TRAVEL TIME RELIABILITY ON FREEWAYS

J.W.C. van Lint, H. Tu, H. J. van Zuylen
Transportation and Planning Department, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, P.O. Box 5048, 2600 GA Delft, The Netherlands, tel. +31 15 2785061, fax +31 15 2783179,
H.vanLint@citg.tudelft.nl, H.tu@citg.tudelft.nl, H.J.vanZuylen@citg.tudelft.nl

Abstract

In contemporary transportation planning route travel times are considered key indicators on the reliability of a road network. Mean and variance of travel times do not provide much insight in the travel time reliability since these metrics tend to obscure rare but relevant high travel times under specific circumstances. We therefore introduce two metrics, based on just three characteristic percentiles (10^{th}, 50^{th} and 90^{th}), which indicate both congestion as well as possible onset (dissolve) of congestion. High values of either metrics indicate high travel time unreliability. We subsequently propose a simple neural network model to predict these percentile values and hence unreliability based on time-of-day (TOD), and day-of-week (DOW).

The conclusion for traffic managers is twofold. The metrics can be used to identify not only the unreliability of travel times for a given DOW/TOD period, but also identify DOW/TOD periods in which it is likely that congestion sets in (or dissolves. Practically, this means identifying the uncertainty of start, end and hence length of morning and afternoon peak hours. Secondly, the neural network based model can serve as a long-term travel time prediction tool, but also as a travel time unreliability prediction tool, using the two metrics presented above on the model outcomes.

Keywords: Travel time; Reliability; Travel time prediction; Reliability metric
Topic area: E2 Performance Measurement

1. Introduction

In contemporary transportation planning route travel times are considered key indicators on the reliability of a road network (Lo, 2002, Cassir et al., 2001). In the Oxford Dictionary *reliable* is defined as "consistently good in quality or performance, and able to be trusted", and reliability as "the quality of being reliable". Although many different definitions for travel time reliability in a transportation network or corridor are proposed (refer to for example (Bell and Cassir, 2000)), in general travel time reliability relates to properties of the (day-to-day) travel time distribution as a function of time of day (TOD), day of the week (DOW), Month of the Year (MOY), and external factors such as weather, incidents and road works. The wider (longer-tailed) the travel time distribution, the more unreliable travel time on a freeway network or corridor is considered. As we will show, the travel time distribution is sometimes sharply skewed, implying that descriptive statistics such as mean and variance (standard deviation) are not very useful in reconstructing the travel time distribution nor as indicators of reliability. Rather, we should focus on *median* and *percentile* values (Bates et al., 2002, Bates, 2001) which provide us with more robust estimates of how likely a specific travel time is given certain circumstances (e.g. TOD, and DOW values), and provide us with easier means to reconstruct the travel time distribution.

In this paper, we explore and aim to predict the distribution of travel times on a densely used freeway corridor, which is part of the northern beltway around the metropolitan area of Rotterdam (The Netherlands). Although no measured travel times (on a large scale) are available, the corridor is equipped with inductive loops about every 500 meters measuring average speeds on a minute basis. A large database of corridor travel times can subsequently be estimated using a speed based travel time estimation algorithm, such as the PLSB trajectory algorithm proposed in (Lint and Zijpp, 2003). With this travel time database we can subsequently look at TOD and DOW trends in travel times and calibrate models to predict the distribution of travel times. The neural network based model we present in this paper could for example be applied in integrated route planning applications on traffic information websites and commercial route planning software.

The article is organized as follows. In the first section characteristics of the day-to-day distribution of travel time is investigated and characteristic metrics describing this distribution are proposed. In the section thereafter we present a neural network based model for predicting the travel time distribution for a given TOD and DOW based on characteristic percentiles. In the final sections we present some results of the model and offer conclusions and further research directions

2. Characteristics of the day-to-day travel time distribution

In the introduction we discussed the *day-to-day* travel time distribution, capturing the variability of travel times for a particular day-of-the-week (say a Monday) within a particular departure time period (for example between 9:00 and 9:15). However, within a single departure time period *k* (in this case 15 minutes) travel times vehicles will experience are also distributed, due to differences between drivers and the complex spatio-temporal dynamics of freeway traffic flow within these periods. In (Lint, 2004a) it is found that within small departure time periods (of 1 minute) this distribution is approximately normal and stationary with respect to both mean and variance, however, in larger departure time periods we may expect this no longer holds.

Nonetheless, here we will explore day-to-day variability travel times *only*, and propose key characteristics to reconstruct and predict this distribution. As we may expect that traffic flow operations on different days are unrelated[1], we can account for the extra within period variability separately. In the recommendations section we will propose a method to do this.

For the analysis below we look at the (estimated) travel times between 6:00 AM and 8:00 PM for one whole year (2002) on the 6,5 kilometre eastbound carriageway of the A20 freeway between Rotterdam Centre and Rotterdam Alexander polder. This densely used free stretch is also part of the northern beltway around the metropolitan area of Rotterdam (The Netherlands). Travel times were estimated with the PLSB trajectory algorithm (Lint and Zijpp, 2003) for every departure minute between 6:00 AM and 8:00 PM, while the TOD period was chosen as 15 minutes. Per 15 minute TOD period the median value (50 percentile) of the available 15 travel times was taken as the travel time for that particular TOD.

Figure 1 shows the variability of travel times on the A20 for three DOW values, that is Thursdays (Figure 1 top), Fridays (Figure 1 middle) and Saturdays (Figure 1 bottom). In the figure 5, 10, 25, 50 (median), 75, 90 and 95 percentiles are shown as a function of time of day

---

1 For example, the dynamics of a traffic jam building and dissolving on day X is not affected or influenced on how a traffic jam builds and dissolves on day Y

(TOD). A 75 percentile value of say 10 minutes for TOD 16:00-16:15 reflects the percentage of days on which during that TOD a median travel time of 10 minutes or less occurred.

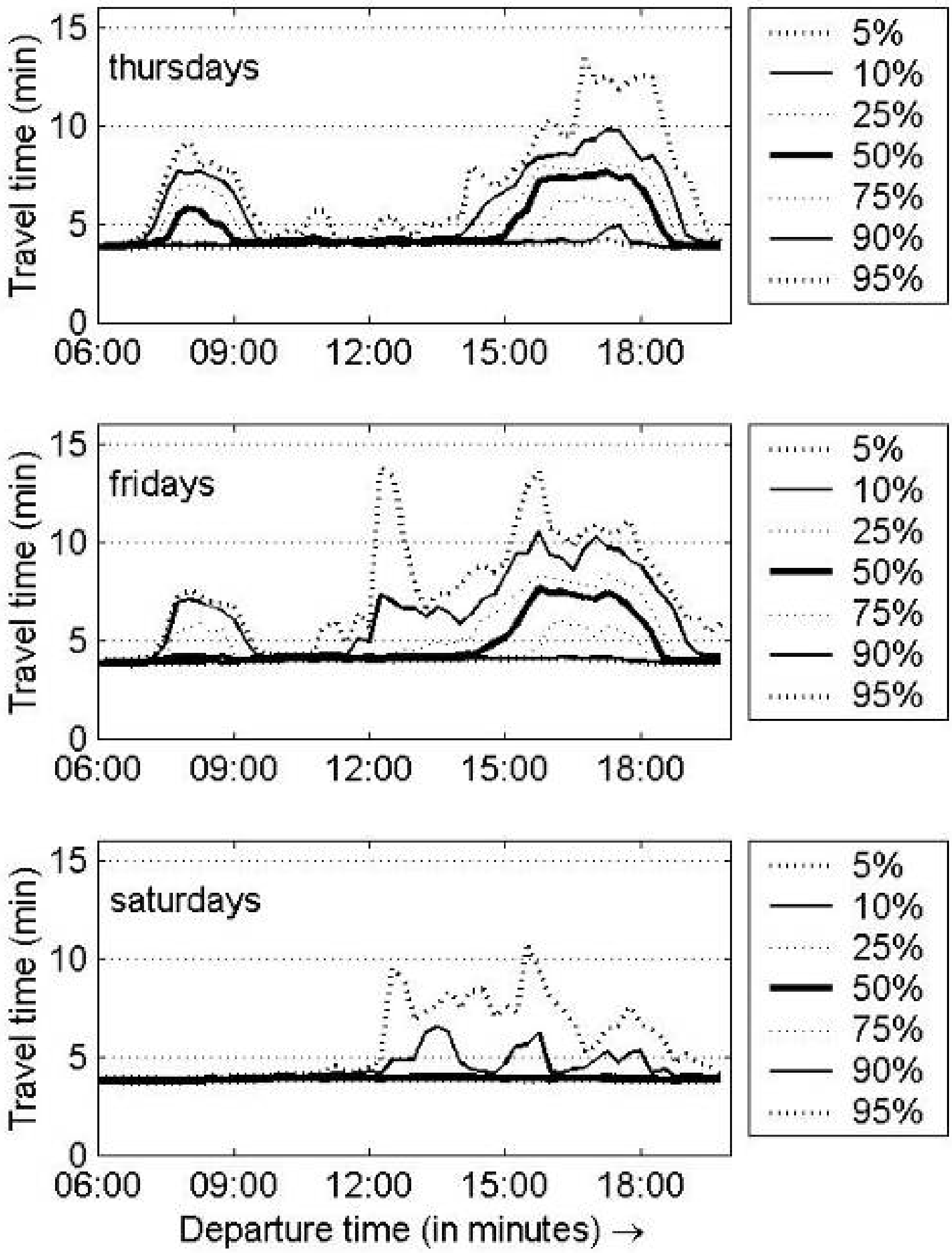

Figure 1: Variability of travel times in 2002 for 15 minute time-of-the-day (TOD) periods between 6AM and 8PM on the northern Rotterdam beltway A20 on Thursdays, Fridays and Saturdays

On Thursdays (which may be classified a typical weekday travel time pattern), a morning and afternoon peak are clearly identifiable, while on Fridays in more than half the cases no morning peak occurs. On Fridays the afternoon peak starts earlier and lasts longer than on other weekdays. On Saturdays, in 75% of the cases no congestion occurred, albeit that there were some serious exceptions. In 5% of the cases travel times were three times higher than the normal (no congestion) case of 4 minutes.

As argued in (Bates, 2001) the 90th percentile is a robust statistic representing the upper bound of travel times occurring for a particular TOD/DOW. We will therefore use it in the reliability metrics we propose below. From Figure 1 it becomes clear that the travel time distribution for many DOW/TOD values are skewed to the left, reflecting the fact that in still a considerable numbers of cases extreme travel times occurred. As such, the skewedness of the travel time distribution for a given DOW/TOD is an indicator for the travel time reliability. The more skewed the travel time distribution, the less reliable the travel times. In (Lam and Small, 2001) the distance between the 90th and 50th percentile is proposed as indicator of reliability, we propose for skewedness the ratio of the distance between the 90th and 50th percentile and the distance between the 50th and 10th percentile:

$$\lambda^{skew} = \frac{T90 - T50}{T50 - T10} \tag{1}$$

In which TXX denotes the XX percentile value. In general, the larger λ^{skew} the higher the probability of extreme travel times (relative to the median) that may occur. Since metric is a ratio it can be interpreted and applied regardless of the absolute magnitude of travel times.

The width of the travel time distribution is also an indicator for travel time reliability. The wider the distribution is (relative to the median), the higher the uncertainty and the lower travel time reliability. We propose the following metric

$$\lambda^{var} = \frac{T90 - T10}{T50} \tag{2}$$

which is the ratio of the range of travel times in which 80% of the observations around the median fall into, and the median travel time. Large values indicate the width of the travel time distribution is large relative to its median value.

Figure 2 shows scatter plots of both λ^{var} (top) and λ^{skew} (bottom) against median travel times (in seconds!). From the top graph we see λ^{var} steeply increases with median travel time until some threshold value of median travel time (in this case 0.2) is reached, after which this metric remains constant or even slightly decrease again. Put simply, for higher median travel time, the travel time distribution is also wider and hence our uncertainty is larger. The metric for skew (λ^{skew}) also steeply increases with median travel time but decreases after reaching a maximum at the (approximately) same threshold value as above. In the intermediate conditions between free-flowing and congested traffic, λ^{skew} is up to ten times larger then in free-flow or congested conditions.

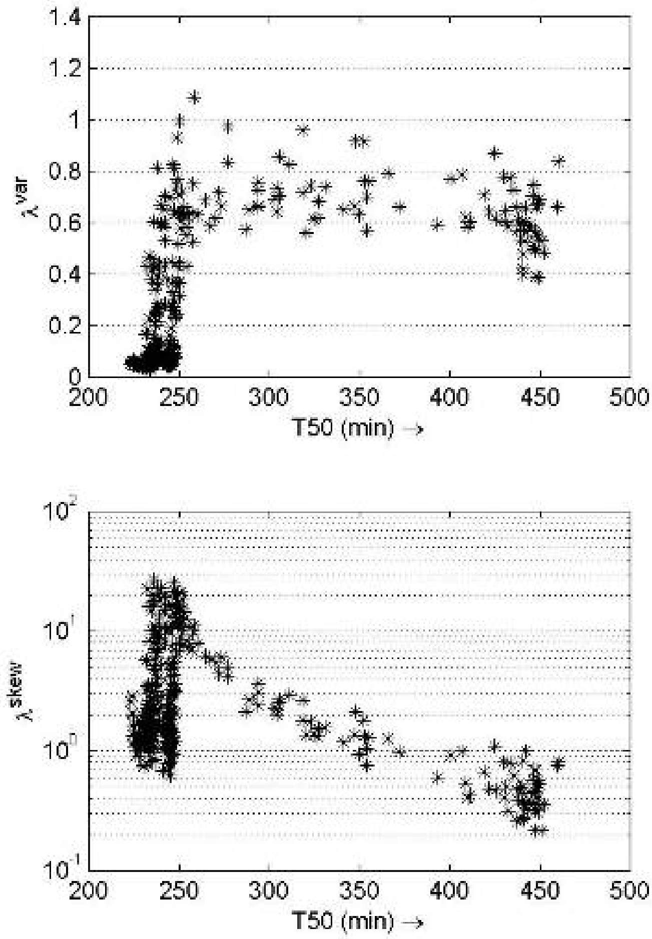

Figure 2: Scatter plots of two measures of reliability (spread and skew) as a function of median travel time.

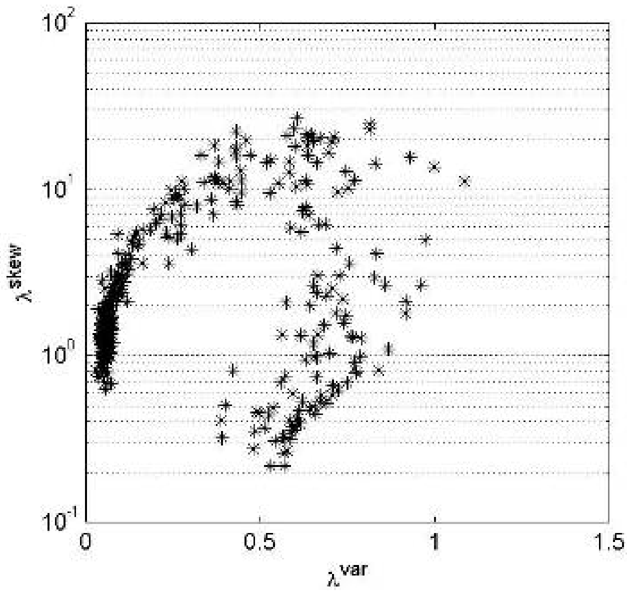

Figure 3: Scatter plot of skew against spread of the travel time distribution.

More counter-intuitively is the relation between λ^{skew} against λ^{var}. Figure 3 (which shows a scatter plot of λ^{skew} against λ^{var}) illustrates how λ^{skew} steeply (near exponentially) increases with λ^{var}. As traffic becomes more severely congested ($\lambda^{var} > 0.2$) the increase in skew slows down, and reaches a maximum (in this case ± 11) between $0.4 \leq \lambda^{var} \leq 0.5$. For $\lambda^{var} > 0.4$ we see two branches of the figure. One follows the previous trend and remains constant or slightly decreases again. The second starts with low values of λ^{skew} (< 0.2) for $\lambda^{var} \approx 0.4$ and increases for $\lambda^{var} > 0.4$.

The explanations to these phenomena are straightforward. In periods just before (after) the regular peak periods, most travel time observations are low, that is near free-flow travel time, while in a few cases congestion has already set in (still persisted) and high travel times occurred. This results in (strongly) left skewed day-to-day distributions[2]. Figure 4 schematically shows how this translates to the results in Figure 3. In case of queue build up / dissolving the travel time distribution is strongly skewed to the left, while in case of congestion the travel time distribution is skewed to the right (the right branch in Figure 3), reflecting the fact that in most of those time periods high travel times occurred, with a few (extreme) exceptions. In free flow conditions the travel time distribution is more or less symmetric (far left in Figure 3). To a degree, the skew in the transient conditions may also be influenced by the (arbitrary) choice of the TOD period size (in our case 15 minutes). Different binning of the data may lead to a slightly different pattern, we argue, however, the general pattern stays the same.

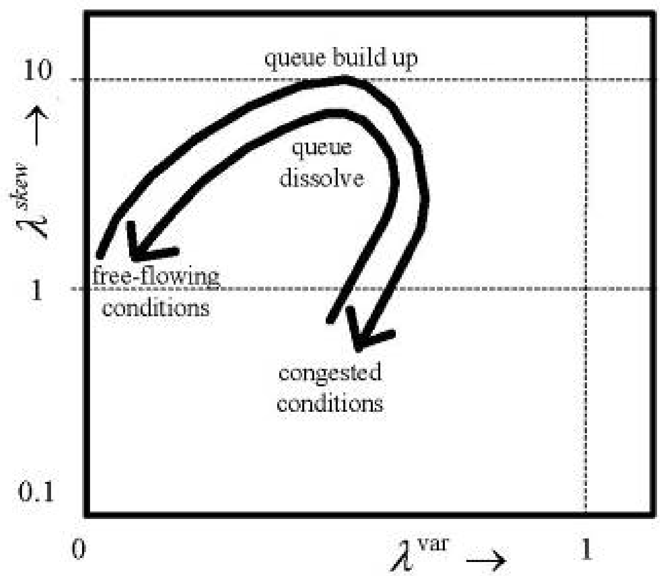

Figure 4: Schematic representation of the relation between skew and spread of the travel time distribution

The conclusion is that three characteristic percentiles provide us with two key indicators for the width and shape of the travel time distribution. The first, λ^{var} globally indicates whether in a certain DOW/TOD period congestion occurs, while the second, λ^{skew}, indicates congestion build

---

2 Similarly, in *within* peak hour time periods most travel time observations are high, except on a few days during which no congestion occurred, yielding strongly right-skewed day-to-day distributions.

up and dissolve. More particularly, high λ^{skew} values indicate whether or not the start or end of the peak hour might occur in a particular TOD, which provides for a means to quantify the uncertainty in the duration of a peak hour on a particular DOW. In all cases, high values for either indicator imply travel times can be classified more unreliable than for low values.

3. Long term travel time prediction model

In this section we derive a simple model that predicts the three percentiles and inherently the reliability metrics proposed above. We may expect that a percentile value can be considered a function of time of day (TOD), day of the week (DOW) and the month of the year (MOY). The latter reflects seasonal differences (e.g. effect of spring and summer holidays, but also climate influences). The general form of this model reads

$$\mathbf{y} = G(TOD, DOW, MOY, \Omega) \tag{3}$$

in which **Ω** denotes a vector of parameters to be estimated from our database. As a preliminary choice we use a simple feed-forward artificial neural network (ANN)[3], which uses as inputs TOD and DOW *only*. We leave out the seasonal component, since we have too few TOD/DOW/MOY measurements (per tupelo only 4 or 5 values) to create statistically sound percentiles. The neural network model has a sigmoid output layer of size three producing the output vector **y** shown in eq. (4).

$$\mathbf{y} = \begin{bmatrix} T50 - T10 \\ T50 \\ T90 - T50 \end{bmatrix} \tag{4}$$

The reason for this particular choice is that this ensures $T10 \leq T50 \leq T90$ (we choose an output transfer function which is always positive). Note that the three percentiles can be easily obtained with this output vector. For the hidden layer (size H) we choose a hyperbolic tangent transfer function. The number of inputs U to that layer equals two (TOD and DOW). Mathematically, this ANN model can be written as

$$\begin{aligned} y_k &= 1/\left(1+e^{-z_k}\right), k = \{1,2,3\} \\ z_k &= -\left[b_{k0} + \sum_{h=1}^{H} b_{kh} \tanh\left(w_{h0} + \sum_{j=1}^{U} w_{hj} u_j\right)\right] \end{aligned} \tag{5}$$

in which y_k denotes one of the three model outputs, u_j denotes one of the U inputs, and $\boldsymbol{b}$ and $\boldsymbol{w}$ the vectors of adjustable parameters of the output and hidden layer respectively. Note both hidden and output layer have a bias weight. We use the Levenberg-Marquardt / Bayesian Regularization (LMBR) algorithm described in (Foresee and Hagan, 1997) to train this model. As shown in (MacKay, 1995) and (Lint, 2004b) the Bayesian statistics behind this algorithm provide a superior and mathematically sound way to prevent ANN over-fitting and, as a bonus, an analytical and unified way of calculating confidence intervals around the predictions. We can therefore set H initially to the arbitrary size of 15. All input and output data is linearly scaled to the interval [0.1,0.9] for better and more stable learning (Hagan and Menhaj, 1994).

---

3 A comprehensive introduction in the field of ANN's can be found in Bishop, C. M. (1995) *Neural Networks for Pattern Recognition,* Oxford University Press, United Kingdom.

4. Results

We have trained the ANN on the same dataset as described earlier (TOD/DOW percentiles for 2002). From the training it appeared that over 95% of the parameters were dubbed "effective" (MacKay, 1995), which is to be expected in such a relatively small model with only two independent variables (TOD and DOW). It also indicates given the data we used and the choice of input and output our preliminary choice of $H=15$ proved just about sufficient. Below we test these results against TOD/DOW percentiles calculated for the whole year of 2001. Figure 5 shows the results of the ANN model of eq. (5) for the same three days as were shown in Figure 1, that is on Thursdays, Fridays and Saturdays.

In the figure, it is clear the ANN provides a smooth approximation of these three characteristic percentile time series. We may conclude the model has certainly learned the key differences between these travel time profiles on different days.

- On Thursdays, in more than 50% of the cases both a morning and afternoon peak occurs, the first between 8:00 and 9:00 AM, the latter between 15:00 and 18:15 PM
- The Friday afternoon peak starts earlier and lasts longer than on other weekdays
- On Saturdays, most of the time no congestion occurs, during the whole afternoon minor congestion may occur in less than half the occasions.

As a more quantitative indication of the performance we finally note that on the test data (2001), the following overall performance was obtained. In the ensuing N denotes the total number of TOD/DOW values[4]. In terms of mean relative error we then find

$$MRE = \frac{100}{N} \sum_{\forall TOD,DOW} \sum_{k=1}^{3} \frac{y_k(TOD,DOW) - T_k(TOD,DOW)}{T_k(TOD,DOW)} = -0.1\%$$

and in terms of relative standard deviation

$$SRE = \frac{100}{N-1} \sum_{\forall TOD,DOW} \sum_{k=1}^{3} \left(\frac{y_k(TOD,DOW) - T_k(TOD,DOW)}{T_k(TOD,DOW)} - \frac{MRE}{100} \right)^2 = 29\%$$

In both cases T (TOD,DOW) is a vector (T10, T50, T90). On the average, the model is almost unbiased, albeit there is a considerable residual error in its predictions, which is largely due to the smooth approximation the model makes.

5. Conclusions

The conclusions of this study can be roughly subdivided in two parts. First, following the generally accepted notion that median and percentile values are more robust statistics than mean and variance for quantifying travel time reliability, we derived two characteristic metrics based on the median, 10th and 90th percentile of the day-to-day travel time distribution.

- λ^{var} provides us with the relative width of the travel time distribution (with respect to the median) in a certain DOW/TOD period. Large values indicate larger uncertainty and hence unreliability.
- λ^{skew} depicts the skewedness of the distribution. In DOW/TOD periods where congestion sets in or dissolves, this metric is large (10-15), while in severely congested conditions the metric very small (<<1). In free flow conditions λ^{skew} is approximately one. Large values depict high uncertainty and hence unreliability.

4 We used daily profiles from 6:00Am-8:00PM, with 15 minute time periods, yielding N = 56*7=392 (TOD/DOW combinations)

The metrics can be used to identify not only the unreliability of travel times for a given DOW/TOD period, but also identify DOW/TOD periods in which it is likely that congestion sets in (or dissolves. Practically, this means identifying the uncertainty of start, end and hence length of morning and afternoon peak hours.

Secondly, we proposed a simple neural network based model to predict the three percentiles and hence the two metrics. On a separate test dataset, the model proved almost unbiased results, but produced considerable variance. We argue, however, the predominant reason is that this (preliminary) model makes a very smooth prediction, which is on the average correct, but results in considerable variance due to the more pronounced variability in the target signal. Even in its current form, the model can serve as a long-term travel time prediction tool, but also as a travel time unreliability prediction tool, using the two metrics presented above on the model outcomes.

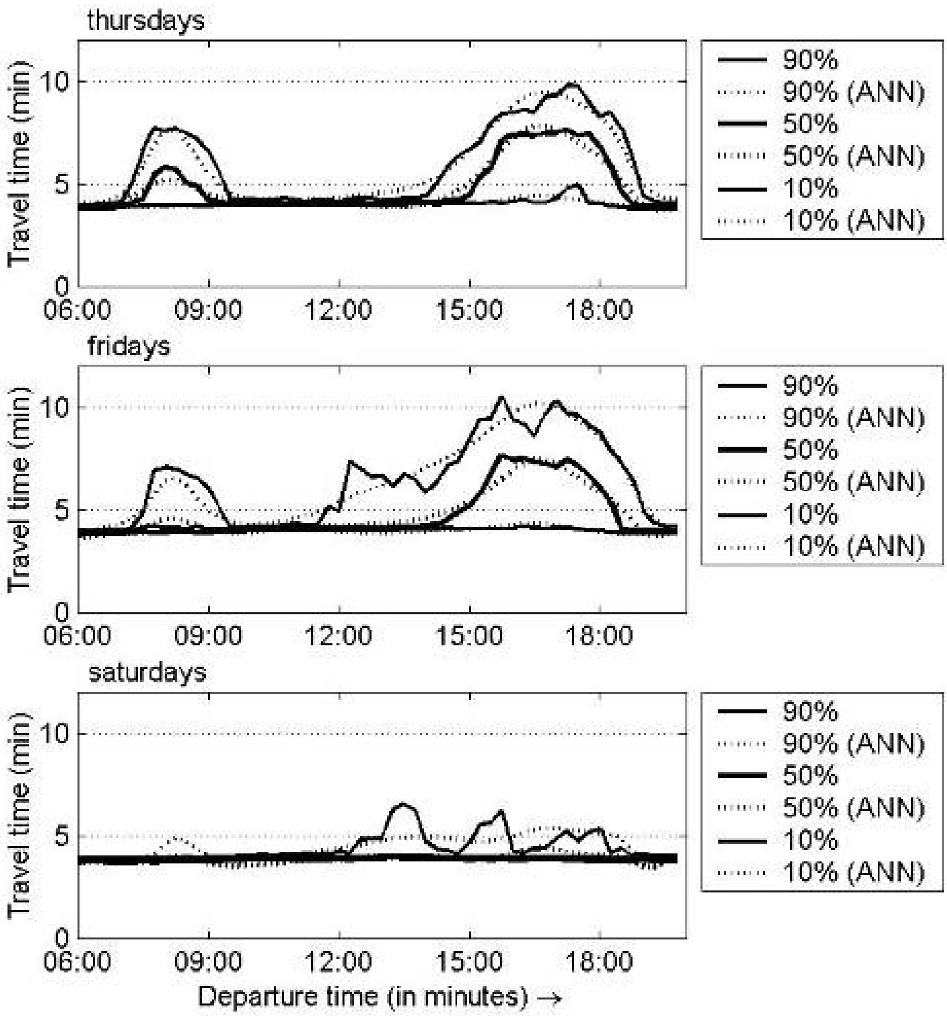

Figure 5: example performance of long term ANN travel time prediction model

6. Further research directions

The validity, robustness and most of all usefulness of the reliability metrics should be further investigated in various application domains, ranging for discrete choice modeling (route and departure time), to real time Advance Travel Information Systems (ATIS) such as web-based or in-car route navigation tools.

Furthermore, the artificial neural network model we developed here is still simple, capturing time-of-day and day-of-week trends only. The model should be augmented with seasonal trends (e.g. month of the year or quarter), and could also include for example external factors such as weather and large-scale road works. For this a large-scale database containing several years of (estimated) travel times and associated external factors on a number of different freeway routes is needed.

Finally, the travel time distribution in a particular TOD/DOW (MOY/…) period could be augmented by the within period (here of magnitude 15 minutes) distribution of travel times. This within period distribution reflects differences in individual driver behavior and more importantly, the dynamics of freeway traffic. Following the same reasoning as in (Lint, 2004a), we could assume that within period variance is independent of day-to-day travel time variance. This is a reasonable assumption, since we may expect for example congestion build up on day X be unrelated (physically!) to congestion build up on day Y. Assuming statistical independence then allows one to super impose day-to-day with within period variance.

Acknowledgment

This research is part of the Research project "Regiolab-Delft" (Van Zuylen and Muller, 2002) and is sponsored by the Dutch Ministry of Public Works, Water Management and Infrastructure.

References

Bates, J., Black, I., Fearon, J. and Porter, S. 2002. Supply Models for Use in Modeling the Variability of Journey Times on the Highway Network, In Proceedings of the European Transport Conference Proceedings, CD-Rom, Hamerton College, Cambridge, England

Bates, J. J. 2001. Reliability: The Missing Model Variable, In Travel Behaviour Research: The Leading Edge (Ed, Hensher, D.), Elsevier Science, Oxford, UK, pp. 527-546.

Bell, M. G. H. and Cassir, C. 2000. Reliability of Transport Networks, Research Studies Press, London, UK.

Bishop, C. M. 1995. Neural Networks for Pattern Recognition, Oxford University Press, United Kingdom.

Cassir, C., Yang, H., Ka Kan, L. O., Tang, W. H. and Bell, M. G. H. 2001. Travel time versus capacity reliability of a road network (Chapter 9), In Reliability of transport networks, Vol. 119-138.

Foresee, F. D. and Hagan, M. T. 1997. Gauss-Newton Approximation to Bayesian Learning, In International Conference on Neural Networks, Vol. 3, pp. 1930-1935

Hagan, M. T. and Menhaj, M. B. 1994. Training Feed-Forward Networks with the Marquardt Algorithm, IEEE transactions on Neural Networks, 5, 989-993.

Lam, T. C. and Small, K. A. 2001. The value of time and reliability: measurement from a value pricing experiment., Transportation Research Part E: Logistics and Transportation Review, 37, 231-251.

Lint, J. W. C. v. 2004a. Quantifying Uncertainty in Real-Time Neural Network based Freeway Travel Prediction, In 83rd Transportation Research Board Annual Meeting, National Academies Press, Washington D.C., USA

Lint, J. W. C. v. 2004b. Reliable Travel Time Prediction on Freeways, TRAIL thesis series T2004/3, Delft: TRAIL Research School

Lint, J. W. C. v. and Zijpp, N. J. v. d. 2003. An Improved Travel-time Estimation Algorithm using Dual Loop Detectors, In Transportation Research Board Annual Meeting, CD Rom, National Academies Press, Washington D.C, USA

Lo, H. K. 2002. Trip travel time reliability in degradable transport networks, In Proceedings of the 15th International Symposium On Transportation And Traffic Theory (ISTTT), Adelaide, South Australia, pp. 541-560

MacKay, D. J. C. 1995. Probable Networks and Plausible Predictions: A Review of Practical Bayesian Methods for Supervised Neural Networks, Network: Computation in Neural Systems, 6, 469-505.

Van Zuylen, H. J. and Muller, T. H. J. 2002. Regiolab Delft, In Proceedings of the 9th World Congress on Intelligent Transport Systems, CD-Rom, Chicago, Illinois, USA.