
Published paper

Watson, SM., Clark, S.D., Redfern, E., Tight, M.R. (1992) *Development of an Influence Statistic for Outlier Detection With Time Series Traffic Data.* Institute of Transport Studies, University of Leeds. Working Paper 366

Working Paper 366

May 1992

DEVELOPMENT OF AN INFLUENCE STATISTIC FOR OUTLIER DETECTION WITH TIME SERIES TRAFFIC DATA

SM Watson, SD Clark, E Redfern and MR Tight

This work was sponsored by the Science and Engineering Research Council

CONTENTS

DEVELOPMENT OF AN INFLUENCE STATISTIC FOR OUTLIER DETECTION WITH TIME SERIES TRAFFIC DATA

INTRODUCTION

As part of a SERC funded project investigating the detection and treatment of outlying time series transport data, the practical applicability of the Influence Statistic described by Watson et al (1991) is assessed here.

Missing or outlying data occur in a variety of transport time series such as traffic counts or journey times for many reasons including broken machinery and recording errors. In practice such data is patched largely by subjective opinion or using simple aggregate methods.

In the analysis of non-transport time series several methods have been recently developed to both detect and treat outliers, including work by Kohn and Ansley (1986), Hau and Tong (1984) and Bruce and Martin (1989). These methods use either an intervention modelling approach (where the outlier is modelled as part of an ARIMA structure) or look at the influence an observation exerts on a particular parameter associated with the model.

An alternative is the Influence Statistic proposed by Watson (1987) and Watson et al (1992) which examines the influence of an observation on the sample autocorrelation function. Initial research showed the statistic has practical application in a transport context, and a replacement procedure based on the method was found to be effective in treating maverick data.

Here we report the results from a wider application of the statistic using traffic count data from the Department of Transport. Further developments are suggested and investigated for the replacement procedure and a comparison is made between possible variations in the method.

SECTION 1 - THE INFLUENCE STATISTIC FOR OUTLIER DETECTION AND REPLACEMENT

The Influence Statistic (IS_t) is based on the influence I on the theoretical autocorrelation function ρ_k of any pair of observations k time points apart, and may be defined.

$$IS_t = \frac{1}{p}\left[\sum_{L_t} I^2 + \sum_{D_{t-1}} I^2\right] \quad \ldots (1)$$

where $I = I(\rho_k, (y_t\, y_{t+k})) = y_t\, y_{t+k} - \rho_k (y_t^2 + y^2_{t+k})/2$ $\sum_{D_{Lt}}$ indicates summation over the L elements in the t'th row and $\sum_{D_{t-1}}$ represents summation in the (t-1)'th diagonal of the

Influence Function Matrix. $P = L + D_{t-1}$ where L is the maximum lag considered. The Influence Function Matrix described by Chernick et al (1982) is an n x L matrix of I values, where n is the length of the time series.

Watson (1987) shows that IS_t involves the sum of squares of products of normal variates and therefore it is possible to derive critical values for the statistic. Critical values vary according to the number of elements summed (P) in the matrix. For a given maximum lag L, P increases where $t<L$ and where $t > L$, P is constant at 2L, the t'th point being the observation of interest in the series. IS_t is a quantative statistic designed to improve on the subjective pattern recognition method advocated by Chernick et al (1982). An example of the critical value curves is given in Appendix A.

In the derivation of theoretical moments and critical values for the statistic, ρ_k is assumed to be constant. In practice a global "sample summary" estimate of ρ_k is needed, and empirical results (Watson, 1987) showed the following measure to be appropriate:-

$$r^* = \frac{|\max r_k| + |\min r_k|}{2} \qquad \ldots (2)$$

where $|\max r_k|$ and $|\min r_k|$ are the absolute values of the maximum and minimum sample autocorrelation values respectively. The r^* value is intended to be a representative measure of the range of values expected to occur in the autocorrelation function because of the assumption of constant r_k. Clearly a "better" summary measure may well exist, although with simulated ARIMA processes the number of outliers detected using r^* only rose consistently above nominal levels for $\phi = \geq 0.6$.

One of the features of the statistic is that a model fit is not required before outlying and influential data are detected. However, examination of the sample autocorrelation function may indicate a suitable value for the maximum lag. Unless significant spikes are seen at a higher lag, L=5 was found empirically to be an appropriate level.

Further manipulation of IS_t suggests a possible replacement or estimation procedure for use with outlying or missing observations. It can be shown (Watson, 1987) that a replacement value is given by

$$(\acute{Z}_t \times S) + \bar{Y} \qquad \ldots (3)$$

where

$$\acute{Z}_t = \frac{Z_{t+k}}{r_k}\left(1-\sqrt{1-r_k^2}\right) \qquad \ldots (4)$$

$\bar{Y}$ and S being the mean and standard deviation of the original series, Z_t representing the transformed data.

Simulation studies with series of normal variates showed that the replacement procedure works very well (see Watson 1987). Initial results with traffic count data have also been successful (Watson et al, 1992). The adjustment procedure (2) does of course depend upon

the sample acf at lag k, giving several possible adjustment values. At this stage the value chosen was that at the lag for which r_k (k = 1,2,...L) was largest. In section 4 we explore alternative approaches to selecting the adjustment.

SECTION 2 - OUTLIER DETECTION WITH TRAFFIC COUNT DATA

Several time series of traffic counts have been obtained from the Department of Transport, consisting of traffic counts for different road types throughout the day. Here we look in detail at trunk road data in a built up area (denoted T/B) on one particular route, where counts have been made in both east and west bound directions. Four time series are available in each direction, with counts throughout the day beginning at 08:00, 12:00, 17:00 and 20:00 in each case. One of the series (T/B East 20:00) is analysed in some detail in Redfern et al (1992), where the outlier detection methods given by ARIMA modelling, Tsay (1988), the Influence Statistic and "by eye" are compared. Sample results for two of the series have also been reported in Watson et al (1992). Here we aim to use the time series to illustrate further development of the outlier detection and replacement procedure.

Time series plots of the data are shown in figures 1-8 (Appendix B). Several features are clear from the plots including non-stationarity (wandering mean and variance) and high extreme values, which could contrive to produce difficulties in model fitting. Missing values were coded as zero for our purposes, although on rural roads a zero value could be a genuine observation. Depending upon the overall level of the time series a zero value may not be influential in a model fit, and in non-transport time series negative values may well be expected as part of the series.

Using the Influence Statistic outlying data can be detected without firstly fitting a model to the series, which for a transport practitioner would reduce the amount of statistical expertise needed in screening the data. However examination of the series sample autocorrelation function is recommended as this should highlight the presence of autocorrelation at high lags. A maximum lag of 5 was found to be adequate in the absence of significant autocorrelation at higher lags.

For the traffic count time series a maximum lag of 8 was used as the autocorrelation function showed a large spike at lag 7. A computer algorithm was used to output values of the Influence Statistic for each observation. One slight disadvantage with the method is that the statistic cannot be computed for either the first or last L observations in the series. This can be overcome by applying the method twice with forward then backward calculation, although a visual inspection of the time series may indicate whether outliers are suspected in the tails of the series.

A summary of the results found is given in Table 1, together with those points detected using the "High Residual" method. The latter process involved fitting an "optimum" Box-Jenkins ARIMA model to the series and highlighting those observations with a residual greater than 3 standard deviations from the residual mean. Critical values are given for each series together with the r* value used to determine the critical value. Note that results shown in Table 1 refer to a single application of the outlier detection method. No attempt has been made to replace the outlying data and re-test at this stage. Consequently the results are at variance with those seen overall in Watson et al (1992), where replacement values were iteratively substituted.

From Table 1 a higher number of outliers or missing observations are seen to be common to both methods for Eastbound traffic than is seen for Westbound traffic, although no obvious reason for this is clear. Overall more observations are picked out solely by the Influence Statistic than are suggested by the High Residual method alone although for individual series this is not always the case. In several series (for example T/B EAST 12:00) a small run of consecutive points are detected. Where this occurs it is unlikely that every point is a genuine outlier, but adjustment of one of the points may render its neighbours insignificant. This phenomena is illustrated further in Section 3. Where observations are picked out by both methods this may be confirmation that these are the "genuine" outliers (or missing data) in the series, and that other consecutive points detected by each method should be viewed with more caution.

ROUTE T/B 912 EAST

SERIES	BOTH METHODS	HIGH RESIDUAL ONLY	INFLUENCE STATISTIC ONLY	CRITICAL VALUE (CV)
08:00	27, 65, 72	88, 122, 128	24, 73, 79, 80	R* = 0.357 CV = 5.4
12:00	72, 88, 89 90	27, 57	2, 79, 86, 92, 93	r* - 0.365 CV = 5.3
17:00	45, 66, 72	27, 52	88	
20:00	52, 66, 72, 88	27, 122	95	r* = 0.380 CV = 5.2

ROUTE T/B 912 WEST

SERIES	BOTH METHODS	HIGH RESIDUAL ONLY	INFLUENCE STATISTIC ONLY	CRITICAL VALUE
08:00	51, 65, 72	27, 93, 96	58, 62	R* = 0.262 CV = 6.2
12:00	72	18, 24, 27, 89, 122	23, 44, 65, 79	r* = 0.442 CV = 4.5
17:00	27, 72	13, 36, 40, 122	30, 35, 79, 85, 86	r* = 0.232 CV = 6.5
20:00	85	2	64, 72	r* = 0.42 CV = 4.7

Table 1: Summary of Outlier detection using the Influence Statistic and high residual methods.

SECTION 3 - OUTLIER REPLACEMENT WITH TRAFFIC COUNT DATA

In the comparison between the Influence statistic and High Residual outlier detection methods, no attempt was made to infill missing values or replace suggested outliers. Using the procedures described by equations 3 and 4 this became possible and may be instrumental in unmasking other potential outliers not apparent on the first sweep of the series.

The results of an 'iterative' outlier detection-then-replacement cycle for the T/B East data are described within Table 2. Following the first pass through the series those points highlighted by the influence statistic as missing or outlying were replaced with 'new' values constructed using the procedure described in Section 2. A further pass through the series then uncovered further suspect points which had been masked by those detected in the initial screening. These were then replaced and further passes made until no observations were found to be significant by the Influence statistic. In Table 2 a comparison is made between the new replacement value and the original observation where zero indicates a missing value. The following features are apparent from Table 2.

i) Up to 5 passes were needed to unmask all potential outliers/missing values although for T/B East 08:00 a single pass was sufficient.

ii) As may be expected, less variation is seen in the values of the replacement than occurs in the original series.

iii) A substantial number of points are masked by those initially detected in all series except T/B EAST 08:00 where no masking is apparent.

A summary of the points detected in all the T/B EAST and WEST series from all passes through the data is given in Table 3. A certain amount of clustering is now apparent with several adjacent or neighbouring points detected by the statistic. Those points detected by both the influence statistic and the High Residual Method in the original list are underlined. A comparison with Table 2 indicates an increase in the number of points detected by both methods.

SERIES	OBS	NEW	OLD
T/B EAST	24	174	67
08:00	27	233	0
PASS 1	65	139	426
	72	275	0
	73	166	83
	79	261	367
	80	172	74
T/B EAST	2	288	537
12:00	72	287	0
PASS 1	79	299	522
	86	269	440
	88	357	530
	89	321	607
	90	269	474
	92	340	446
	93	275	543
PASS 2	1	261	467
	27	238	0
	57	274	486
	122	238	0
PASS 3	76	254	388
PASS 4	66	282	415
T/B EAST	45	530	1164
17:00	66	430	995
PASS 1	72	338	0
	88	333	1118
PASS 2	122	387	0
	136	354	643
PASS 3	52	444	952
PASS 4	17	333	607
	108	370	607
PASS 5	129	339	476
T/B EAST	52	137	628
20:00	66	136	535
PASS 1	72	184	0
	88	131	701
	95	110	68
PASS 2	73	136	352
PASS 3	87	174	358

Table 2: Iterative detection and replacement for T/B EAST series.

T/B EAST	
SERIES	**ALL PASSES**
08:00	24, 27, 65, 72, 73, 79, 80
12:00	1, 2, 27, 57, 66, 72, 76, 79, 86, 88, 89, 90, 92, 93, 122
17:00	17, 45, 52, 66, 72, 88, 108, 122, 129, 136
20:00	52, 66, 72, 73, 87, 88, 95
T/B WEST	
08:00	27, 44, 51, 58, 62, 65, 72, 75
12:00	2, 23, 27, 30, 37, 44, 45, 51, 58, 65, 66, 72, 79, 86, 87, 88, 93
17:00	6, 7, 27, 30, 35, 36, 40, 72, 79, 85, 86, 122
20:00	64, 71, 72, 77, 85

Table 3: Potential outliers/missing values from all passes through series (points detected by Influence Statistic and High Residual method underlined)

SECTION 4 - DEVELOPMENT OF REPLACEMENT PROCEDURE USING THE INFLUENCE METHOD

Where outlying or missing data have been detected in a series, a suggested replacement has been computed using y_t^1 computed at the lag at which the largest sample autocorrelation exists. Clearly other criteria could be used in the computation of a replacement and the existing methodology may not be the optimum. Here we examine three possible alternatives and use a simulation study to assess their relative performance. In Section 5 the alternatives are applied to time series of traffic count data.

Different solutions we considered for a replacement value are given by the following:-

1. Replacement using the largest adjustment computed over L ie using max $|Z_t^1|$ where Z_t^1 is given by (4).

2. Replacement using the adjustment at the lag for which max $|r_k|$ occurs.

3. Replacement using an averaged adjustment computed over the range of lags considered.

   ie using $\frac{1}{L}\sum_{i=1}^{L} Z_t^1$

4. Replacement using an overall adjustment computed from the average of the adjustment at the lag at which max $|r_k|$ occurs and the adjustment at the lag at which min $|r_k|$ occurs.

ie using $\dfrac{[Z_t^1]_{\max|r_k|} + [Z_t^1]_{\min|r_k|}}{2}$

Each possible replacement has been selected in order to provide an estimate for a missing or outlying point which could reasonably reflect a value within the original series without undue influence on r_k.

A comparison between the performance of each replacement is initially made using simulated low order stationary AR and MA series. Our primary concern is with the treatment of transport data, which are seen to be generally non-stationary with seasonal components. A comparison made initially with more 'well behaved' time series is less likely to be influenced by the peculiarities of individual series.

100 time series have been simulated (length 200) for each of five AR (1) models and five MA (1) models. Theoretical parameters were chosen across the range possible within stationarity bounds.

A single pass of the influence statistic was made through each of the series to establish the number of potentially outlying data within the simulated time series. The results are shown in Table 4 below, with the frequency distribution of the number of suspect points in each series. It can be seen from Table 4 that 35 of the 100 AR(1) series generated with ϕ_1 = -0.6 contained 2 potential outliers, for example.

		No of Outliers in Series											
MODEL		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**>10**
AR (1)	Φ_1 = 0.9	13	21	18	20	14	5	4	3	1	0	0	1
	Φ_1 = 0.3	18	14	21	15	12	9	5	1	3	1	1	0
	Φ_1 = 0.2	16	22	24	19	2	12	4	1	0	0	0	0
	Φ_1 = -0.45	19	19	27	19	10	5	1	0	0	0	0	0
	Φ_1 = -0.6	10	18	35	19	9	3	3	2	0	0	1	0
MA (1)	θ_1 = 0.99	14	20	27	12	11	8	7	0	1	0	0	0
	θ_1 = 0.67	34	16	18	8	8	2	3	3	1	4	2	1
	θ_1 = 0.36	10	18	28	17	15	7	2	3	0	0	0	0
	θ_1 = -0.1	17	18	30	12	10	7	4	2	2	0	0	0
	θ_1 = -0.2	13	17	19	15	13	8	8	1	2	3	1	0

Table 4: Frequency of No of Potential Outliers in Simulated Series.

As the series were generated of length 200, a maximum of 10 outliers may be expected to occur by chance within each. Although this number was exceed by one Autoregressive and one Moving Average Series, a large proportion of the simulated data not containing any points highlighted by the statistic.

In order to attempt to assess the relative merits of each replacement criteria, a single data point (t=90) was removed and then a replacement computed according to each method. The 90th data point was chosen as one which had not been highlighted as a potential

outlier in any of the simulated series, and is not contained within the initial or final stretch of the data. The four possible replacements were then compared with the original data, and that replacement closest to the true value noted as the 'best'. A summary of the results is shown in Table 5 below.

MODEL		REPLACEMENT CRITERIA			
		1	2	3	4
AR (1)	$\Phi_1 = 0.9$	38	41	16	25
	$\Phi_1 = 0.3$	34	46	27	11
	$\Phi_1 = 0.2$	34	47	25	8
	$\Phi_1 = -0.45$	52	47	31	8
	$\Phi_1 = -0.6$	58	52	26	7
	TOTAL	216	233	125	59
MA (1)	$\theta_1 = 0.99$	47	52	30	11
	$\theta_1 = 0.67$	48	56	26	7
	$\theta_1 = 0.36$	44	49	19	14
	$\theta_1 = -0.1$	33	34	32	13
	$\theta_1 = -0.2$	33	34	30	17
	TOTAL	205	225	137	62

Table 5: Successful missing value estimation for alternative replacement criteria.

The contents of Table 5 indicate the number of times, out of the 100 replications for each model a particular criteria best estimated the pseudo-missing value. For example, in the 100 series generated as AR (1) with $\phi_1 = 0.9$, replacement criteria (2) was most successful in estimating the missing value in 41 series and criteria (3) was most successful in 16 series. For each series, the "best" estimate of the missing value was defined as the closest estimate to the original value in absolute terms. Where different criteria gave equally good estimates, both were marked as successful.

From Table 5 it is clear that overall criteria (2) performed most successfully for both simulated Autoregressive and Moving Average Series, although criteria (1) also performed well for many series. As the adjustment at the lag for which max $|r_k|$ occurs and the largest adjustment computed gave the same solution in some cases, this may be expected.

An additional measure of how well each method replaces outliers may be given by the size of the replacement with respect to the variance of the individual series. Replacements computed which fall outside say, 3 standard deviations either side of the series mean could be considered extreme themselves. In fact upper and lower limits were computed for each of the 1000 simulated series using a 2 standard deviation criteria, and for each series all four replacement methods gave solutions within the mean ± 2 standard deviations limits.

A further comparison was drawn between criteria by superimposing an additive outlier at the 90'th point in each of the simulated series. The size was made relative to the scale of each series by adding (or subtracting) a multiple of the standard deviation of the

original data. Results for outliers of size 3σ, 5σ and 7σ for the simulated autoregressive and moving average series are shown in Tables 6 and 7 respectively. Again the number of times each method most closely replaced the original data was noted.

Perhaps the most noticeable feature of both tables in comparison with results shown in Table 5 (no imposed outlier) is that Method 1 now performs considerably less well than the alternatives.

For both autoregressive and moving average data criteria 2 was best able to replace the artificial outlier, and was consistent in doing so as the size of the outlier was increased. This confirms earlier findings as reflected in Table 5. Increasing the magnitude of the outlier appears to have little effect on the relative performance of each criteria, with method 3 most often the 'second best' in replacement.

	SERIES PARAMETER					
	Method	**ϕ_1 = 0.9**	**ϕ_1 = 0.3**	**ϕ_1 = 0.2**	**ϕ_1 = -0.45**	**ϕ_1 = -0.6**
3σ	1	7	1	7	0	2
	2	50	58	48	52	51
	3	17	32	24	34	24
	4	29	9	23	14	24
5σ	1	8	1	8	1	2
	2	49	58	47	49	52
	3	16	30	22	34	25
	4	30	11	25	16	23
7σ	1	6	1	8	2	2
	2	48	59	48	48	54
	3	18	29	20	33	22
	4	31	11	26	17	23

Table 6: Performance of Replacement Methods with Additive Outlier - Autoregressive Series

	SERIES PARAMETER					
	Method	**θ_1 = 0.99**	**θ_1 = 0.67**	**θ_1 = 0.36**	**θ_1 = -0.1**	**θ_1 = -0.2**
3σ	1	2	1	1	2	4
	2	54	43	51	48	46
	3	32	31	31	27	26
	4	13	24	17	24	26
5σ	1	2	2	1	2	4
	2	53	44	49	49	48
	3	32	31	29	25	26
	4	13	23	21	24	24
7σ	1	2	2	1	2	4
	2	54	46	45	50	50
	3	31	27	27	22	22
	4	13	25	27	26	26

Table 7 Performance of Replacement Methods with Additive Outlier - Moving Average Series

SECTION 5 - COMPARISON OF REPLACEMENT RESULTS FOR TRAFFIC COUNT DATA

Using the T/B East and West bound traffic counts, the performance of each of the four replacement criteria can be assessed with real data. For each time series, three points (observations 60, 70 and 100) were deleted to create artificial missing values. These observations were chosen as they had not been detected as potential outliers within any of the series and they are not within the tails of the series. Each of the four replacement criteria were then applied to estimate the pseudo-missing values. Results are given within Table 8, where the original value of the 60'th, 70'th and 100'th observation is shown with the estimates from each of the four replacement criteria. An asterix indicates the closest estimates(s) in each case. From the summary row at the bottom of Table 6, criteria (2) and (3) were most successful in estimating the deleted point, with (2) performing best overall. This result is consistent with those described in section 4 using simulated time series.

T/B EAST 08:00		CRITERIA		
OBS (VALUE)	**(1)**	**(2)**	**(3)**	**(4)**
60 (242)	214.9	220.3 *	220.2 *	217.6
75 (203)	225.4	221.4	220.4 *	223.0
100 (306)	218.8	263.1 *	227.4	241.6
T/B EAST 12:00		**CRITERIA**		
OBS (VALUE)	**(1)**	**(2)**	**(3)**	**(4)**
60 (264)	282.9	266.4	268.3	264.5 *
75 (268)	297.7	297.7	275.7 *	288.5
100 (315)	255.6	256.1	261.0 *	255.9
T/B EAST 17:00		**CRITERIA**		
OBS (VALUE)	**(1)**	**(2)**	**(3)**	**(4)**
60 (446)	360.8	379.6 *	376.7	370.2
75 (397)	344.0	366.0 *	356.8	357.7
100 (260)	383.0	325.6 *	350.0	346.3
T/B EAST 20:00		**CRITERIA**		
OBS VALUE	**(1)**	**(2)**	**(3)**	**(4)**
60 (145)	186.3	143.9 *	152.7	139.3
75 (142)	120.8	122.2	126.5 *	125.3
100 (90)	142.4	142.4	120.6 *	137.8

Table 8: Comparison of replacement criteria Using traffic count data

T/B WEST 8:00		**CRITERIA**		
OBS VALUE	**(1)**	**(2)**	**(3)**	**(4)**
60 (165)	153.6	166.1 *	161.3	161.0
75 (185)	149.2	156.2 *	153.9	155.6
100 (134)	157.2	143.6 *	152.1	146.8
T/B WEST 12:00		**CRITERIA**		
OBS VALUE	(1)	(2)	(3)	(4)
60 (408)	359.4	380.4	386.5	402.9 *
75 (425)	343.5	365.7 *	358.5	357.3
100 (589)	330.6	419.7 *	347.8	377.6
T/B WEST 17:00		**CRITERIA**		
OBS (VALUE)	(1)	(2)	(3)	(4)
60 (376)	298.5	317.6 *	304.4	308.8
75 (312)	300.0	314.9 *	306.5	307.4
100 (387)	292.7	325.6 *	302.6	312.8
T/B WEST 20;00		**CRITERIA**		
OBS (VALUE)	(1)	(2)	(3)	(4)
60 (140)	126.9	138.7 *	141.9 *	143.0
75 (132)	144.0	136.0 *	140.8	140.3
100 (152)	140.3	141.6	144.9 *	142.5
	(1)	**(2)**	**(3)**	**(4)**
'BEST' ESTIMATE	0	16	8	2

Table 8 continued

CONCLUSIONS

An outlier/missing value replacement procedure is suggested here which generates a range of possible values according to the value of the sample acf at the k'th lag. The procedure is an extension of the outlier detection statistic, IS_t which highlights potential outliers according to their influence on the sample acf.

A comparison is made between those outliers/missing values detected using the influence statistic, IS_t and those found using the High Residual Method. The latter involved the fitting of optimum Box-Jenkins (1976) models to the time series and flagging points with a residual greater than 3 standard deviations from the residual mean. From the 8 time series of traffic counts considered, a number of points were found to be common to both methods. Overall more observations were picked out by the Influence Statistic with several small runs of consecutive points.

An iterative cycle of detection and replacement was then applied to uncover further potential outliers which may have been masked by the original group. Replacements were made using the $|\max r_k|$ criteria and up to 5 passes through the data were needed. This iterative cycle increased the total number of points detected and increased the number of points found to be common to both the Influence Statistic and High Residual Methods.

Given that several replacement values were possible using the procedure, it was felt that the original choice need not necessarily be optimum. Three additional alternatives are therefore defined to assess whether this is the case. A comparison was made using firstly 1000 simulated time series with low order underlying AR and MA structures. The simulated series were found to contain few outliers and could be expected to be more 'well behaved' than real life traffic count data. Criteria (2) and (1) ($\max |r_k|$ and 'largest adjustment' respectively) performed best on the simulated data ie estimated a pseudo-missing point closest to its true value.

An additional check was made on whether the replacements computed by each method fell within 2 standard deviations of the series mean. For all methods on each simulated series this was the case.

Artificial additive outliers were then created within the 1000 series of magnitude 3σ, 5σ and 7σ. Method 2 was again the most successful in estimating the 'true' data value, but method 3 was now second most successful, whilst method (1) performed poorly across all series and each magnitude of outlier. An even more comprehensive assessment of each replacement criteria could be made by superimposing different types of outliers on the simulated series ie Innovative Outliers level change and so on (see for example Redfern et al, 1992)

A repeat of the comparison was made using 8 time series traffic counts. Here 3 pseudo missing values were generated in each series and all four replacement criteria re-assessed. Similar results were found to these achieved using the simulated series, with the optimum replacement criteria being that suggested by the procedure at the lag at which $|\max r_k|$ occurs. In practice any of the replacement methods can be easily programmed into an algorithm which calculates the Influence Statistic. Results described here indicate that whilst all four criteria would give a suitable replacement for missing or outlying data, the max $|r_k|$ adjustment would probably be optimum.

REFERENCES

BOX, G.E.P. and JENKINS, G.M. (1976) "Time Series Analysis, Forecasting and Control". HOLDEN-DAY.

BRUCE, A.G. and MARTIN, R.D. (1989) "Leave-K-out Diagnostics for Time Series." JRSS.B, 51, 363-424.

CHERNICK, M.R., DOWNING, D.J., and PIKE, D.H. (1982) "Detecting Outliers in Time Series Data." J.A.S.A., Vol 77, 380, 743-747.

HAU, M.C. and TONG, H. (1984) "Outlier Detection in Autoregressive Time Series Modelling." Technical Report 15. Department of Statistics, Chinese University of Hong Kong, Hong Kong.

KOHN, R. and ANSLEY, C.F. (1986) "Estimation Prediction and Interpolation for ARIMA Models with Missing Data." J.A.S.A. 81, 751-761.

REDFERN, E.J., CLARK, S.D., WATSON, S.M. and TIGHT, M.R. (1992a) "A Comparative Study of Outlier Detection and Missing Value Estimation Methods Applied to Time Series Transport Data". Computational Statistics Vol 1, 349-354. Ed Y. DODGE AND J. WHITTAKER. THYSICA-VERLAG.

REDFERN, E.J., CLARK, S.D., WATSON, S.M. and TIGHT, M.R. (1992b) "Identifying Outliers and Other Events in Seasonal Time Series". (To be published)

TSAY, R.S. (1988) "Outliers, level shifts and variance changes in time series". Journal of Forecasting, 7, 1-20.

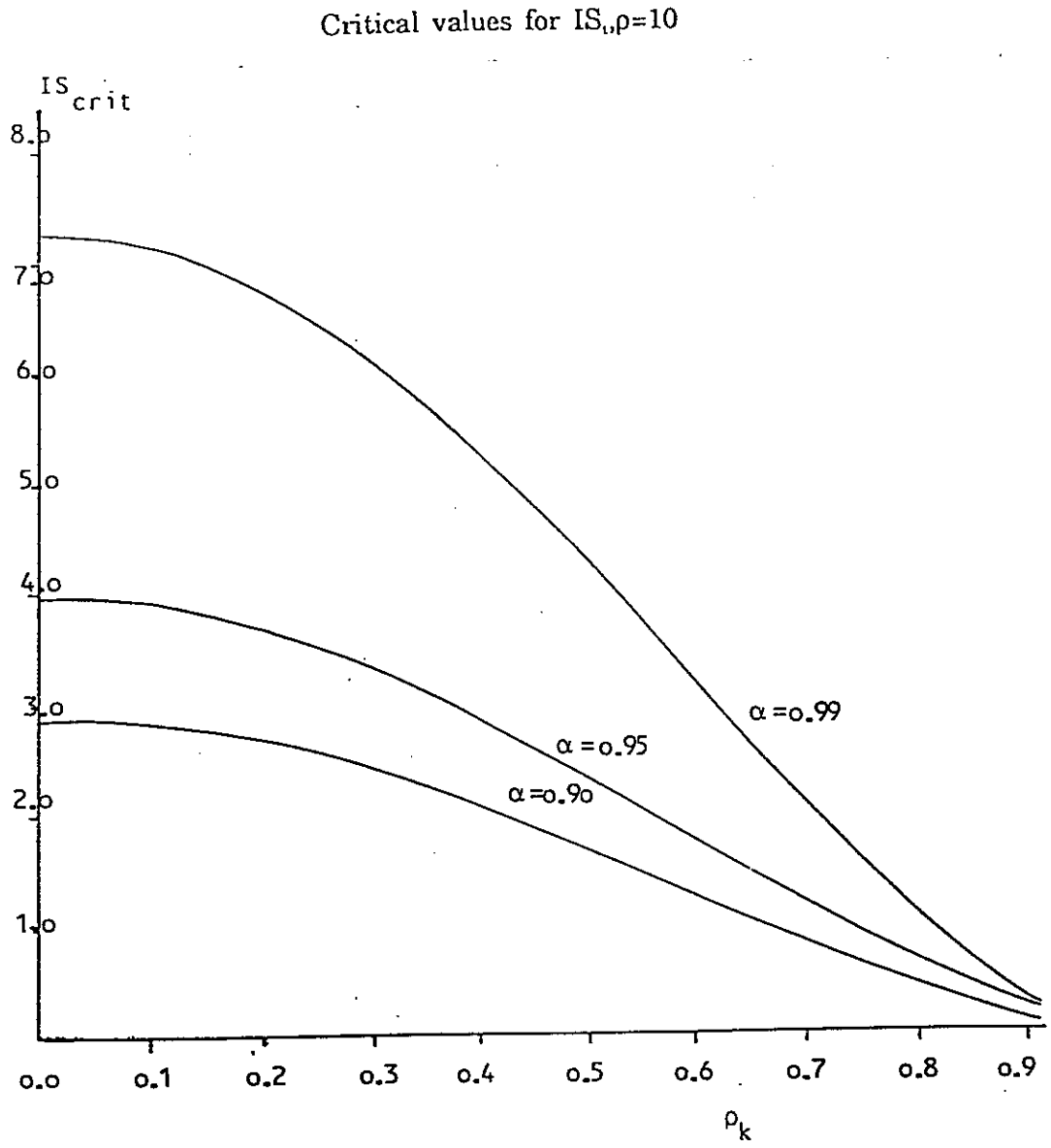
WATSON, S.M. (1987) "Non-normality in Time Series Analysis." Unpublished PhD Thesis, Trent Polytechnic.

WATSON, S.M., CLARK, S., TIGHT, M.R. and REDFERN, E. (1992). "An Influence method for Outlier Detection Applied to Time Series Traffic Data". WP365, Institute for Transport Studies, University of Leeds, Leeds.

smw/wp366

Appendix A

Example of critical value curves.



Appendix B

Figs 1-8, Time series plots of traffic count data.

FIG 1 Class T/B, East at 08:00

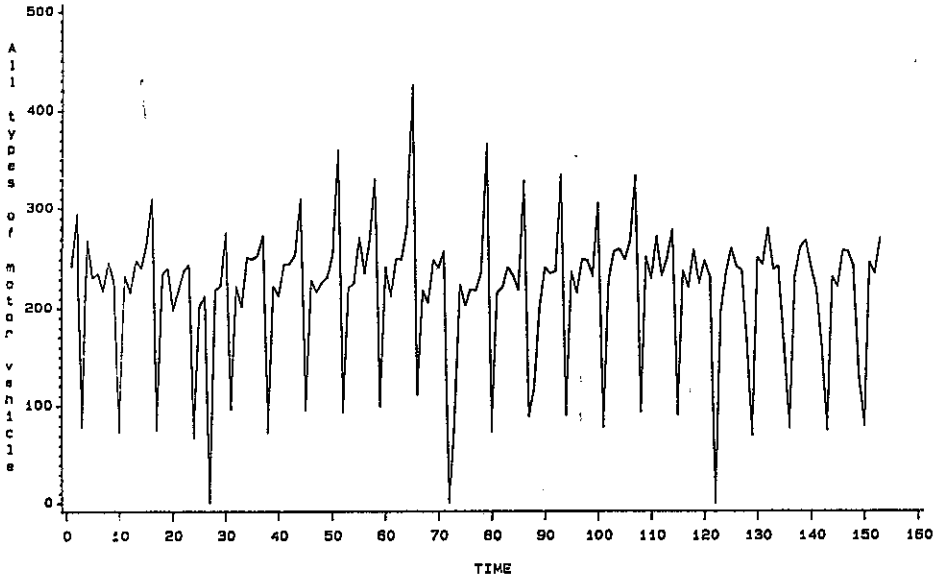


FIG 2 Class T/B, East at 12:00

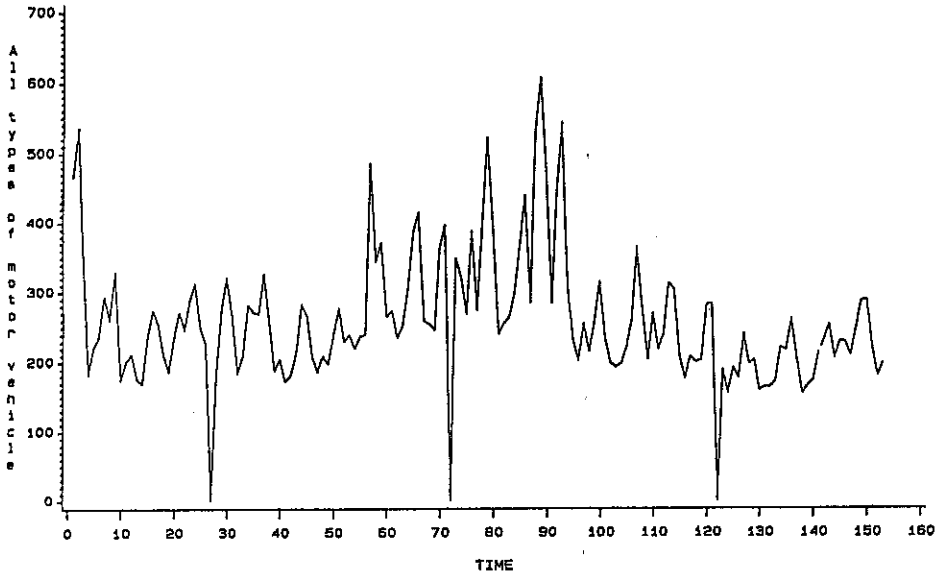


FIG 3 Class T/B, East at 17:00

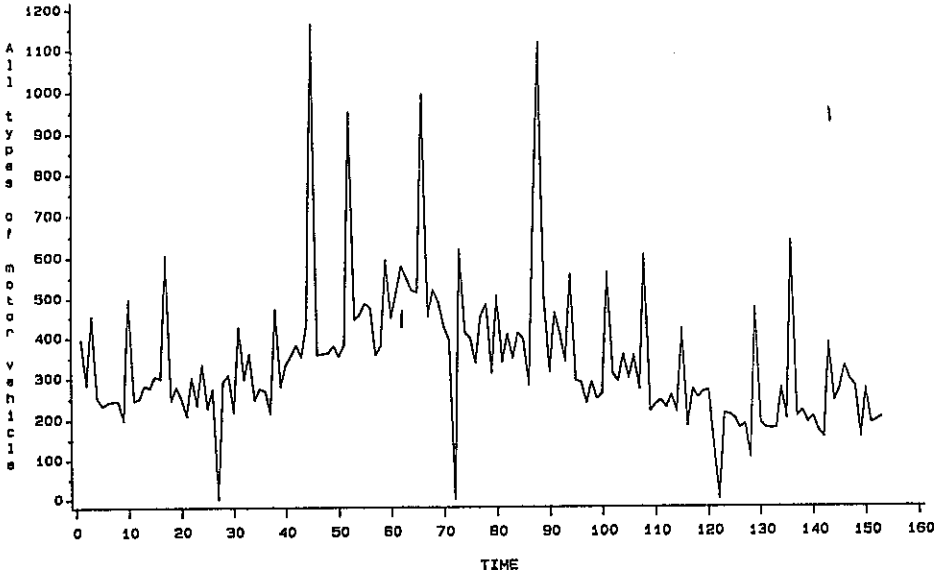
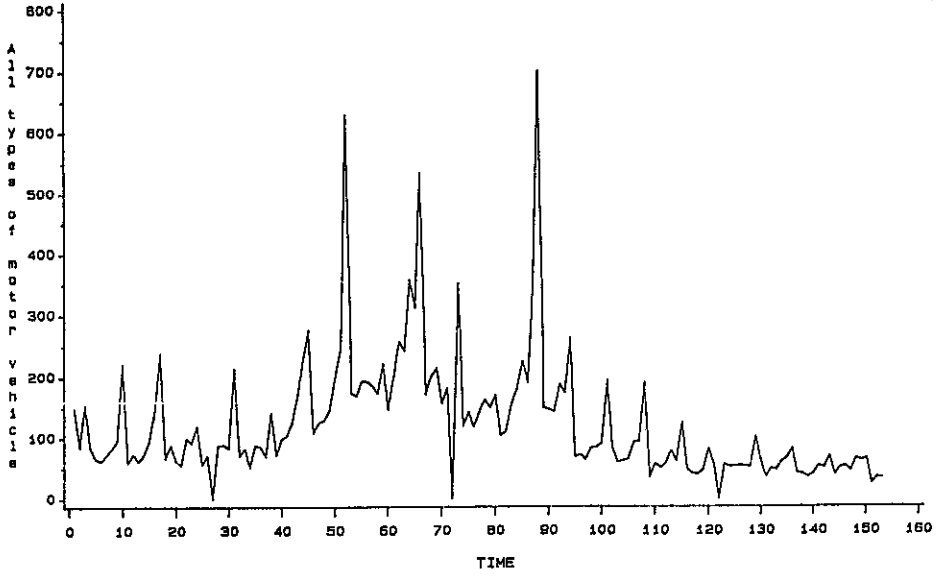


FIG 4 Class T/B, East at 20:00



Appendix B

Figs 1-8, Time series plots of traffic count data.

FIG 5 Class T/B, West at 08:00

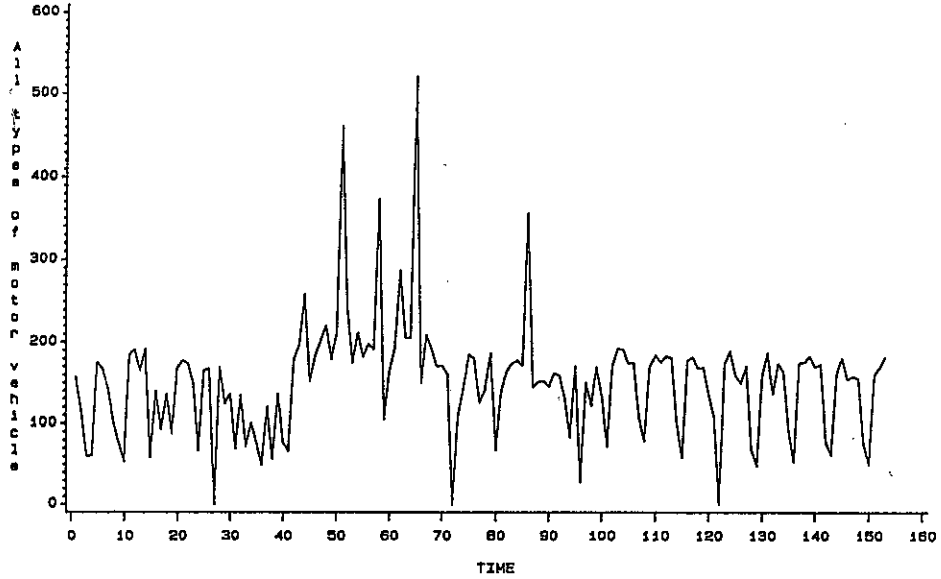


FIG 6 Class T/B, West at 12:00

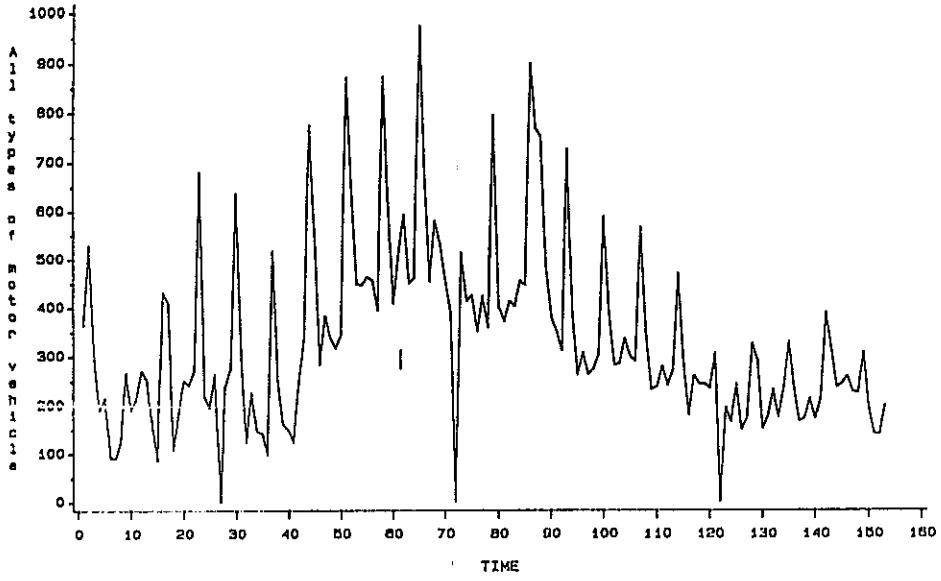


FIG 7 Class T/B, West at 17:00

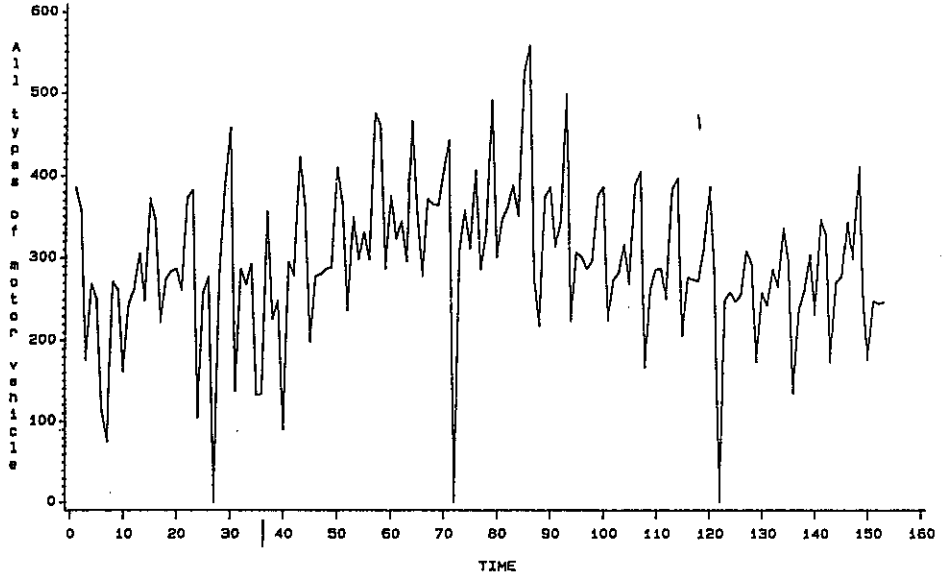


FIG 8 Class T/B, West at 20:00