Data, modelling and inference in road traffic networks

BY RICHARD. J. GIBBENS1,* AND YUNUS SAATCI2

1Computer Laboratory, University of Cambridge, William Gates Building, 15 JJ Thomson Avenue, Cambridge CB3 0FD, UK
2Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK

In this paper, we study UK road traffic data and explore a range of modelling and inference questions that arise from them. For example, loop detectors on the M25 motorway record speed and flow measurements at regularly spaced locations as well as the entry and exit lanes of junctions. An exploratory study of these data helps us to better understand and quantify the nature of congestion on the road network. From a traveller’s perspective it is crucially important to understand the overall journey times and we look at methods to improve our ability to predict journey times given access jointly to both real-time and historical loop detector data. Throughout this paper we will comment on related work derived from US freeway data.

Keywords: road networks; exploratory data analysis and statistical models

1. Introduction

Road traffic data are increasingly becoming available for UK road networks in suitable forms and sufficient quantities that create interesting and evolving challenges for modellers.

In this paper, we look at the use of just one form of data now routinely gathered from loop detectors located on the UK’s strategic road network. The data are collected by the MIDAS system operated by the Highways Agency (2005).

The paper begins with some exploratory data analysis for loop detector data gathered on the southwest quadrant of the M25 London orbital motorway. We consider a basic speed–flow relationship and illustrate how this aids our understanding of the nature of congestion. We also look at some performance metrics derived from the MIDAS data, which help quantify the magnitude of delays experienced by motorists on this heavily congested section of motorway.

The paper continues in §3 with a consideration of how the speeds translate into journey times for motorists and in particular we consider the variability in these times. This leads on to a study of methodologies for journey time prediction.

* Author for correspondence (richard.gibbens@cl.cam.ac.uk).

One contribution of 16 to a Discussion Meeting Issue ‘Networks: modelling and control’.
Journey time prediction using sources of real-time measurement data has the potential to assist travellers through the provision of more accurate estimates of journey times. Improving the accuracy of the prediction by suitable methods that make use of real-time data helps to reduce the overall uncertainty of journey times. Rice & van Zwet (2004) describe a simple-to-implement prediction methodology and report successful results with US data in comparison with the more sophisticated and harder-to-implement methods. In this work, we have examined in detail the performance of these methodologies when used with real-time UK MIDAS loop detector data. A preliminary account of this investigation is given in Gibbens & Werft (2005) and Werft (2005). A fuller account of some of these issues has been reported in Gibbens & Saacti (2006).

Section 3 describes the basic model and defines the prediction methodologies considered. Section 4 presents the results of our numerical investigations into journey times and the comparison between the methodologies.

The studies reported here have been motivated by related studies with US loop detector data, especially with data collected on the California freeways available through the Freeway Performance Measurement System (PeMS; http://pems.eecs.berkeley.edu/). A recent survey of this work is included in Varaiya (2008). Earlier work of particular relevance is in Chen et al. (2001a,b).

2. Exploratory data analysis using MIDAS loop detector data

Figure 1a shows the pattern of speed and flow measurements during the morning period (05.00–11.00) on Wednesday 14 July 2004 recorded at a single location between junctions 11 and 12 on the clockwise carriageway of the M25 motorway. The speed \( v(t) \) is the average speed of vehicles across all clockwise lanes during the time-interval \([t, t+1]\) minutes. We use miles per hour (mph) as the units of speed. The corresponding flow \( q(t) \) is the total number of vehicles passing the given location across all clockwise lanes during the same 1 minute interval. Figure 1 shows a free-flow regime where initially average speeds are around the level 65–70 mph and remain so as flow begins to build up. Once flow has built up sufficiently at a time between 06.00 and 07.00 the speeds are seen to collapse to lower levels. During this congested regime that ends at approximately 11.00, speeds and flows vary significantly minute by minute.

Let us define the \emph{density} \( \rho(t) \) of vehicles per mile per lane by

\[
\rho(t) = \frac{q(t)}{nv(t)}, \tag{2.1}
\]

where \( n \) is the number of lanes (\( n=4 \) in this example). Figure 1a describes a free-flow regime when the density of vehicles is sufficiently low and then once the density exceeds a critical value the congested regime takes hold.

These ideas are further illustrated by the fitting of simple models. A model suggested in the traffic literature (see May 1990; Bellemans 2003) is

\[
v(t) = v_{\text{free}} \left(1 - \left(\frac{\rho(t)}{\rho_{\text{jam}}}\right)^{\alpha}\right)^{\beta} . \tag{2.2}
\]

Phil. Trans. R. Soc. A (2008)
Here $v_{\text{free}}$ is the limiting speed as the density drops to zero. The quantity $\rho_{\text{jam}}$ is the density when the speed finally drops to zero and the road is totally jammed in a gridlocked state. The parameter values used in figure 1b were $v_{\text{free}} = 71.5$ mph and $\rho_{\text{jam}} = 218.9$ vehicles per mile per lane. A nonlinear least-squares fit to the data produced the estimated parameter values of $\alpha = 1.71$ and $\beta = 10.3$. The fitted expression (2.2) can now be used for the relationship of flow with density using expression (2.1). The maximum flow occurs at a critical value of density $\rho_{\text{crit}}$ and gives a natural measure to the road’s capacity for flow. The value of the critical density $\rho_{\text{crit}}$ is given analytically by

$$\rho_{\text{crit}} = \rho_{\text{jam}} \left( \frac{1}{1 + \alpha \beta} \right)^{1/\alpha}.$$

The estimated value of $\rho_{\text{crit}}$ is 9.5 vehicles per mile per lane and the estimated capacity is 6385.1 vehicles per hour.

Figure 1 has examined the relationship between speeds and flows at a single location and on a single day. We now extend the scope to a region of road covered by many loop detector sites and over many days within a single year. We also look at alternative performance metrics to the speed and flow.

Figure 2 describes loop detector data taken from the M25 clockwise between junctions 9 and 14 on weekdays. A total of 32 loop detector sites were used and data were recorded for 247 weekdays in the year 2003 during the 7 hour morning
Figure 2. (a,b) Daily measurements of vehicle miles travelled (VMT), vehicle hours travelled (VHT) and vehicle hours delay (VHD) taken over 247 weekdays in 2003 and from 32 loop detector sites situated over a stretch of M25 clockwise between junctions 9 and 14 of length approximately 10 miles. (a) The VHD with the VMT. The fitted curve shows the rapidly increasing nature of the delay with additional VMT. The median VHD is 2229 and the median VMT is 351 469. (b) The delay with the VHT, with median 7446 hours. The median delay is nearly one-third of the median hours travelled. (c) Variation between the VHT (black solid line, observed; grey solid line, ideal) and the VMT (dashed line, observed) over the day (Monday, 6 January 2003) on a section of the clockwise carriageway of the M25. The VMT increases rapidly from 05.00 to 06.30 when it peaks at around 70 000 VMT per hour (shown on the right-hand scale). After 06.30, the VMT declines at a slower rate. The VMT increases before levelling off at approximately 07.00 and remains at a high level for several hours before declining. The VHT by the same vehicles at a reference speed of 67 mph would peak well below the observed level and remain at a much lower level throughout the morning busy period.
period from 05.00 to noon. Each loop detector site is located within a cell of length \( h = 500 \), \( m \approx 0.311 \) miles. The total vehicle miles travelled (VMT) is the aggregate over time-intervals \( t \) and loop detector sites of the product \( q(t) \times h \). The vehicle hours travelled (VHT) is given by the ratio of VMT to the average speed, that is \( \text{VHT} = \frac{\text{VMT}}{v(t)} \). The delay caused by congestion can be assessed by the difference between the VHT and that which would arise from the same VMT at a reference speed, here taken to be 67 mph (a value given by Target 1 in the Public Service Agreement produced by the Department for Transport 2004). Thus, the vehicle hours delay (VHD) is given by

\[
\text{VHD} = \text{VHT} - \frac{\text{VMT}}{67}. \tag{2.4}
\]

Figure 2a shows the VHD against VMT for the 247 days included in this study. Delay increases rapidly with the VMT with the median VHD 2229 vehicle hours per day and the median VMT some 351,469 vehicle miles. Figure 2b shows the VHD against the VHT. The median VHT is 7446 vehicle hours. Thus, the median delay is nearly one-third of the median hours travelled each day. While many models could be fitted to these data for the VHD, we have just shown in figure 2a that a particularly simple one with a nonlinear least-squares fit to the expression \( \text{VMT} / (C - \text{VMT}) \). The estimated value of \( C \) for these data was approximately 511,000 vehicle miles. As indicated by the model, a small growth in the VMT demanded could be expected to translate into substantial increases in delay.

Figure 2c looks at the daily profile of these performance metrics. The left-hand scale records the VHT per hour at 1 minute intervals as well as the difference between the VHT and the congestion delay given by expression (2.4). The right-hand scale shows the VMT per hour. We can see that the VMT per hour increases rapidly at approximately 06.00, peaking shortly thereafter. The VHT per hour also rises rapidly and remains at high levels for several hours before declining.

3. Journey time prediction methodologies

(a) Basic model and notation

The basic model and terminology are taken directly from Rice & van Zwet (2004) and are briefly summarized here as follows.

We suppose that there is a velocity field \( V(d, \ell, t) \) specifying the average speeds of vehicles for days \( d \in D \), at loop detectors \( \ell \in \{1, \ldots, L\} \) and for times (of day) \( t \in T \). There may be many days \( d \) and journeys are traversed from loop 1 to loop \( L \). The time of day epochs \( t \) are taken as every minute in the case of MIDAS data.

We define \( T(d, t) \) for the time of travel from loop 1 to loop \( L \) starting at time \( t \) on day \( d \). \( T(d, t) \) can be determined (approximately) from the velocity field on any day \( d \) in the past.

We also define a frozen-field travel time \( T^*(d, t) \) given by

\[
T^*(d, t) = \sum_{\ell=1}^{L-1} \frac{2d_\ell}{V(d, \ell, t) + V(d, \ell + 1, t)}, \tag{3.1}
\]
where \(d_l\) is the distance between loops \(l\) and \(l+1\). This quantity will play a pivotal role in the prediction methodologies. Note that it may be very easily determined with simple arithmetic operations from speed measurements as part of an online algorithm for journey time prediction.

The historical average travel time \(\overline{T}(t)\) for a journey starting at time of day \(t\) is given by

\[
\overline{T}(t) = \frac{1}{|D|} \sum_{d \in D} T(d, t),
\]

where \(|D|\) is the number of days in the set \(D\).

The task of a journey time prediction method is to estimate \(T(d, t + \delta)\) for time lag \(\delta \geq 0\) given only information known at time \(t\) on day \(d\). Time \(t\) is the decision time for estimating a journey beginning after a lag of \(\delta\) at time \(t + \delta\).

Two naive estimates of the journey time \(T(d, t + \delta)\) are as follows:

(i) \(T^*(d, t)\), the frozen-field estimator evaluated at the decision time, \(t\), and

(ii) \(\overline{T}(t + \delta)\), the historical mean estimator for journeys starting at time of day \(t + \delta\).

The frozen-field estimator \(T^*(d, t)\) assumes, therefore, that speeds remain held permanently fixed at their time \(t\) values throughout the journey. We would expect that this estimator would behave best at small values of \(\delta\), where it is able to capture from the real-time measurements known up to time \(t\) specific features of the traffic profile on day \(d\). As \(\delta\) increases, these (frozen) features become less relevant compared with the information captured by the long-run historical average estimator \(\overline{T}(t + \delta)\).

(b) Linear regression method using varying coefficients

Rice & van Zwet (2004) observed in US loop detector data a strong linear relationship between the frozen-field estimator \(T^*(d, t)\) and the exact observed journey time \(T(d, t + \delta)\) of the form

\[
T(d, t + \delta) = \alpha(t, \delta) + \beta(t, \delta) T^*(d, t) + \epsilon,
\]

where \(\epsilon\) is a mean zero random variable and the coefficients \(\alpha(t, \delta)\) and \(\beta(t, \delta)\) vary with both the decision time \(t\) and the lag before the journey begins \(\delta\). Further details of such varying coefficients models are given by Hastie & Tibshirani (1993). The parameters of such a linear model may be fitted through a weighted least-squares procedure which minimizes

\[
\sum_{d \in D, s \in T} [T(d, s) - \alpha(t, \delta) - \beta(t, \delta) T^*(d, t)]^2 K(t + \delta + s),
\]

where \(K(x) = (1/\sqrt{2\pi}) e^{-x^2/2\sigma^2}\) is the Gaussian density with mean zero and variance \(\sigma^2\).

The purpose of the Gaussian density \(K(\cdot)\) is to produce smoothed estimates of the regression coefficients \(\widehat{\alpha}(t, \delta)\) and \(\widehat{\beta}(t, \delta)\) as both the decision time \(t\) and the lag \(\delta\) vary. The degree of smoothing is adjusted by the choice of the variance parameter \(\sigma\). This methodology then yields a regression-based journey time estimator \(\overline{T}(d, t + \delta)\) given by
\[ \hat{T}(d, t + \delta) = \hat{\alpha}(t, \delta) + \hat{\beta}(t, \delta) T^*(d, t). \] (3.5)

Observe that putting \( \alpha(t, \delta) = \alpha'(t, \delta) T'(t + \delta) \) shows that the estimator \( \hat{T}(d, t + \delta) \) is, in fact, a particular data-dependent linear combination of the two naive estimators.

(c) Nearest neighbour methods

An alternative family of prediction techniques is given by the nearest neighbour method. In the simplest form of the nearest neighbour method the estimator of journey time \( T(d, t + \delta) \) is given by first finding the previous day \( d' \) which most closely matches the observed speeds up to time \( t \) on day \( d \), according to some well-defined distance measure. Hence, if day \( d' \) minimizes the distance to \( d \) among all previous days then the nearest neighbour estimator \( T_{\text{NN}}(d, t + \delta) \) is given by

\[ T_{\text{NN}}(d, t + \delta) = T(d', t + \delta). \] (3.6)

Rice & van Zwet\(^1\) offer several options for the distance \( m(d_1, d_2) \) between two days \( d_1 \) and \( d_2 \). Two such options considered for evaluation are given as follows:

\[ m_1(d_1, d_2) = \sqrt{\sum_{t-w \leq s \leq t} [T^*(d_1, s) - T^*(d_2, s)]^2} \] (3.7)

and

\[ m_2(d_1, d_2) = \sum_{\ell \in L, t-w \leq s \leq t} |V(d_1, \ell, s) - V(d_2, \ell, s)|, \] (3.8)

where \( w \) is a window size parameter.

The nearest neighbour method can be readily extended to the \( k \)-nearest neighbour (\( k \)-NN) method. First, the \( k \)-closest days \( d_1, d_2, \ldots, d_k \) are found. Then, the predictors derived from each similar day are combined in a weighted averaging scheme, where the weights are inversely proportional to the distance of each day to the present day \( d \). The predictor for the \( k \)-NN method \( T_{\text{NN}}^k(d, t + \delta) \) is hence given by

\[ T_{\text{NN}}^k(d, t + \delta) = \sum_{i=1}^{k} w_i T(d_i, t + \delta), \] (3.9)

where \( w_i \propto m(d, d_i)^{-1} \) and the distance function is \( m(d_1, d_2) \). Thus, the simplest nearest neighbour method corresponds to the \( k \)-NN method with \( k = 1 \).

We notice that determining the estimator \( T_{\text{NN}}^k \) involves evaluating a distance for each day according to the distance function as well as ranking those distances to find the \( k \)-closest days.

\(^1\)Rice & van Zwet also consider a third class of estimators based on a principal components procedure. We have not considered such estimators here as Rice & van Zwet did not find them to improve over the regression or nearest neighbour estimators.

*Phil. Trans. R. Soc. A* (2008)
4. Numerical results

(a) The MIDAS dataset

The data considered in this report consist of speed measurements collected per minute from 63 MIDAS loop detector sites located on lane 2 (where the slow lane is numbered 1) of the clockwise carriageway between junctions 9 and 14 on the M25 London orbital motorway. The spacing between the loops $d_l$ is taken as 500 m. The data considered ranged from 05.00 to 20.00 (i.e. 900 1 minute intervals) on weekdays in 2003. Missing values reduced the original 261 weekdays down to 231 days.\(^2\) The split between days of the week was 39 Mondays, 142 midweek days (i.e. Tuesdays, Wednesdays and Thursdays) and 50 Fridays. The resulting data formed a velocity field $V(d, l, t)$ with dimensions $231 \times 63 \times 900$.

For comparison, the study by Rice & van Zwet (2004) included 34 days and 116 loop detectors along 48 miles of I-10 in Los Angeles.

Figure 3 shows a spatio-temporal plot of the speeds for a single day (Monday, 6 January 2003). During the period 06.30–10.00, and for much of the road under consideration, vehicles are travelling at relatively low speeds with a backward-propagating wave-like profile (see also our earlier discussion in §2). Horizontal stripes can be seen in the plot to roughly coincide with the bottlenecks being formed in the vicinity of junctions.

(b) Journey times

From the velocity field, a travel time $T(d, t)$ can be constructed for the journey from loop 1 to loop 63, which starts at time $t$ on day $d$. Figure 4a(i) shows how the journey times vary during the day for each of the individual 39 Mondays. Journey

\(^2\)Missing values within the MIDAS speed data that formed significant blocks over time and loops caused that day to be rejected. More commonly, missing values occurred throughout parts of the day at one or more non-adjacent sites. Less frequently, many sites produced missing values for just a single minute. In both of these cases, the missing values were imputed by straightforward linear interpolation.
times are naturally seen to increase during the morning busy period. (Several exceptions occur on Bank Holiday Mondays.) During the middle portion of the day and again between 17.00 and 19.00 there are significant numbers of days when journey times have increased. However, these increases are much less pronounced than they are in the morning. By contrast, in the dataset considered by Rice & van Zwet (2004) the maximum congestion is in the period from 15.00 onwards.

Figure 4 shows a strong day-of-week effect on journey times and we have used these three categories of weekdays (namely, Mondays, midweek days and Fridays) to separately estimate journey times.

The key linear relationship identified by Rice & van Zwet (2004) that underlies the prediction methodology is between the quantities $T^*(d, t)$ and $T(d, t + \delta)$. Figure 5 shows the scatter plots of these two quantities where the decision time $t$ is 08.00, the lag $\delta$ ranges from 0 to 120 min and the data are confined to just the 39 Mondays. Each plot also shows the historical mean estimator as a horizontal line. We notice how the slope of the regression line diminishes as the lag increases.

Equation (3.4) was used to fit the regression coefficients $\alpha(t, \delta)$ and $\beta(t, \delta)$ by a standard weighted least-squares procedure. The regression-based journey time estimator $\hat{T}(d, t)$ was then obtained from the fitted coefficients using

---

*Figure 4. (a)(i) Journey times on 39 Mondays during 2003 starting at times ranging from 05.00 to 20.00. (ii) The distribution of journey times by means of box-and-whiskers plots. Journey times are not only longer during the morning rush hour period but also more spread out. (b)(i,ii) Journey times on 142 midweek days (Tuesday, Wednesday and Thursday). Median journey times rise during the morning and evening rush hours and there are many outlier days with longer journey times. (c)(i, ii) Journey times on 50 Fridays. Median journey times rise significantly from midday onwards along with a very wide variation in journey times.*
The smoothness of the surfaces is controlled by the parameter $s$ that was taken as 10 minutes. (The choice of such parameters is discussed in Gibbens & Saacti (2006) where the sensitivities to changes in the parameters is also explored.)

An important consequence that would follow from the adoption of Gaussian errors in the statistical model for $T(d, t)$ in (3.3) is that the many powerful techniques and the tools of Gaussian models can then be applied. In particular, the same statistical model may also be used to construct a prediction interval (also shown in figure 5 by the outer pair of sloping lines). The prediction interval illustrated here gives a region that we expect, given the statistical model, to contain the exact journey time with a probability of 90%. The level of 90% is for illustration only. It could be either higher or lower corresponding to intervals that are wider or narrower, respectively.

Figure 5. The linear relationship between the frozen-field estimator $T^*(d, t)$ and the journey time $T(d, t + \delta)$. Here the decision time $t$ is fixed at 08.00 on Mondays and the lag $\delta$ increases from 0 to 120 min. Both the historical mean and the least-squares regression are shown. Solid lines, linear regression estimator $\hat{T}$; dashed lines, upper and lower 90% prediction interval; horizontal dot-dashed lines, historical mean estimator $\hat{T}$.

The smoothness of the surfaces is controlled by the parameter $\sigma$ that was taken as 10 minutes. (The choice of such parameters is discussed in Gibbens & Saacti (2006) where the sensitivities to changes in the parameters is also explored.)

An important consequence that would follow from the adoption of Gaussian errors in the statistical model for $\hat{T}$ in (3.3) is that the many powerful techniques and the tools of Gaussian models can then be applied. In particular, the same statistical model may also be used to construct a prediction interval (also shown in figure 5 by the outer pair of sloping lines). The prediction interval illustrated here gives a region that we expect, given the statistical model, to contain the exact journey time with a probability of 90%. The level of 90% is for illustration only. It could be either higher or lower corresponding to intervals that are wider or narrower, respectively.
It may be worth concluding this section by describing how the regression estimator would be implemented. Using historical data, such as those shown in figure 4, the regression model is fitted and the sloping lines in figure 5 are computed. This part of the calculation is done offline and the results are saved for use by the online part of the algorithm. At the decision time \( t \) the frozen-field estimator \( T^* \) is obtained from the current speed measurements (in our example journey, the estimator involves a simple calculation (given by equation (3.1)) using the speed values recorded by the 63 MIDAS loop detectors). The regression estimator \( \hat{T} \) and the prediction interval are then looked up from the saved results of the offline calculation. For example, consider a lag of \( \delta = 60 \) minutes as shown in the central panel of figure 5. If the online calculation of \( T^* \) yields a value of 30.00 min then the regression estimator is \( \hat{T} = 22.31 \) min and the 90% prediction interval is (15.39, 29.24). If the frozen-field estimator was instead a value of 60.00 then the regression estimator would be \( \hat{T} = 34.51 \) min and the 90% prediction interval would be (27.56, 41.45). The historical mean estimator \( \bar{T} \) is computed from historical measurements alone and, in both these cases, it is 28.08 min, independent of online measurements.

(c) Comparison of methodologies

Figure 6 shows how the root-mean-square prediction errors for the three estimators varies as \( t \) varies throughout the period between 05.00 and 20.00 and with the lag \( \delta \) increasing from 0 to 120 min. The historical mean estimator is not affected by the choice of lag \( \delta \) except that the curves shown shift leftwards by the amount \( \delta \). The regression-based estimator has the lowest root-mean-square prediction error. During the period 06.30–10.00 on Mondays the regression-based estimator has more than halved the error compared with the historical mean. Later in the day, when journey times are far less variable there is less benefit to be obtained from the regression approach compared with simply using the historical mean. As the lag \( \delta \) is allowed to increase the error in the regression-based estimator \( \hat{T} \) approaches that of the historical mean. Figure 6 also includes the nearest neighbour estimator \( T^{k\text{NN}} \) calculated with \( k = 4 \), a window size parameter of \( w = 20 \) min and the \( m_1(\cdot) \) distance function. The performance of the \( T^{k\text{NN}} \) estimator is quite similar to the regression estimator.

The middle and right panels in figure 6 show the prediction errors for the cases of midweek days and Friday, respectively. A similar comparison applies in these two categories. However, the prediction error with the historical mean estimator is rather greater in the case of Friday afternoons than that for Mondays. Hence, there is considerable scope for using real-time information to reduce the prediction error of journey times as can be seen with both the regression and the nearest neighbour estimators.

The findings shown in figure 6 taken together show that when the prediction error in the historical mean is high it is possible for the regression and nearest neighbour methods to substantially reduce the error, at least for short to medium lags. For longer lags, over 2 hours (say), all estimators will approach the performance of the historical mean.

It is quite surprising that despite investigating a wide choice of parameters (\( k \) and \( w \) for the nearest neighbour estimator and \( \sigma \) for the regression estimator) we were unable to observe any significant improvement of the nearest neighbour
procedure over the regression procedure. Of course, it may be that certain additional information concerning the presence of specific incidents on the road could be used to improve the nearest neighbour estimator. The regression procedure has rather minimal online requirements, as discussed above, compared with the nearest neighbour procedure which must compute an online search for the \( k \)-closest days.

5. Conclusions

In this paper, we describe our findings from using MIDAS loop detector data for journey time prediction. We have found that the simple-to-implement regression-based method of Rice & van Zwet (2004) works well in our example scenario of UK data taken from the M25 London orbital motorway in 2003.
This paper looked at the variability of journey times across days in three day categories: Mondays; midweek days; and Fridays. The regression-based estimator together with a \(k\)-nearest neighbour estimator was studied and the results compared in terms of the root-mean-square prediction error. It was found that where the variability was greatest (typically during the rush hour periods or periods of flow breakdown) the regression and nearest neighbour estimators reduced the prediction error substantially compared with a naive estimator constructed from the historical mean journey time. Only as the lag between the decision time and the journey start time increased to beyond approximately 2 hours did the potential to improve upon the historical mean estimator diminish. Thus, there is considerable scope for prediction methods combined with access to real-time data to improve the accuracy in journey time estimates. In doing so, they reduce the generalized cost of travel. The regression-based prediction estimator has a particularly low computational overhead, in contrast to the nearest neighbour estimator, which makes it entirely suitable for an online implementation.

Finally, the studies described here demonstrate both the value of preserving historical archives of transport-related datasets and provision of access to real-time measurements.

The authors acknowledge support and funding from the Department for Transport (Horizons research grant H05-217) and from the EPSRC (research grant GR/S86266/01). The authors are especially grateful to the Highways Agency for use of the MIDAS loop detector data. All views expressed within this paper are those of the authors.

References


Werft, W. 2005 Travel time prediction in road networks. MPhil in Statistical Science, Statistical Laboratory, University of Cambridge.