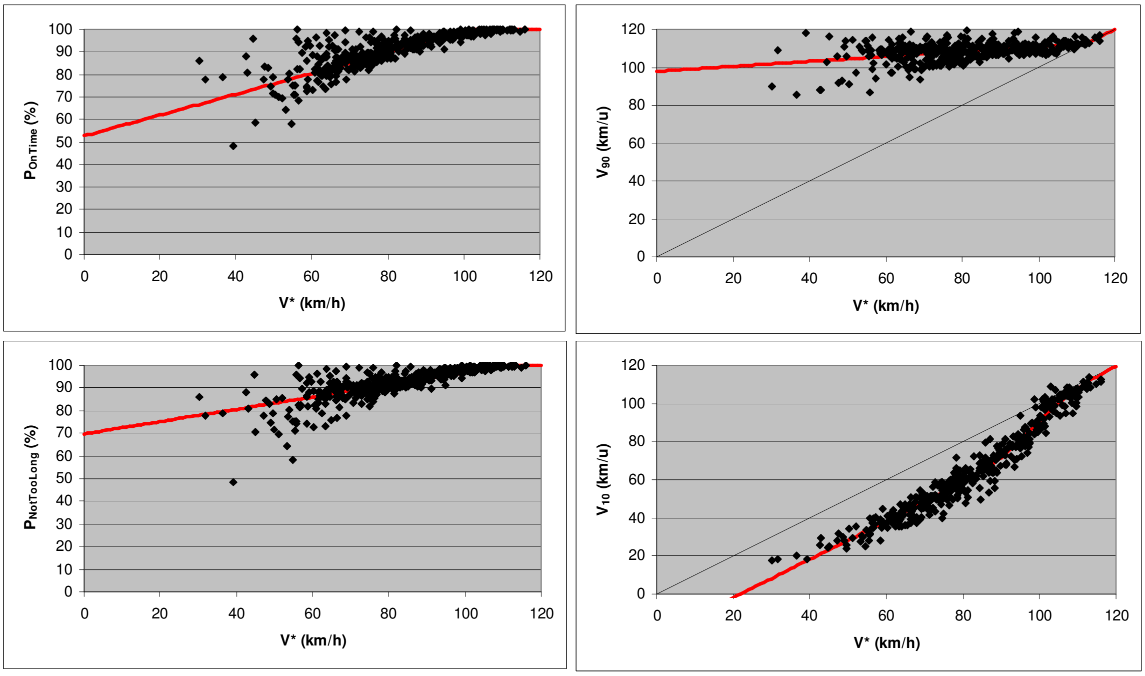

Figure 4: **Four reliability indicators in the dataset as a function of the corrected speed V*. All data points have been scaled to a standard route length of 20 km/h and a standard speed limit of 120 km/h. The thick lines indicate the best fitting model**

4.3 Interpretation

Both the standard policy reliability indicator P_{OnTime} and the alternative indicator $P_{NotTooLong}$ are strongly related to (some kind of) speed. This suggests that the best way to improve the reliability is to increase the average speed on the highways, i.e. relieving congestion. So, improving journey times and reliability are strongly correlated.

However, this is not the whole story. The fitted curves suggest that reducing the speed limit can also improve the reliability, though this conclusion cannot be drawn from the current data set. The set only contained a limited number of routes with a speed limit under 120 km/h, and for those routes who had a reduced speed limit, this limit only applied to short parts of the routes.

Another way to improve the reliability is to change exogenous factors, like the impact of incidents and of road works, as will be discussed in the next section.

4.4 Exogenous Factors

The data set also contained the number of days with a (normal) accident (including car break downs at the shoulder of the road), with a large accident (causing severe traffic jams), with road works and/or with rain along each of the routes. From this we calculated the probability of encountering a road accident (normal or large), road works or rain for each of the routes. We hypothesised that routes with higher probabilities of (for instance) road accidents would have less reliable journey times.

We tested this hypothesis and checked whether differences in these probabilities explained some of the remaining variation in the reliability indicators. However, no clear correlation was found.

This might seem striking at first sight, but it can be understood as follows: compare two routes of equal length and with equal speed limits, the first with a higher probability of road accidents than the second. The first route will have both a lower reliability indicator and a lower mean speed as a result of the higher number of road accidents. However, both routes will follow the same relation between mean speed and reliability as was found before, see Figure 5. Therefore, the fitted relation cannot be improved by using the probability of having road accidents, road works or rain as an extra variable.

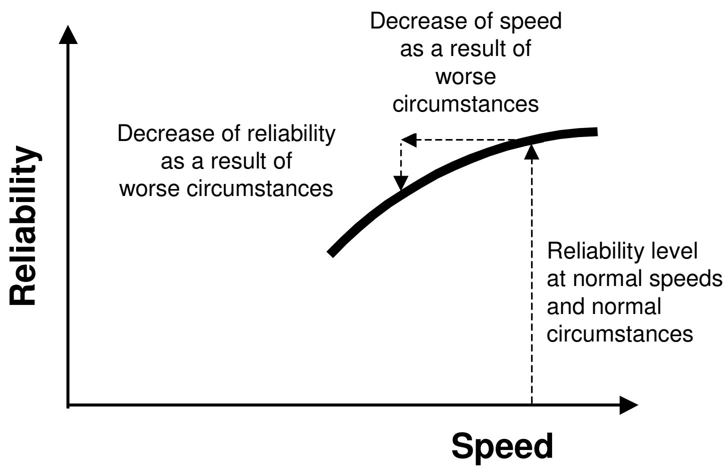

Figure 5: Effect of worse circumstances on both speed and reliability

However, we were able to estimate the amount of shift along the curve due to different probabilities of the exogenous factors. The speed can be described as:

$$V^* = V_{perfect} - \Sigma f_{exo} \cdot \Delta V_{exo}$$

in which $V_{perfect}$ is the mean speed under perfect conditions (no accidents, road works, rain in the whole year); f_{exo} is the frequency with which a certain exogenous factor occurs and ΔV_{exo} is the impact of the exogenous factor on the mean speed. Values for these variables are given in Table 2. We note that the values for road works might be biased, since they are based on the rush hour data only, and it is known that these are usually planned outside the peak, especially larger road works. Those road works that take place during the peak are usually minor and are at locations with less dense traffic.

	f_{exo}	ΔV_{exo}
Accident	0.1184	19.16
Large accident	0.0065	54.10
Road works	0.1164	5.34
Rain	0.0716	10.56

Table 2: Frequency and impact of some exogenous factors (averaged over all routes)

Incident management and clever road works planning might reduce the frequency and/or the impact on the mean velocity. As a result of a slightly increased mean velocity the reliability indicators will increase. However, the final impact on reliability is rather limited: if ΔV_{exo} is halved for accidents (as a result of improved incident management), the mean velocity will increase by about 1.1 km/h over a whole year and as a result, the standard policy reliability indicator will increase by about 0.5% (for a route of 20 km length).

5. CONCLUSIONS

We have been able to find empirical relations for four reliability indicators

- reliability indicator = probability of a trip being “on time”: P_{OnTime}
- alternative reliability indicator = probability of a trip being not “too long”: $P_{NotTooLong}$
- percentile-10 of the speed distribution: V_{10}
- percentile-90 of the speed distribution: V_{90}

The variables for this function were:

- mean travel time T_{mean}
- mean speed limit $V_{max,mean}$
- journey length L

This function has been used to construct a tool called LMS-BT that is able to predict future reliability levels. It is a simple and pragmatic tool to forecast future levels for the reliability of journey times, based on an empirical model. This tool is post-processing tool: national and regional models (LMS and NRM) are used to predict future traffic demand and congestion under different scenarios, the output of these runs are used as input for the LMS-BT tool to predict future travel time reliability in terms of the policy reliability indicator P_{OnTime} and its alternative indicator $P_{NotTooLong}$. In addition, the tool also calculates the cost of unreliability with the help of the predicted percentile-10 and percentile-90 of the speed distribution and an input Value of Reliability (VOR). The VOR are based on a separate study (Hamer et al., 2005).

This tool is also capable of predicting reliability impacts for a range of exogenous factors, like accidents (normal and large), road works and rain. When the frequency of occurrence or the impact of an exogenous factor is changed, the mean speeds (as predicted by the national and regional models) are corrected and new values for the reliability are calculated.

Increasing the mean speed on the highways (i.e. reducing congestion) is a direct method to improve the reliability of travel times. Improved incident management and better road works planning also has a positive (indirect) effect on reliability. On the basis of the current research, this effect seems rather limited, however, the underlying data set might be biased so that we cannot draw any firm conclusions on this. There are also indications that reducing the maximum speed leads to more reliable journey times, but further research is necessary to confirm this.

REFERENCES

Goudappel Coffeng (2003), *Betrouwbaarheid van Reistijden*, Ministry of Transport, Waterworks and Infrastructure – Adviesdienst Verkeer en Vervoer, Rotterdam.

Hamer, R., Jong, G. de, Kroes, E.P. (2005), *The Value of Reliability in Transport — Provisional values for the Netherlands based on expert opinion,* RAND Technical Report Series, TR-240-AVV.

Kouwenhoven, M.L.A., Grol, H.J.M. van, Kroes, E.P. (2004), *Ontwikkeling ex ante Instrument Betrouwbaarheid,* RAND Technical Report Series, TR-233-AVV.

Ministry of Transport, Waterworks and Infrastructure (2004), *Nota Mobiliteit. Naar een betrouwbare en voorspelbare bereikbaarheid, deel 1,* The Hague.

NOTES

[1] Note that T_{median} and V_{median} are directly related ($V_{median} = L / T_{median}$), but that this is not true for T_{mean} and V_{mean} : $V_{mean} = (\Sigma V) / N = (\Sigma (L / T)) / N \neq L / T_{mean} = L / (\Sigma T / N)$

[2] Induction loop data on speed and intensity was available with a 15-minute resolution. These speeds were averaged over a peak by taking the mean weighed with the intensity. The indicators P_{OnTime} and $P_{NotTooLate}$ were first calculated for each period of 15 minutes in each peak, before being averaged over the whole peak (again weighed with the average intensity in each 15-minute period).