UNIVERSITY OF LEEDS

This is a repository copy of *Factors Affecting Travel Times on Urban Radial Routes.*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/2362/

Monograph:
May, A.D. and Montgomery, F.O. (1984) Factors Affecting Travel Times on Urban Radial Routes. Working Paper. Institute of Transport Studies, University of Leeds , Leeds, UK.

Working Paper 177

Reuse
See Attached

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

White Rose university consortium
Universities of Leeds, Sheffield & York

White Rose Research Online
http://eprints.whiterose.ac.uk/

ITS

Institute of Transport Studies
University of Leeds

This is an ITS Working Paper produced and published by the University of Leeds. ITS Working Papers are intended to provide information and encourage discussion on a topic in advance of formal publication. They represent only the views of the authors, and do not necessarily reflect the views or approval of the sponsors.

White Rose Repository URL for this paper:
http://eprints.whiterose.ac.uk/2362/

Published paper
May, A.D. and Montgomery, F.O. (1984) *Factors Affecting Travel Times on Urban Radial Routes.* Institute of Transport Studies, University of Leeds, Working Paper 177

White Rose Consortium ePrints Repository
eprints@whiterose.ac.uk

Working Paper 177
July 1984

FACTORS AFFECTING TRAVEL TIMES ON URBAN RADIAL ROUTES

by

A D May and F O Montgomery

ABSTRACT

A D May and F O Montgomery, Working Paper 177, Factors Affecting Travel Times on Urban Radial Routes.

This paper presents the results of a study of peak period travel times and potential causal factors conducted over a one year period on five radial routes. It describes the hypotheses tested and the survey design adopted and then presents results for inter-vehicle variation in travel time within 250 15 minute periods and for inter-period variation based on 1018 moving observer runs. Moving observer results are found to depend on flow over the previous half hour, and a modified time series analysis used to demonstrate this is outlined. The contriubtion of other explanatory variables is also presented.

1. INTRODUCTION

1.1 Purpose and structure of the paper

This paper is based on a study conducted in Leeds U.K., during 1982/3 with the objectives of developing a cost-effective survey strategy for travel time surveys on urban radial routes and of understanding the nature and causes of variations in travel time on such routes. In this paper we concentrate on the second of these objectives, having discussed some of the issues related to the first objective elsewhere (May and Montgomery, 1983).

The hypotheses which the study was designed to test are first described, followed by summaries of the data collected on travel time variations within short (15 minute) periods, and on the variations between such periods. The paper then concludes with analyses of the nature of these two types of variation and the causes of the latter type.

1.2 Limitations of the study

The data described here relate to in-vehicle travel times by car for inbound journeys in the morning peak period (07.15 - 09.45) on five radial roads in Leeds. The detailed findings may not be transferable to other places, but the general conclusions should be of wider relevance. The survey locations and the types of travel time data collected were largely determined by the study's first objective, as described elsewhere (May and Montgomery, 1983).

2. STUDY HYPOTHESES

2.1 Types of variability in travel time

The variation in travel times on a particular route can be thought of as having three components:

a) variation within small time periods;

b) variation between periods (within days);

c) variation between days.

2.2 Inter-vehicle variation

Provided that the period of measurement is sufficiently small, it is hypothesised that variation (a) above is caused by differences in characteristics of the cars; driving styles; traffic lanes chosen; and traffic conditions met. These variations are here termed inter-vehicle variations. Obviously as the measurement period becomes larger, the time-related effects, as described in the next section, come into play, and the choice of measurement period must be based on an assessment of the size of these effects relative to the inter-vehicle variation. In this study a period of 15 minutes was chosen, since it was the minimum period for which most travel time survey data were required, and seemed likely to exclude most of the time-related effects.

2.3 Inter-period variation

(b) and (c) above constitute the inter-period variation, and can be analysed by examining the variation between such periods of some measure of central tendency of travel times. The mean and the median may both be useful measures. We suggest that most of this inter-period variation will be explained by the following factors:

i) Effect of traffic volume

One would expect that, other things being equal, travel times would increase with increasing flow. Numerous past studies have shown, however, that the relationship is quite complex for single links. (Branston, 1976). The relationship along a route consisting of several links may be simpler (because high travel times on one link may be compensated for by shorter times on the next), or more complicated because the effect of flow in preceding time periods cannot be ignored.

ii) Effect of traffic composition

The proportion of buses and commercial vehicles in the traffic stream may affect car travel times because these vehicles tend to be slower moving, have less acceleration and are more difficult to overtake. We have excluded motorcycles and pedal cycles from consideration because, from observation, they do not appear to interfere with the flow of other traffic. This may not hold true in areas with very high cycle use.

iii) Effect of specific incidents

Incidents or unusual conditions, such as road accidents, burst water mains, roadworks, bus strikes, etc. obviously affect travel times and must be recorded if they occur. However the data collected under these conditions may be so atypical that they cannot be said to form part of the overall distribution of travel times, and would have to be analysed separately.

iv) Effect of weather

There is some evidence to suggest that travel times are affected by heavy rain, road surface conditions and visibility. (Freeman et al, 1971)

v) Effect of time of day

As well as the above four effects, there are several reasons why travel times may depend on the time of day. These include: altered network characteristics due to bus lanes or parking restrictions which may apply for only part of the survey period; changes in signal linking plans; changes in the proportion of regular commuters who know the route well and are generally "in a hurry"; changes in the proportion of turning movements at junctions; changes in pedestrian activity affecting the probability of being stopped at protected crossing facilities.

vi) Effect of day of week

As we are only considering week days, this effect may be quite small, with most of the day to day variation being explainable in terms of the first four factors.

vii) Effect of time of year

Most of the seasonal differences should be explainable in terms of flow, composition and weather, but there may still be an effect due to changes in the journey purpose mix, particularly relating to school and university term times.

viii) Effect of secular trends

There are many reasons why travel times may be getting steadily longer or shorter, such as improvements in vehicle performance, gradual changes in the age distribution of drivers, deterioration in the road surface, marginal improvements to signal linking schemes, or gradual deterioration as base data for signal settings becomes more outdated.

2.4 Inter-route variation

Each of the above factors may be expected to influence travel times on an individual link or route. In addition, geometry and patterns of demand will cause variations in the nature of those influences between links and between routes. A total of five routes were studied, each of which was divided into between 10 and 15 links. While the data are available for a link-by-link analysis, this paper looks only at the variation in whole routes' travel times. The routes selected represent a wide range of urban arterials from high standard with grade separated junctions down to a narrow twisting route with many junctions passing through a suburban shopping area. If the relationship between travel time and the foregoing factors is found to be stable between routes (or if the differences can be explained by some factor such as frequency of junctions) then it could be expected to have wider applicability.

3. SURVEY DESIGN

3.1 Selection of study dates

Given the need to survey among routes and days of the week, we adopted a Latin square design as below.

Routes	Week No.				
	1	2	3	4	5
Kirkstall Road	Mon	Tue	Wed	Thu	Fri
Otley Road	Tue	Wed	Thu	Fri	Mon
Roundhay Road	Wed	Thu	Fri	Mon	Tue
York Road	Thu	Fri	Mon	Tue	Wed
Hunslet Road	Fri	Mon	Tue	Wed	Thu

Such an arrangement immediately suggested partitioning the survey into 5-week blocks. The advantage of this design is that the between-route variation can be separated out from the week to week variation and the day of week variation. Rather than collect data by day of week throughout the year, we concentrated the five davs per week data collection into two blocks which we expected to be as different as possible in terms of traffic and driving conditions. For the remainder of the year, one day of the week, Thursday, was selected to measure seasonal effects, with an overlap period at the end of the survey to allow for secular trends to be identified. Table 1 lists the dates for which data was collected and indicates that a total of 84 route-davs' data were collected, giving information for a minimum of 16 davs per route.

Table 1. The survev programme

Weeks beginning	Latin Square Treatment (1)
10 May - 14 June 1982 (2)	Full
21 June - 19 July	Thursday only
6 Sep. - 4 Oct.	Thursday only
11 Oct. - 8 Nov.	Thursday only (3)
15 Nov. - 13 Dec.	Thursday only
17 Jan. - 21 Feb. 1983 (2)	Full
28 Feb. - 28 Mar.	Thursday only
9 May - 20 June (2) (4)	Thursday only

1) Full treatment is as indicated at the head of this section; Thursday surveys were conducted on the appropriate route for the week in question.
2) With one week omitted for school half term.
3) With one week of full surveys to replace NPM surveys on week of 7th June.
4) With an extra week omitted for the general election.

3.2 Survey methods

In order to obtain data on inter-vehicle variations within 15 minute periods the number plate matching (NPM) method was used. High sample rates were achieved by using tape recorders, and an accurate time for the passage of each vehicle was obtained by interpolation between bench mark times spoken onto the tapes at roughly 5 minute intervals.

The NPM method was, however, too labour intensive and too dependent on light and weather conditions for use throughout the year. It was therefore used only in the first five week block (except for week 4 which had to be repeated in November due to technical problems which arose with the recorders).

The main method used throughout the whole survey was the moving vehicle observer method (MVO), as described by Wardrop and Charlesworth (1954). Two survey cars were used for overlapping periods of two hours each providing runs at approximately 10 minute intervals during the periods of highest flow. A total of 1018 runs were made on the five routes.

3.3 Supplementary data

Flow data were obtained from the NPM and MVO surveys, and the local authority's automatic traffic counters and UTC detectors. Composition was obtained from NPM and MVO. The MVO surveys provide information on incidents, weather and parking, and the UTC incident log provided details of reported accidents, breakdowns, roadworks and detector/signal malfunctions.

4. INTER-VEHICLE VARIATION

4.1 The data used

NPM data were collected for 10 periods in each of 5 days on each of 5 routes, giving a total of 250 survey periods. Using an initial range of permitted travel times of between 4 and 30 minutes, a total of 8596 matches were achieved. Table 2 indicates the crude success rate of matching, expressed simply as the proportion of vehicles entering the study section which were matched. While the full population of vehicles has inevitably not been recorded, the number matched exceeds 10% on all but 41 of the 250 survey

Table 2. Proportion of cars entering study section which were matched

Route	Week No. 1	2	3	5	6
Kirkstall Road	.304	.235	.285	.385	.393
Otley Road	.219	.158	.145	.280	.278
Roundhay Road	.146	.130	.115	.194	.110
York Road	.208	.199	.167	.195	.182
Hunslet Road	.081	.085	.100	.145	.086

4.2 Treatment of outliers

One of the main problems in using number plate matching to obtain travel time statistics is the presence of outliers. If number plates are matched and a high maximum travel time is allowed, the resulting distribution of travel times has a long thin right-hand tail. The travel times in this tail are either genuinely very slow-moving vehicles, vehicles stopping en route or deviating from the route, or the result of spurious matches.

The main cause of spurious matches is the fact that usually only part of the full registration plate is recorded. The usual method of overcoming this problem is to specify an upper limit to the travel times and thereby cut off the tail of the distribution. However, this is inevitably a somewhat arbitrary procedure, and raises the dilemma of whether a separate cut-off should be applied to each date/time period, or whether the same cut-off should be used throughout. The second option is computationally simpler, but means that some time periods will include more outliers than others, whereas with the first option we run the risk that a major determinant of mean travel time is the arbitrary upper cut-off value rather than any real effects.

In this study we are looking at possibly quite small variations in travel time between days/time periods, and it is therefore important that the variations we observe are real and not artefacts of the analysis procedure. The crux of the problem is that the most commonly used measure of 'average', the mean, is heavily influenced by outliers (especially as these tend to be on one side) and is therefore unsuitable unless the data is firstly sifted in a somewhat arbitrary fashion. An alternative approach is to use the median journey time in place of the mean, in the knowledge that the former is relatively unaffected by outliers. This however raises the problem that standard statistical techniques are based on mean, not median values.

The approach which we have adopted consists of the following.

a) After matching with a high upper cut-off value, the median and quartiles are calculated for each quarter-hour period.

b) Assuming that the natural logarithms of travel times in each period are approximately normally distributed, the standard deviation for each period is calculated (SD = interquartile range/1.350).

c) For each period all log-travel times greater than $\pm$ 3 standard deviations from the median are discarded.

d) The mean and standard deviation of the truncated data are calculated for each period.

Examination of the results of this approach, which is similar to that of Anscombe (1967), showed that the mean of the truncated data was generally somewhat lower than the mean of the original, whereas the median of both sets was generally very close to the mean of the truncated set. This is in accordance with expectations, indicating the stability of the median, and the symmetry of the truncated data.

Figure 1 presents the distribution of the standardised log-travel times, Z, given by

$$Z = (\bar{L}_k - L) / S_k \qquad (1)$$

where $L = \log_e(\text{travel time in seconds})$

$\bar{L}_k$ = mean of L for period k

S_k = standard deviation of L for period k.

The distribution contains a total of 7788 observations remaining after truncation for the 250 survey periods. It also presents the similar distribution for the 7690 observations remaining after truncation using a normal, rather than log-normal distribution. Comparison of the two distributions confirms the findings of other researchers (Richardson and Taylor, 1978) that travel times (in their case from MVO data) are better approximated by a log-normal distribution.

4.3 Inter-period variation from NPM data

The means of the 250 log-normal distributions provide a first indication of the variations between periods, days and routes. To enable inter-route comparisons, these have been expressed in terms of slowness (seconds / km) as defined by Hewitt (1974). Figure 2 presents the results for two routes, Otley Road, which exhibits substantial variation from day to day and between the shoulders and peak of the peak, and Kirkstall Road, which exhibits very little variation in either.

FIG 1a NPM DATA ALL ROUTES - ALL DAYS

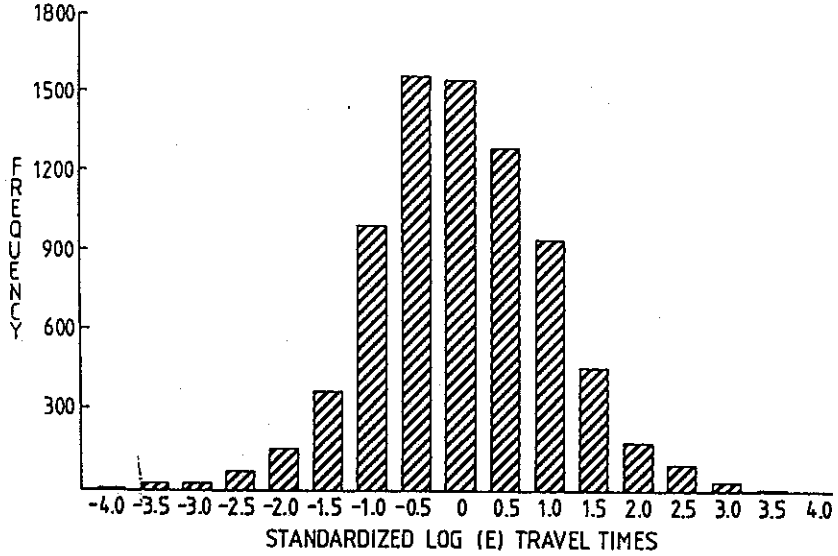

FIG 1b NPM DATA ALL ROUTES - ALL DAYS

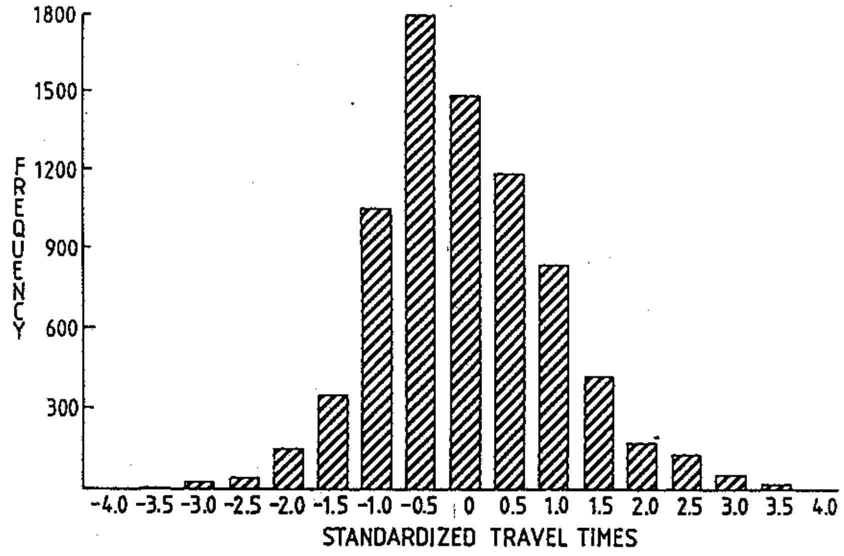

FIG 2a NPM SURVEYS OTLEY ROAD CARS

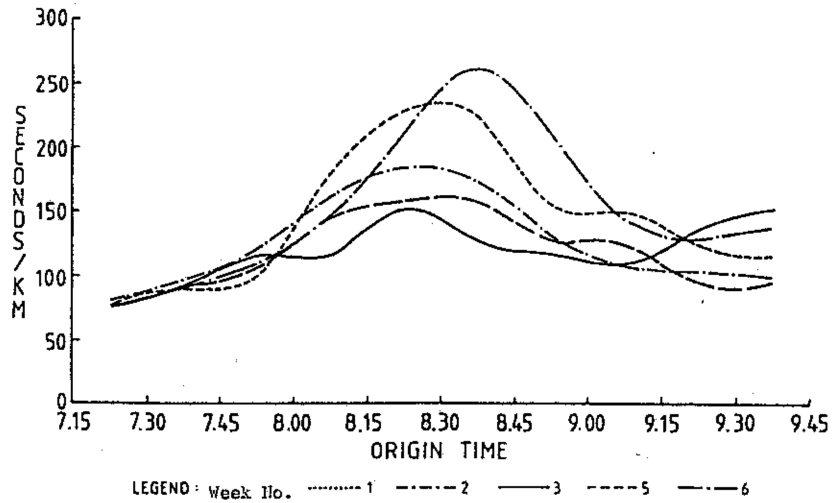

FIG 2b NPM SURVEYS KIRKSTALL ROAD CARS

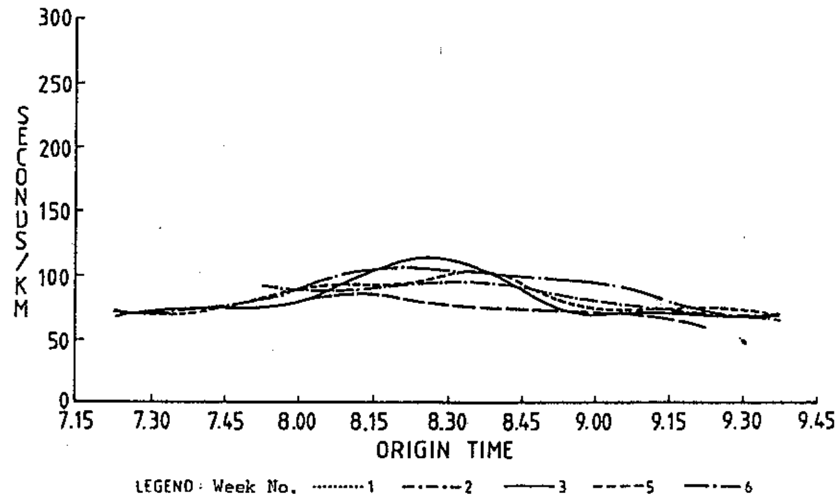

Table 3 indicates the off peak (07.15 - 07.30) slowness for each route and the highest and lowest peak slownesses (among the five days of observation for each route). It is noticeable that the off peak values are very similar between routes, but that some routes (notably Otley Road and Hunslet Road) exhibit variations from day to day in peak slowness which are greater than the difference between minimum peak slowness and off peak slowness.

Table 3. Off peak and range of highest peak slownesses

Route	Mean slowness (seconds/km)		
	Off peak	Range of highest peak values	
Kirkstall Road	70	90	110
Otley Road	80	150	260
Roundhay Road	70	110	150
York/Selby Road	80	130	200
Hunslet Road	100	140	260

A paired t-test of NPM medians with adjusted MVO values for the same periods found no significant difference between the two measures of slowness for the 250 survey periods, despite the fact that the vehicle populations for which they are measured are strictly different (May and Montgomery, 1984).

4.4 The nature of inter-vehicle variations

If t_{m-s}, t_m and t_{m+s} are the travel times whose logarithms are $\bar{L}_k - S_k$, $\bar{L}_k$ and $\bar{L}_k + S_k$ respectively, then

$$t_{m+s} / t_m = t_m / t_{m-s} = e^{S_k} \qquad (2)$$

and the non-dimensional parameter e^{S_k} provides a useful measure of the distribution of travel times. Table 4 indicates for each of the five routes and five days the range of values of e^{S_k} for period 1 (07.15 - 07.30, pre-peak) and for period 6 (08.30 - 08.45, the height of the peak). Three of the five routes have little inter-vehicle variation in slowness in period 1 on any day. Otley Road has much greater variation on some days than others suggesting, as does Table 3, that its operation is much less stable. Hunslet Road performs very differently from the other four, with high inter-vehicle variation on all days. The period 6 values are surprisingly similar to those for period 1, suggesting that the range of slownesses is proportional to mean slowness. The only noticeable differences are that Otley Road has more stable inter-vehicle variation at the height of the peak, while the reverse is true for Roundhay Road, and that York Road has more stable inter-vehicle variation but at a higher level than the pre-peak. The causes of these differences, which will depend on the micro-level performance of the routes, are beyond the scope of this study, but suggest an interesting area for further research.

Table 4. Ranges of values of e^{S_k} for periods 1 and 6 over five days' data

	e^{S_k} for period 1		e^{S_k} for period 6	
	min.	max.	min.	max.
Kirkstall Road	1.0795	1.1416	1.0457	1.1075
Otley Road	1.1003	1.6302	1.0730	1.3431
Roundhay Road	1.1568	1.1943	1.0834	1.3678
York Road	1.0791	1.2776	1.2306	1.2712
Hunslet Road	1.9650	2.1031	1.1941	2.0188

5. INTER-PERIOD VARIATION

5.1 Preliminary analyses

In Section 2.3 we outlined the factors thought a priori to be important in affecting the variation of average journey times between periods. Of these, it was thought that flow would be the most important. Table 5 shows the variables used to measure the various factors.

Fig. 3

MVO SLOWNESS - TOTAL MVO FLOW - OTLEY ROAD

REGIME ENDS ARE 8.00 & 9.00

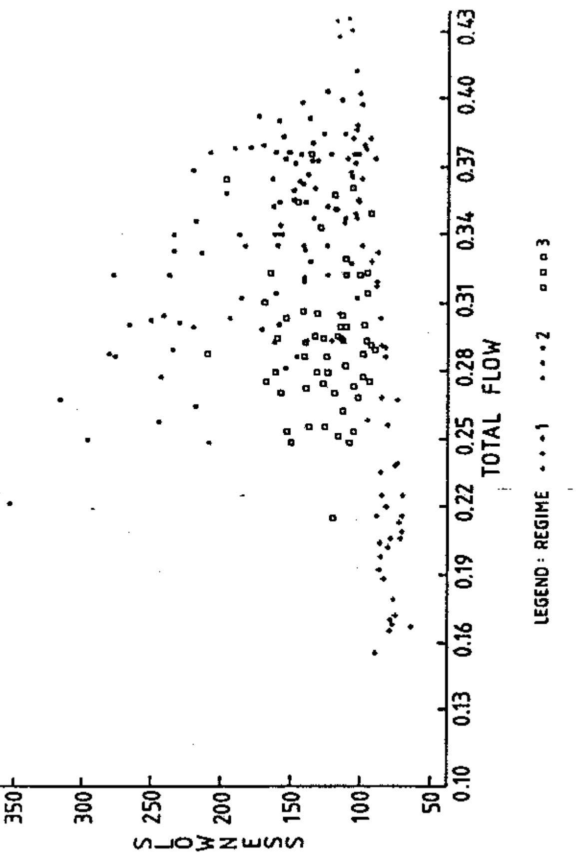

Fig. 4

TOTAL MVO FLOW - TIME OF DAY - OTLEY ROAD

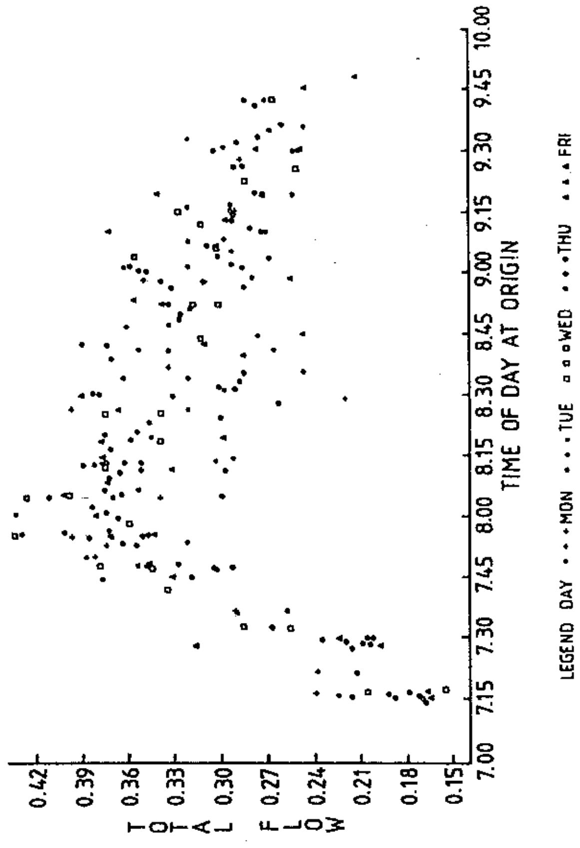

Table 5. Variables used in analyses of inter-period variation

Factor	Variables
Flow	Total flow in vehicles/sec.
Composition	Proportion of buses, proportion of c.v.
Route	Dummy
Time of year	Dummies for season, survey block, school term
Driver	Dummy
Day of week	Dummy
Time of day	Dummies for hour, before dawn, school patrols active
Weather	Dummies for rainfall, road surface, temperature, visibility
Incidents	Dummies for major and minor incidents*

* Based on observer's assessment.

Initial multiple linear regressions of slowness or log(slowness) against the variables in table 6 showed several of the variables (for route, time of year, day of week, time of day) as being more significant than total flow. This result was contrary to expectations, and it was thought that some of the variables were showing high significance because of cross-correlation with flow. This did not, however, explain why flow itself was poor as an explanatory variable.

Plots of slowness against flow by route showed very little correlation, but when the points were differentiated by time of day, an interesting pattern emerged as shown for Otley Road on Fig. 3. This pattern, which was repeated for the other routes, seemed to indicate a narrow range of slowness before 8 am, a very wide range between 8 am and 9 am, and a smaller range after 9 am. At no time, however, did there appear to be a relationship with flow.

Plots of flow by time of day were then produced, as in Fig. 4 for Otley Road, and showed the same general shape as the plots of slowness by time of day (Fig. 2). This suggested that any relationship between slowness and flow might involve a time lag.

5.2 Distributed lag structures

The use of lagged variables is a fairly recent phenomenon, and the literature on the subject is mainly from the field of econometrics, where they are used extensively in time series analysis. Some relevant texts on the subject are Judge, Griffiths et al (1980), Kendal (1976) and Dhrymes (1981). The general form of the distributed lag model is

$$Y_t = \alpha + \sum_{i=0}^{n} \beta_i X_{t-i} + e_t \qquad (3)$$

where

Y_t is the value of the dependent variable at time t,

X_{t-i} is the (lagged) value of the independent variable at time t-i,

and β_i are their coefficients

e_t is the error term.

The two main classes of lag structure are

a) infinite ($n = \infty$), where the effect of the independent variable never completely vanishes;

b) finite, where the independent variable is assumed to have zero effect after some (perhaps unknown) number of time periods n.

This formulation presents problems if the β_i are estimated by least-squares methods. If an infinite structure is assumed, there are an infinite number of β_i to be estimated; even if a finite structure is assumed, it is usually the case that the lagged variables are very highly cross-correlated. When this is the case estimation by least-squares would lead to imprecise estimates of the β_i. In order to overcome these problems, it is usual to impose some structure on the coefficients, reflecting the researcher's prior feeling about the distribution of the lag effect over time.

For the infinite case, one of the most useful assumptions is the geometric lag structure as described by Koyck (1954). This assumption allows Eq. (3) to be transformed in such a way that the problem of estimating an infinite number of parameters is reduced to estimating two. However the assumption of a monotonically decreasing function is somewhat restrictive and serious problems of auto-correlation can be introduced by the transformation, which may cause the least squares estimates of β_i to be inconsistent, inefficient and biased. For the finite case, several lag structures have been proposed including the arithmetic, inverted V and polynomial. The arithmetic lag involves lag weights decreasing according to a straight line function (a special case of the polynomial) and is therefore even more restrictive than the geometric. The inverted V is the simplest available structure which can cater for a rise and fall in lag weights, but again is rather restrictive in its formulation. With the polynomial lag structure as described by Almon (1965), the lag weights are a function of a polynomial of degree q where

$$\beta_i = C_o + C_1 i + \dots C_q i^q \tag{4}$$

This model has been widely used due to its flexibility, the relatively small number of parameters to be estimated, and the ease with which they can be estimated. The method is, however, not without its problems, including the unknown properties of the estimates of β and the effect of mis-specifying the maximum lag length.

5.3 Application of Lag Structures to Slowness-Flow models

There are several differences between typical econometric time series data and traffic slowness-flow data, the two most important of which are:

a) In econometric data the observations are equally spaced in time/or very nearly so), for example yearly, quarterly, monthly, weekly. In traffic data from MVO surveys, however, the observations are quite unequally spaced.

b) Econometric models usually refer to a single series, with consequent problems of high cross-correlation between the lagged variables. In the data as described in this paper, however, we are dealing with several series, one for each day and this may reduce the extent to which cross-correlation is a problem.

Almost all the literature on the fitting of distributed lag models assumes a constant time interval between observations, hence if we ignore the fact that MVO data is not equally spaced, we may have problems of mis-specification. The options available are therefore:

a.1) Assume that MVO runs are equally spaced in time, knowing this is not true;

a.2) Incorporate the actual lag length in the lag function;

a.3) Transform the "by run" data into "by period" data by calculating the mean slowness and flow for each 15 minute period, so that the interval between observations is constant.

The main problem caused by (b) above concernes the treatment of missing values. In any time series it is obvious that all lagged values will be missing for the first observation, similarly values for lags greater than one interval will be missing for the second observation. In econometric models it is usual to discard the initial observations containing missing values, and this is not a great loss as long as the maximum lag length is small compared to the number of observations. However in MVO data this procedure would mean discarding observations with missing values at the start of each series (day) so that the proportion of data lost would be much greater. Three options are available to overcome this problem:

b.1) Leave missing values as missing, and suffer the loss of data at the beginning of each day;

b.2) Assume missing values are zero. As we are dealing with flow before 7.15 am this is not an entirely unreasonable assumption;

b.3) Assume missing values are equal to the flow measured on the first run.

Initial regressions of slowness and log (slowness) for each run against flow for that run and the four preceding runs (without imposing any specific structure of the kind outlined in 5.2), showed a tendency for the weights to rise and then fall as the lag increased, so that by run (t-4) the parameters were not significant at 5%. When the data for individual runs was transformed to data for individual periods (option a.3), the parameters were significant only up to period (t-2) for all routes, suggesting that the effect of flow on slowness persists for a maximum of 30 minutes. Both the rise and fall, with increasing lag, and the finite lag effect suggested the use of the polynomial lag structure.

5.4 Estimation of polynomial lag parameters

If the observations are uniformly spaced (options a.1 and a.3) then we can apply Eq. (4) where i is the number of intervals between the dependent variable and the lagged independent variable.

Therefore for the example of a quadratic polynomial function:

$$\beta_i = C_o + C_1 i + C_2 i^2 \tag{5}$$

If we are working with lags up to t-3, then substituting the above in Eq. (3) gives:

$$\begin{aligned} Y_t = \alpha + C_o X_t &+ (C_o + C_1 + C_2)\, X_{t-1} \\ &+ (C_o + 2C_1 + 4C_2)\, X_{t-2} \\ &+ (C_o + 3C_1 + 9C_2)\, X_{t-3} \end{aligned} \tag{6}$$

Collecting terms gives

$$\begin{aligned} Y_t = \alpha &+ C_o(X_t + X_{t-1} + X_{t-2} + X_{t-3}) \\ &+ C_1(X_{t-1} + 2X_{t-2} + 3X_{t-3}) \\ &+ C_2(X_{t-1} + 4X_{t-2} + 9X_{t-3}) \end{aligned} \tag{7}$$

Which can be estimated by least squares. The coefficients C can then be substituted back in Eq. (5) to obtain the lag weights.

A similar procedure can be adopted for option a.2, replacing i in Eq. (7) by L_i, the actual lag length in minutes. While the coefficients can, as before, be estimated by least squares, the values of β_i cannot be estimated because the L_i themselves are not constant. Similarly, the variability in L_i makes it difficult to determine the number of lagged terms to include.

5.5 Choice of appropriate lag structure: an empirical example

The following describes the search process followed in seeking the 'best' model form to explain the relationship between slowness and flow. Regressions were performed using PROC GLM and PROC STEPWISE of SAS 79 (1979). For the most part, in the interests of clarity, only one route, Otley Road, is described.

Treatment of missing values As already mentioned in 5.3, three options (b.1 - 3) were available for the treatment of missing values of lagged flow at the start of each survey day. All three options were tested by regressing slowness against lagged flow up to run (t-3), assuming equal spacing between runs (option a.1).

Table 6 shows the values of r^2 obtained under each option where it can be seen that treating early missing values as zero or as the first recorded value causes an improvement for all routes except York Road. This is not

Table 6. Values of r^2 under different missing values treatments

Route	Missing values treated as		
	Missing (Option b.1)	= 0 (Option b.2)	= Run 1 value (Option b.3)
Kirkstall Road	.244 (120)	.274 (168)	.297
Otley Road	.238 (158)	.406 (209)	.370
Roundhay Road	.277 (160)	.322 (211)	.389
York Road	.292 (122)	.253 (173)	.202
Hunslet Road	.136 (206)	.183 (257)	.208

Number of observations in brackets.

surprising since the morningpeak period starts earlier on York Road than on the other routes, and hence flows at the start of the survey period are already quite high and rising, whereas on the other routes they are quite low. It is not clear which of the two options in Table 6 (b.2 and b.3) is the better overall, but option b.2 produces a slightly more consistent improvement over all routes, and was therefore adopted for the rest of the analysis.

Form of dependent variable Earlier work, as reported in 4.2, suggested that the variation in slowness was better described by a log-normal than a normal distribution. If this was so then one would expect a model with log (slowness) as the dependent variable to perform better than one with simply slowness. This hypothesis was confirmed using several regressions carried out under various lag structures. Table 7 shows the results for Otley Road.

Table 7. Values of r^2 under various lag structures for alternative forms of dependent variable (Otley Road only)

Dependent variable	Option	a.1			a.2		a.3	
	Lag Structure	A	B	D	A	D	A	D
Slowness		0.020	0.406	0.414	0.020	0.346	0.000	0.341
Log (Slowness)		0.066	0.508	0.506	0.066	0.475	0.045	0.435

A: No lag assumed
B: Unconstrained lag lag structure assumed)
C: Quadratic polynomial lag structure
D: Cubic polynomial lag structure
Missing values: = 0 (Option b.2)

Note that while a log-likelihood function (Maddala, 1979) is strictly a better test than r^2, the differences in r^2 values are sufficient to demonstrate the superiority of log (slowness).

Choice of lag structure Table 8 shows for Otley Road data the values of r^2 obtained for various lag structures and data sets, using log (slowness) as the dependent variable, and option b.2.

The results shown, which were mirrored for the other routes, demonstrated the importance of including a lagged relationship, and indicated that option a.1, which assumed equal intervals, was superior. While there is little to choose between the lag structures, the quadratic is easier to use and has been employed for what follows.

Table 8. Values of r^2 for various lag structures and data set options

Data set option	Lag Structure [1]	A	B	C	D
a.1	-	0.066	0.508	0.496	0.506
a.2	-	*	*	*	0.475
a.3	-	0.045	0.435	*	0.435

1: See Notes to Table 7
Dependent variable: log (slowness)
Missing values: 0 (Option b.2)
* Not tested

Lagged flow relationships After substituting the regression coefficients in Eq. (6) the following equations were obtained for the 5 routes in the order of Table 6.

$$L_t = 4.07 + 0.59Q_t + 0.40Q_{t-1} + 0.21Q_{t-2} + 0.02Q_{t-3} - 0.17Q_{t-4} \quad (8)$$

$$L_t = 4.37 - 0.88Q_t + 0.61Q_{t-1} + 1.22Q_{t-2} + 0.95Q_{t-3} - 0.20Q_{t-4} \quad (9)$$

$$L_t = 4.11 + 0.81Q_{t-1} + 0.98Q_{t-2} + 0.51Q_{t-3} - 0.60Q_{t-4} \quad (10)$$

$$L_t = 4.11 + 0.92Q_t + 0.64Q_{t-1} + 0.36Q_{t-2} + 0.08Q_{t-3} - 0.20Q_{t-4} \quad (11)$$

$$L_t = 4.32 + 0.70Q_{t-1} + 0.82Q_{t-2} + 0.36Q_{t-3} - 0.68Q_{t-4} \quad (12)$$

Where L_t is log (slowness) on run t; Q_{t-i} is total flow measured on run (t-i)

These equations demonstrate that lagged flows have no effect beyond run (t-3). Kirkstall Road and York Road have a less lagged effect than the others, and Otley Road and Hunslet Road have greater intercepts (which are consistent with their higher pre-peak slownesses in Table 3). There are clearly differences between routes which are not explained solely in terms of flow.

5.6 Effect on other variables

Inclusion of the other variables in Table 5 indicates the relative effect of other contributory causes and the amount of variation which they explain. For the example of Otley Road an r^2 value of 0.78 was achieved by including all the variables significant at the 5% level. In descending order of contribution they were the dummies for winter, school crossing patrols, Tuesday, visibility, the 0800-0900 regime, rainfall and term time. All of these are self-explanatory except Tuesday, which is not obviously different in terms of activities and flow patterns. It is interesting to note the relative importance of these factors and the contribution of the 0800-0900 regime, which suggests that differences in trip purpose, trip pattern and traffic control have an effect above that of flow.

5.7 Conclusions

The main conclusions to be drawn from the work reported in this section are

(i) that slowness is influenced by flows over the previous 30 minutes rather than solely current flows;

(ii) that these flows explain about half the variation in slowness;

(iii) that other factors explain just over half of the remaining variation;

(iv) that of these the seasonal (winter) effect was most important but that other driving condition, traffic control and movement pattern-related variables were also dominant;

(v) that there is still some as-yet unexplained inter-route variation;

(vi) that the application of time series analysis to this area provides a powerful evaluative tool.

REFERENCES

Almon, S. (1965). The Distributed Lag Between Capital Appropriation and Expenditures. Econometrica 33, 178-196.

Anscombe, F.J. (1967). Topics in the Investigation of Linear Relations Fitted by the Method of Least Squares. Journal of Royal Statistical Society, Series B, 29 1-52.

Branston, D. (1976). Link Capacity Functions: A Review. Transp. Research, Vol. 10, No.4.

Dhrymes, P.J. (1981). Distributed Lags - Problems of Estimation and Formulation. Amsterdam, North-Holland.

Freeman, Fox, Wilbur Smith & Assoc. (1971). Speed/Flow Relationships in Surburban Areas. Report to Dept. of Environment/Road Research Lab.

Hewitt, R.H., Chambers, J.B. and White, A.N. (1974). Graphical Solution of Moving Observer Surveys. The Highway Engineer. June 1974.

Judge, G.G., Griffiths, W.E., Hill, R.C. and Lee, T.C. (1980). The Theory and Practice of Econometrics. John Wiley and Sons.

Kendal, M. (1967). Time-Series. Charles Griffin and Co. Ltd.

Koyck, L.M. (1954). Distributed Lags and Investment Analysis. Amsterdam, North-Holland.

Maddala, G.S. (1979). Econometrics. McGraw-Hill.

May, A.D. and Montgomery, F.O. (1983). Travel Time Monitoring in Urban Areas. Proc., PTRC Summer Annual Meeting, Univ. Sussex, July 1983. PTRC Education and Research Services Ltd., London.

May, A.D. and Montgomery, F.O. (1984). Cost Effective Travel-Time Monitoring Strategies for Urban Radials. To be published.

Richardson, A.J. and Taylor, M.A.P. (1978). Travel Time Variability on Commuter Journeys. High Speed Ground Transportation Journal.

SAS79 (1979). SAS Users' Guide, 1979 Edition. SAS Institute Inc., Cary, N.C., U.S.A.

Wardrop, J.G. and Charlesworth, G. (1954). A method of estimating speed and flow of traffic from a moving vehicle. Journal Institution of Aviation Engineers, Part 11, 3(1).