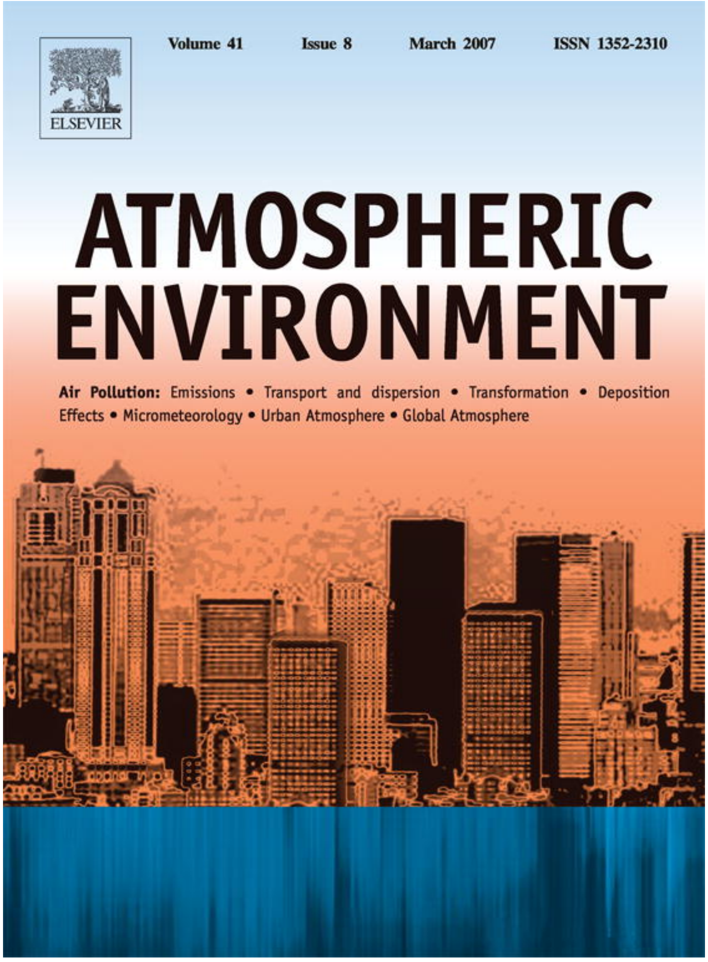
Volume 41
Issue 8
March 2007
ISSN 1352-2310
ELSEVIER
ATMOSPHERIC ENVIRONMENT
Air Pollution: Emissions • Transport and dispersion • Transformation • Deposition
Effects • Micrometeorology • Urban Atmosphere • Global Atmosphere

Available online at www.sciencedirect.com
ScienceDirect
Atmospheric Environment 41 (2007) 1784–1791
ATMOSPHERIC ENVIRONMENT
www.elsevier.com/locate/atmosenv

Investigation into the use of the CUSUM technique in identifying changes in mean air pollution levels following introduction of a traffic management scheme

Benjamin Barratt[a,*], Richard Atkinson[b], H. Ross Anderson[b], Sean Beevers[a], Frank Kelly[a], Ian Mudway[a], Paul Wilkinson[c]

[a]*Environmental Research Group, King's College London, Franklin Wilkins Building, 150 Stamford Street, London SE1 9NH, UK*
[b]*Department of Community Health Sciences, St George's, University of London, Cranmer Terrace, London SW17 0RE, UK*
[c]*London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK*

Received 29 September 2005; received in revised form 3 August 2006; accepted 8 September 2006

Abstract

There is an increasing need for statistical techniques to identify and quantify the effects of traffic management schemes on ambient pollution levels. Cumulative sum (CUSUM) charts have been used extensively in industrial process control to detect deviations in production parameters from pre-determined values. This study investigates the use of the CUSUM procedure to identify change in ambient air pollution levels following the introduction of a traffic management scheme at a specific location in Central London.

The CUSUM methods of Lucas first compute the standardised deviations of time series observations from the desired process mean. These are accumulated over time to compute the CUSUM at each time point.

Data for the analysis were taken from a kerbside monitoring site on Marylebone Road, a six lane trunk route in Central London. In August 2001 the lane adjacent to the monitoring site was designated as a permanent bus lane. The CUSUM analysis clearly identifies a sustained decrease in carbon monoxide concentrations beginning in 2002. However, seasonality and other factors precluded precise characterisation of the timing of the change. When the analysis was repeated using a reference mean that extrapolated the pre-intervention trend in carbon monoxide concentrations, the CUSUM chart no longer identified a sustained decrease.

CUSUM appears to offer a simple and rapid method for identifying sustained changes in pollution levels, but the range of confounding influences on carbon monoxide concentrations, most notably underlying trends, seasonality and independent interventions, complicate its interpretation. Its application in assessing the presence or timing of a stepped change in pollution or similar environmental time series data is recommended in its basic form only where the predicted change is large by comparison with other independent influences. The authors believe that further development of the technique beyond this initial study is worthwhile in order to improve the technique's sensitivity.

Keywords: CUSUM; Traffic management; Air pollution; Air quality management

*Corresponding author. Fax: +44 (0) 207 848 4045.
E-mail address: benjamin.barratt@erg.kcl.ac.uk (B. Barratt).

1352-2310/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2006.09.052

1. Background

There is an increasing need for statistical techniques to identify and quantify the effects of traffic management schemes on ambient pollution levels, a prime example in the UK being the introduction of the London Congestion Charging Scheme. While emissions modelling techniques can very precisely predict the likely effects of such schemes on pollution levels, it is desirable to have evidence in the form of monitoring results once a scheme is enacted. Such evidence forms the first step in an assessment of any potential health effects from traffic management schemes. However, the identification of what may be subtle changes in mean air pollution concentrations against a background of changing trends and strong seasonal patterns may not be a simple task.

The purpose of this study was to investigate the ability of the Cumulative sum (CUSUM) technique to identify a step change in pollution levels following the implementation of a relatively simple traffic management scheme in Central London. If appropriate, the technique could then be developed further for application in more complex situations such as assessment of the effects of the Congestion Charging Scheme in London.

A traffic management scheme, introduced at a single point in time, may result in a single discrete change in mean pollution levels. Cumulative sum (CUSUM) techniques, first proposed by Page (1954), were developed for use in industrial process control to detect deviations in production parameters from pre-determined values. A useful review of the statistical methods used in process quality control including CUSUM techniques is given by Ryan (2000). These techniques would seem to offer a useful approach to the identification of changes in mean pollution concentrations following a traffic management intervention—one is aiming to detect a (long-term) shift in mean pollutant concentrations from the pre-implementation concentrations.

CUSUM techniques have been used in other fields, for example, they have been used to study the performance of medical professionals, assess learning curves for trainees and the efficacy of treatments amongst others. Rossi et al. (1999) applied the CUSUM technique in an environmental epidemiological setting where they compared mortality rates in two distinct time periods in an area of Tuscany characterised by the presence of chemical plants. As far as we are aware the CUSUM technique has not been used to analyse changes in air pollution levels associated with the introduction of traffic management schemes.

In this study we use a CUSUM procedure to assess changes in ambient mean air pollution levels following the introduction of a traffic management scheme at Marylebone Road, Central London; in August 2001 a permanent bus lane was introduced adjacent to a kerbside continuous air pollution and traffic monitoring site. The primary aim of the analysis was to identify any change in pollutant concentrations. A secondary aim was to use the CUSUM technique to pinpoint the date on which the change in mean pollution levels occurred and to compare this date with the date of implementation of the scheme. We used daily average carbon monoxide concentrations since emission sources for CO are predominately motor vehicles.

2. Method

2.1. The CUSUM procedure

CUSUM methods apply to observations recorded over time (daily, weekly, monthly). The observations may be physical measurements, counts or rates and may be grouped (in production batches for example) or individual observations (e.g., as here, daily average concentrations of a pollutant at a monitoring station). The CUSUM methods of Lucas (1982) and Lucas and Crosier (1982) as applied to individual observations first compute the standardised deviations of observations from the desired process mean:

$$z_i = \frac{x_i - \bar{x}}{\hat{\sigma}_x}, \tag{1}$$

where x_i is the observed value at time i, $\bar{x}$ is the desired process mean and $\hat{\sigma}_x$ is an estimate of the standard deviation of the observed values. These are accumulated over time to compute the CUSUM, S, at each time point i as follows:

$$S_i = S_{i-1} + z_i, \quad \text{where } S_0 = 0. \tag{2}$$

Thus, if there is a shift in the process mean away from the target then z_i will tend to be larger or smaller than the target average and the CUSUM will steadily increase or decrease. Depending upon the magnitude of the shift in the mean the CUSUM may not detect the change immediately, requiring a number of observations at the new level before it begins to pick up the change in mean. An example of these computations are given in Table 1 and illustrated in Fig. 1. Forty data points were randomly generated from a normal distribution, mean = 0, standard deviation = 1 and after observation number 10 the mean was increased to 0.5. The table shows the observed values, x_i, the z_i and cumulative sum S_i. The observed values are plotted in Fig. 1.

Lucas (Lucas, 1982) proposed a pair of cumulative sums where the first is for detecting mean increases (S_{Hi}) and the second is for detecting mean decreases (S_{Li}) defined as

$$S_{Hi} = \max[0, (z_i - k) + S_{Hi-1}] \quad \text{and}$$
$$S_{Li} = \min[0, (z_i + k) + S_{Li} - 1]. \tag{3}$$

The parameter, k, is the reference value or allowable slack in the process and is usually set to be one half of the mean shift (in z units) one wishes to detect, i.e.

$$k = \frac{1}{2}\Delta, \quad \text{where } \Delta = \frac{x - \bar{x}}{\hat{\sigma}_x}. \tag{4}$$

The usual choice of $k = 0.5$, is therefore the appropriate choice for detecting a 1-sigma shift in the process mean.

The CUSUM scheme usually includes confidence limits, $\pm h\sigma_x$ with $h = 4$ or 5 to indicate when a change in the process mean has been detected, the actual change occurring some time earlier.

The values of S_{Hi} and S_{Li} are also given in Table 1 and are plotted in Fig. 2. This figure shows the start of an upward trend from about observation number 13. The CUSUM crosses the upper CUSUM limit at observation number 29. The limits are used as a means of differentiating between random changes in mean values and sustained changes in mean values. It is evident that the control chart does not immediately detect the change in the mean but

Table 1
Example generated data observations with calculated CUSUM parameters

Obs	x_i	z_i	S_i	S_{Hi}	S_{Li}
1	−0.70	−0.70	−0.70	0.00	−0.20
2	−0.71	−0.71	−1.41	0.00	−0.41
3	0.20	0.20	−1.21	0.00	0.00
4	0.25	0.25	−0.96	0.00	0.00
5	0.58	0.58	−0.38	0.08	0.00
6	−0.45	−0.45	−0.83	0.00	0.00
7	−0.39	−0.39	−1.22	0.00	0.00
8	1.13	1.13	−0.09	0.63	0.00
9	−0.80	−0.80	−0.89	0.00	−0.30
10	0.48	0.48	−0.42	0.00	0.00
11	0.09	0.09	−0.33	0.00	0.00
12	2.25	2.25	1.93	1.75	0.00
13	−0.36	−0.36	1.57	0.90	0.00
14	−0.04	−0.04	1.53	0.36	0.00
15	1.02	1.02	2.55	0.88	0.00
16	0.14	0.14	2.69	0.52	0.00
17	1.36	1.36	4.05	1.38	0.00
18	−0.11	−0.11	3.94	0.77	0.00
19	−0.41	−0.41	3.53	0.00	0.00
20	1.85	1.85	5.38	1.35	0.00
21	2.27	2.27	7.66	3.12	0.00
22	0.00	0.00	7.65	2.62	0.00
23	−0.49	−0.49	7.16	1.63	0.00
24	0.05	0.05	7.21	1.17	0.00
25	1.66	1.66	8.87	2.33	0.00
26	0.50	0.50	9.37	2.33	0.00
27	1.56	1.56	10.92	3.39	0.00
28	2.08	2.08	13.00	4.97	0.00
29	−1.36	−1.36	11.64	3.10	−0.86
30	1.31	1.31	12.94	3.91	0.00
31	−0.41	−0.41	12.53	2.99	0.00
32	−0.10	−0.10	12.43	2.40	0.00
33	0.26	0.26	12.69	2.15	0.00
34	1.18	1.18	13.86	2.83	0.00
35	1.01	1.01	14.87	3.34	0.00
36	0.51	0.51	15.38	3.35	0.00
37	2.63	2.63	18.01	5.48	0.00
38	1.79	1.79	19.80	6.76	0.00
39	−1.15	−1.15	18.65	5.11	−0.65
40	1.52	1.52	20.17	6.13	0.00

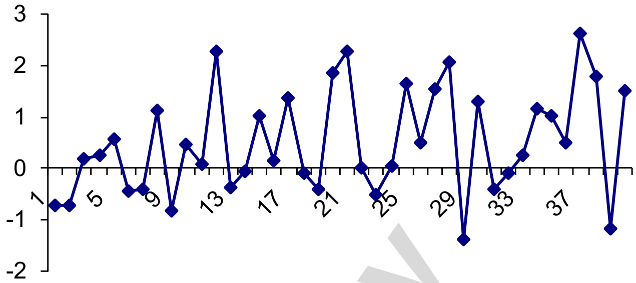

Fig. 1. Data points generated from a Normal Distribution mean 0 and standard deviation 1 (points 1–10) and from a Normal Distribution mean 0.5 and standard deviation 1 (points 11–40).

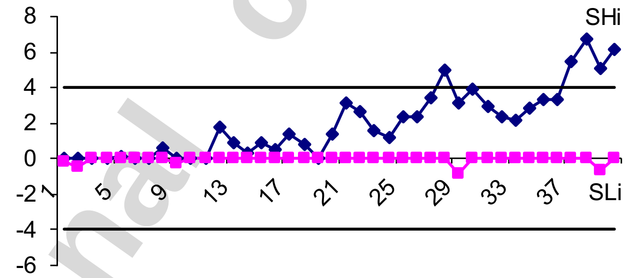

Fig. 2. CUSUM chart for S_{Hi} and S_{Li} for data in Table 1.

requires a run of values at the new process mean to indicate that the change has taken place.

Certain parameters must be considered when applying the technique: length of data set, selection of reference mean, averaging period (daily mean, monthly mean, etc.) and the 'slackness' factor, k. In its original application in quality control, the method is applied to a data set until the process is identified as out of control, then the process is halted, re-adjusted and re-started. In application to environmental time series data, the data set must be sufficiently long to differentiate between transient or periodic changes and sustained step changes. Long datasets are more able to illustrate sustained changes, but are more susceptible to underlying trends and further interventions independent of that under analysis.

2.2. Application to air pollution data

To test the use of the CUSUM method, we applied it to the detection of change in monitored concentrations of carbon monoxide following the introduction of a bus lane on Marylebone Road, a congested six lane east/west trunk route in Central London. The nearside lane in each direction was designated a bus lane on 18 August 2001, after which only buses and taxis have been allowed to enter the lane (Green, 2002). The bus lane is permanently in operation 24 h a day, 7 days a week and is strictly observed due to a system of enforcement cameras and automatic fines.

We selected carbon monoxide (CO) as an indicator of traffic related pollution as it has relatively straightforward atmospheric chemistry, low background concentrations and a strong traffic-related signal at the kerbside. Measurements of CO levels were available from a continuous monitoring site located on the southern side of the road with the carbon monoxide sampling inlet approximately 1 m from the kerb. The monitoring site is part of the UK Government's Automatic Urban and Rural Network (AURN) and is operated to defined QA/QC standards. All data used in this study are fully ratified. The carbon monoxide valid capture rate for the period of study was 96%, equating to 1056 valid days (>75% 15 min mean capture rate per day) out of a possible 1096. The measure used in this analysis was the daily mean concentration measured in $mg\,m^{-3}$.

In addition, traffic count data were obtained from an induction loop system fixed within the road surface adjacent to the monitoring site. The system monitors vehicle number, type (by axle length) and speed for each of the six lanes.

3. Results

3.1. Analysis of change following introduction of the bus lane

Figs. 3 and 4 show time series data between 1 January 2000 and 31 December 2002 for the daily total number of vehicles travelling past the vehicle

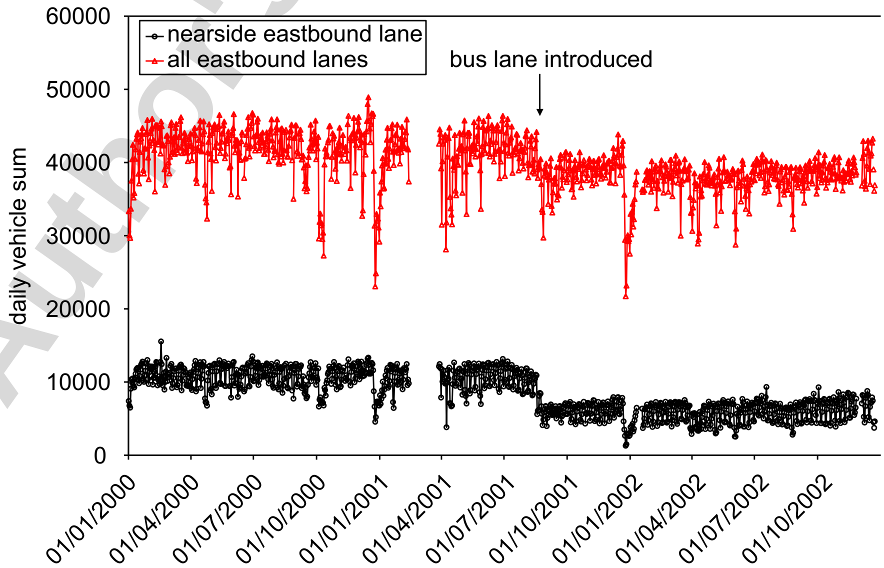

Fig. 3. Time series plot of daily sum vehicle count measured at Marylebone Road between 1 January 2000 and 31 December 2002.

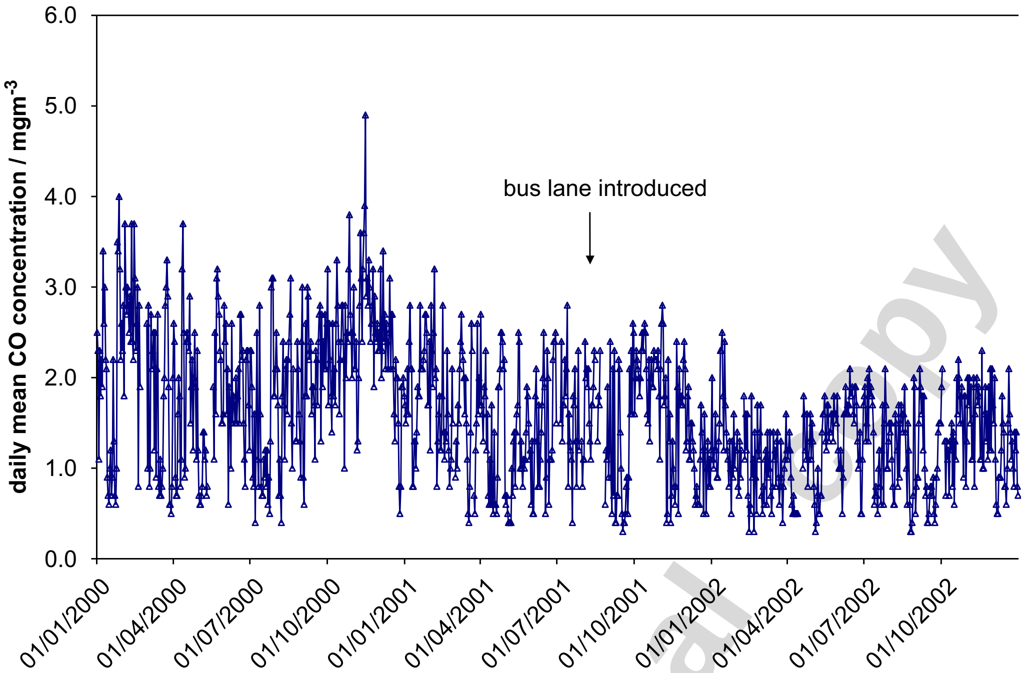

Fig. 4. Time series plot of daily mean CO measured at Marylebone Road between 1 January 2000 and 31 December 2002.

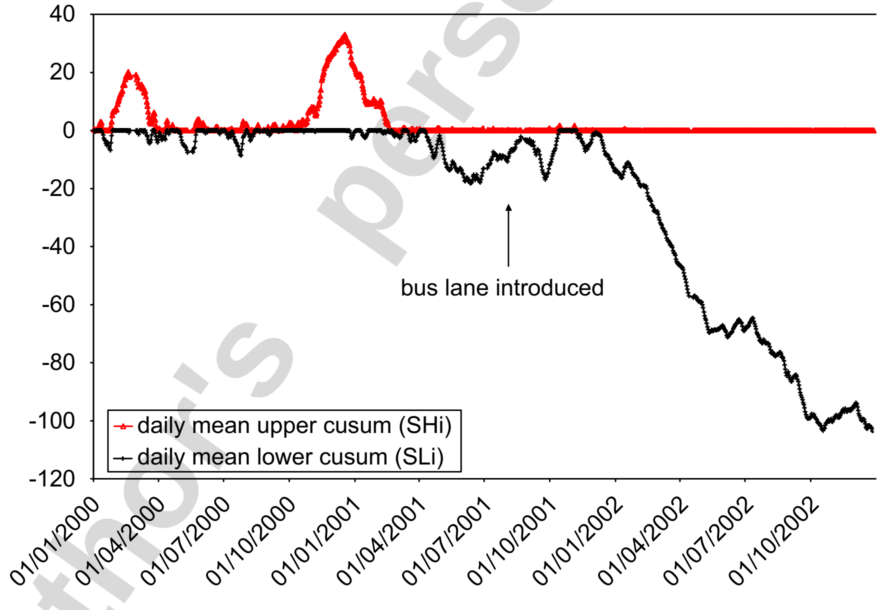

Fig. 5. CUSUM chart for S_{Hi} and S_{Li} for CO data at Marylebone Road.

monitors and the daily mean CO concentration measured at the kerbside monitor. Fig. 3 clearly shows a fall in the number of vehicles using the nearside lane following the introduction of the bus lane. Fig. 4 illustrates two main features of the CO measurements. First, the daily mean concentrations are decreasing throughout the 3-year period and secondly, the data exhibit considerable seasonality—concentrations increasing in the winter months—a feature particularly evident in later months of 2000.

For the CUSUM calculations (Fig. 5), the reference mean ($\bar{x} = 2.13\,\mathrm{mg\,m^{-3}}$) and standard deviation ($\hat{\sigma}_x = 0.92\,\mathrm{mg\,m^{-3}}$) were calculated from the pre-implementation period 1st January 2000 to 17th August 2001. The chart does identify a decrease in carbon monoxide concentrations with a sustained divergence of the lower CUSUM curve from the beginning of 2002, approximately 6 months after introduction of the bus lane. However, the CUSUM curves are clearly influenced by the strong seasonal nature of the CO data with clear peaks during the winter months.

3.2. Accounting for background trends in co concentrations

In order to try and differentiate between any decreases in the mean CO concentrations that might be attributable to the introduction of the traffic control measure from those occurring generally across London we plotted time-centred running annual mean CO concentrations at monthly increments at five Inner London roadside sites not subject to local traffic management schemes during the period 1999–2003 (Fig. 6). A clear linear decrease in pollutant concentrations was observed. A simple regression model calculated this decrease as $-0.368\,\mu g\,m^{-3}\,day^{-1}$. As there was no reason to assume that the background changes in CO levels did not also apply to Marylebone Road we adjusted the reference mean ($\bar{x}$) used in the computation of the CUSUMs to take account of this trend. The resulting CUSUM chart is shown in Fig. 7. It illustrates the fall in the mean concentrations shown in Fig. 5 but it is not sustained. This failure of the CUSUM procedure to identify a sustained change in CO concentrations may be an indication that the reduction in roadside CO concentrations is either not maintained in the long term (and therefore may not be related to the intervention) or insufficiently large to be picked up by the procedure with the input parameters used. Alternatively, seasonal variation in CO concentrations and/or inappropriate adjustment for background trends has obscured a sustained change in concentrations.

The CUSUM procedure assumes independent normally distributed data, whereas air pollution measurements tend to have a skewed distribution and has a high degree of autocorrelation between daily mean values. The CO data did show evidence of skewness and strong autocorrelation at lag 1. To assess the sensitivity of the results to this modest departure from normality the analysis was repeated using log transformed daily CO concentrations. There was little change in the CUSUM chart and therefore no change in the conclusions from the original analyses (data not shown). In order to assess the impact of autocorrelation the data were condensed to weekly averages and the CUSUM analysis repeated. This batch processing did have the effect of producing a lower standard deviation

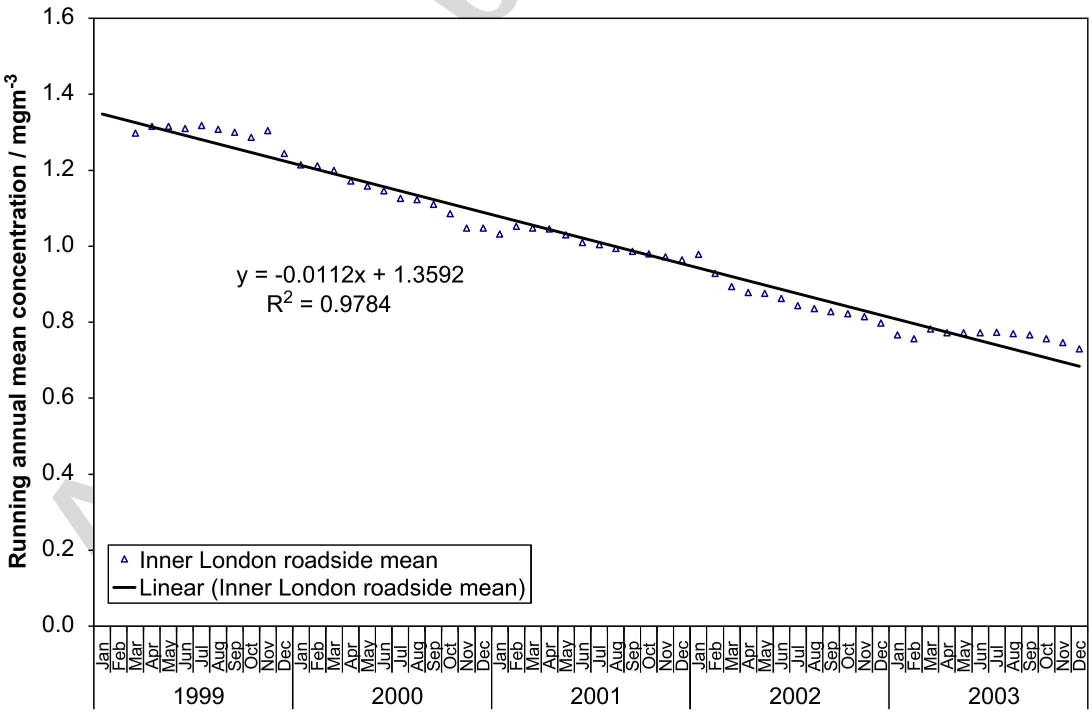

Fig. 6. Monthly increment running annual mean CO concentrations at five Inner London Roadside sites (not including Marylebone Road).

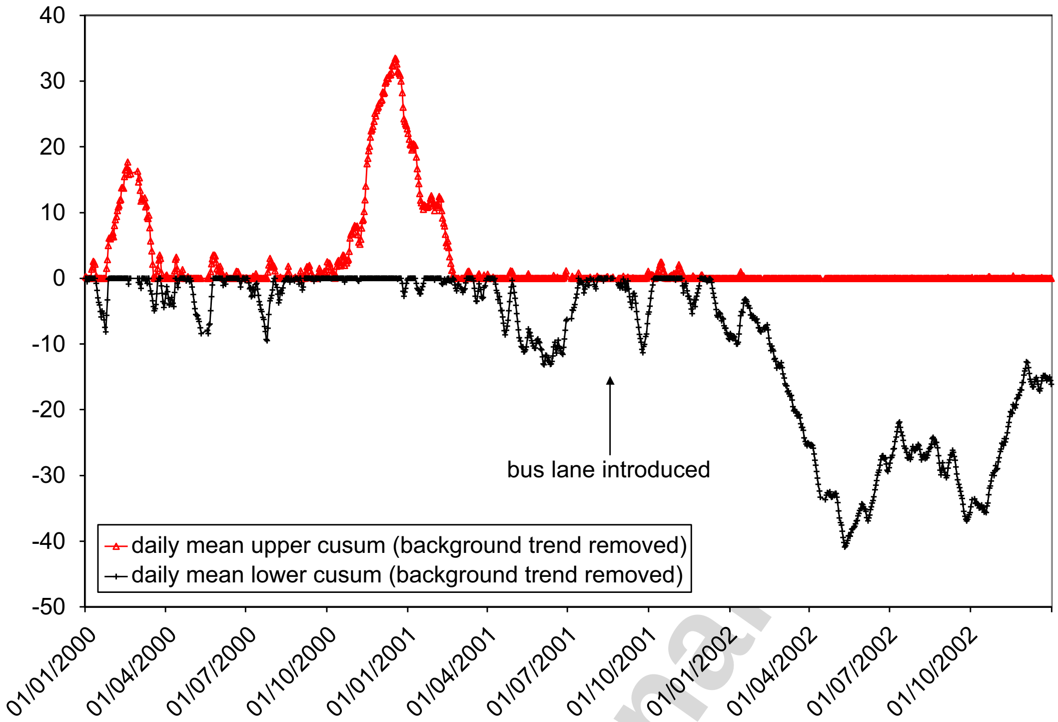

Fig. 7. CUSUM chart for $S_{\mathrm{H}i}$ and $S_{\mathrm{L}i}$ for CO data at Marylebone Road adjusted for background trends.

and a smoother CUSUM trace, clarifying the results of the daily mean analysis (data not shown).

4. Discussion

In this study we attempted to apply a statistical technique normally used in quality control processes to identify sustained change in ambient pollution levels caused by a traffic management intervention. In its basic form, it appears to be a simple method to identify subtle but sustained step changes in pollution levels, but the range of confounding influences on concentrations, most notably underlying trends, meteorological conditions including seasonality and independent interventions, complicate its interpretation and act to obscure the identification of change points. In the basic form presented here, the CUSUM procedure was unable to differentiate between these confounding influences and the target intervention.

The detection of small changes in pollution may therefore require modification of the technique to allow adjustment for time-varying influences on pollution levels. However, a major advantage of the CUSUM technique is that it can be applied to time series data very quickly using simple equations, and the need for modelling to account for confounding factors dilutes this advantage, potentially to the point where other modelling methods may become more suitable in many circumstances.

Despite the complexities of applying the technique to air pollution series, CUSUM can provide useful evidence about change of concentrations from pre-existing levels. Its interpretation is most secure when detecting a change in pollution that is large by comparison with seasonal fluctuations and the shifts of long-term trends. In the more usual circumstance of studying comparatively subtle change in pollution, the secure interpretation of CUSUM requires adaptation of the technique, specifically to use estimates of the standard deviation and mean of pollution levels that take proper account of the underlying temporal variation in them. How this is most efficiently implemented will require further testing, ideally using data from a range of different settings.

We conclude that further development in the technique beyond this initial study is worthwhile, particularly to address issues relating to the setting of the control parameters and the methods of data pre-processing (modelling). Its application in assessing the presence or timing of a stepped change in pollution or similar environmental time series data is recommended in its basic form only where the predicted change is large by comparison with other independent influences.

References

Green, D., 2002. Marylebone Road ('Supersite') Annual Report 2001. King's College, London Environmental Research Group.
Lucas, J.M., 1982. Combined Shewhart–CUSUM quality control schemes. Journal of Quality Control 14, 51–59.
Lucas, J.M., Crosier, R.B., 1982. Fast initial response for CUSUM quality control schemes: Give your CUSUM a head start. Technometrics 24, 199–205.
Page, E.S., 1954. Continuous inspection scheme. Biometrika 41, 100–115.
Ryan, T.P., 2000. Statistical Methods for Quality Improvement, second ed. Wiley, New York.
Rossi, G., Lampugnani, L., Marchi, M., 1999. An approximate CUSUM procedure for surveillance of health events. Statistics Medicine 18, 2111–2122.