

Modelling Daily Multivariate Pollutant Data at Multiple Sites

Gavin Shaddick

Department of Mathematical Sciences, University of Bath, UK.

Small Area Health Statistics Unit, Department of Epidemiology and Public Health, Imperial College School of Medicine at St. Mary's, London, UK.

Jon Wakefield

Departments of Statistics and Biostatistics, University of Washington, Seattle, USA.

Small Area Health Statistics Unit, Department of Epidemiology and Public Health, Imperial College School of Medicine at St. Mary's, London, UK.

Summary. This paper considers the spatio-temporal modelling of four pollutants measured daily at eight monitoring sites in London over a four-year period. Such multiple pollutant datasets measured over time at multiple sites within a region of interest are typical. Here, the modelling was carried out in order to provide the exposure for a study investigating the health effects of air pollution. Alternative objectives include the design problem of the positioning of a new monitoring site, or for regulatory purposes in order to determine whether environmental standards are being met. In general, analyses are hampered by missing data due, for example, to a particular pollutant not being measured at a site, a monitor being inactive by design (for example, a six-day monitoring schedule), or because of an unreliable or faulty monitor. Data of this type is modelled here within a dynamic linear modelling framework, in which the dependencies across time, space and pollutants are exploited. Throughout the approach is Bayesian, with implementation via Markov Chain Monte Carlo.

Keywords: Dynamic linear models; Environmental statistics; Hierarchical models; Isotropy; Spatial modelling; Stationarity.

1. Introduction

The potential effects of ambient air pollution on health is a major issue in public health, and the subject has received a great deal of attention in recent years. A number of epidemiological studies have found consistent associations between daily levels of pollution and adverse health effects for both mortality and morbidity. For summaries see Pope et al. (1995) and COMEAP (1995).

While the long-term (or chronic) effects of air pollution are also of interest, the great majority of studies consider short-term, or acute, effects, due to data availability. While increases in adverse health events are consistently reported, it is unclear as to the actual pollutant(s) responsible, and the size of the effect. In conducting such time series studies to investigate the relationship between air pollution and a health outcome, for example, respiratory mortality, it is important to have a good measure of the level of pollution on each of the study days. Often daily measurements are available from a number of monitoring sites across the study area. Each of these monitors may measure different sets of pollutants, there may be periods of missing data, and each of the recorded measurements is subject to error.

In studies investigating the health effects of daily changes in air pollution the exposures are essentially treated as constant across the study area. Many studies are carried out in urban areas where there may be more than one monitoring site. In this case a daily average of readings from all sites is often used (for example, Schwartz and Dockery (1992); Schwartz (1993)). The use of such an average, or the readings from just one site, may be criticised on the grounds that pollution levels vary within an urban environment. There may be spatial differences due, for example, to local traffic conditions and point sources of pollution, and the ability of the pollutants to disperse because of the surrounding buildings.

In this paper a (hierarchical) dynamic linear model (DLM) is suggested for the analysis. Such DLMs are described in, for example, Pole et al. (1994) and Gamerman and Migon (1993), who specifically consider hierarchical DLMs. At stage one of this hierarchy, each of the individual pollutants on any day are modelled as a function of the true underlying level, corrupted by measurement error. These true underlying levels are assumed to have structure in both space and time, and this structure is

modelled at stage two of the hierarchy along with the relationship between pollutants. A Bayesian approach is adopted throughout, with prior distributions being assigned to the unknown parameters and missing observations at stages one and two. Essentially, the multiple pollutants at any time point are modelled as arising from a multivariate Gaussian random field. This model addresses each of the inadequacies described in the observed data and, specifically, allows information from multiple sites on different pollutants to be combined, in order to provide an accurate level of pollution at each of the observed sites, or at locations previously unmeasured, though the latter is more dependent on modelling assumptions. At any location, monitored or otherwise, a measure of the uncertainty associated with the level can also be obtained. The latter is particularly useful for accounting for the variability in the pollution level, formally via errors-in-variables modelling (e.g., Carroll et al. (1995)), or informally when regression coefficients describing the relationship between health risk and pollution are interpreted.

The Data

This approach was developed to model pollution data collected at eight monitoring sites within London measuring the pollutants particulate matter (PM_{10}), carbon monoxide (CO), nitrogen oxide (NO) and sulphur dioxide (SO_2) over the period 1994–1997. Table 1 summarises the periods of operation, and pollutants measured at each of the sites. It is noted that all four pollutants were measured at four sites only, the periods of operation vary between one and four years, and the percentage of missing values within these periods of operation can be large. For example, 37% of the daily measurements for CO were missing at Hillingdon. Readings of zero for any of the pollutants were considered to be missing values, rather than implausibly low readings. Figure 1 shows time series plots for PM_{10} at all eight sites, and clearly shows the sparsity of data in the first two years. There is no apparent trend in level over the time scale of the study.

Pollutant Dependence

The model exploits the dependencies that exist between pollutants, temporally and spatially, due to the processes by which they are formed. For example, a major source of particulate matter (PM_{10}), and of the other pollutants considered here are combustion processes and, in particular, diesel combustion. Carbon monoxide is a toxic gas emitted as a result of combustion processes which, in urban areas, is almost entirely produced from road traffic emissions, as are oxides of nitrogen, NO_x . Sulphur dioxide, SO_2 , a corrosive acid gas, is primarily caused by power stations burning fossil fuels which contain sulphur

Associations between the levels of pollutants are also observed because of their relationship with meteorological conditions, such as wind direction and speed and temperature. In periods of higher temperature and during temperature ‘inversions’ (where primary pollutants to be trapped near the ground because the usual decrease in temperature with altitude, which causes hot pollutant gases to rise, is absent) leading to enhanced production of secondary pollutants (Onursal and Gautan (1997)). Dependencies with temperature and certain pollutants will also be due, in part, to changing use of heating materials. Table 2 shows the sample correlations between the pollutants (logged values) and temperature, at the Bloomsbury site. It is noted that the associations between the pollutants appear relatively constant across sites (not shown), as do the associations with temperature.

Temporal Dependence

The pollutants considered here have varying atmospheric lifetimes, but all can linger in the environment. For example, nitrogen oxides have a lifetime of approximately one day before being converted to nitric acid, whereas CO can survive in the atmosphere for up to a month, before eventually oxidising to carbon dioxide. The atmospheric lifetime of particulate matter is strongly related to particle size, but may be as long as 10 days for particles of about 1mm in diameter. From these considerations strong daily dependence would be expected between pollutant measures on consecutive days.

Spatial Dependence

There is very high correlation between the readings for each of the pollutants from different monitoring sites. Figure 4(a) shows the position of the eight sites included in the study, and the correlations between the logged measurements of PM_{10} measured at the eight sites are presented in Table 3. High correlations between all pairs of measurements are observed at all sites. In Figure 2 the correlation between daily measurements of each (log) pollutant and the distance between the site at which they were measured, is plotted, for two different time periods. As expected, measurements from sites in close proximity to each other are, in general, more highly correlated than those further away, and this relationship appears relatively constant across time periods and pollutants.

Measurement error

Data from continuous monitoring can generally be assumed to have a relatively small measurement error if, as is often the case, the sites are subject to strict quality assurance and validation procedures (as are the London data considered here). The measurements of PM_{10} from different sample based monitoring techniques can be affected by several factors, such as operating temperature and filter media and history. Determination of the accuracy and precision of any given concentration is, therefore, liable to encompass a margin of error, the target figure in the data quality standards is a precision of $< 5\mu\text{gm}^{-3}$ for concentrations of less than $100\mu\text{gm}^{-3}$ (DETR (1998)). For SO_2 , the accuracy depends on the accuracy of the calibration standards and analyser stability, but based on long-term comparisons is estimated at about $\pm 1\%$ (DETR (1998)).

Spatio-Temporal Modelling

The modelling of environmental variables in time and space has a considerable literature and only key papers are highlighted here. In an early Bayesian application, Handcock and Wallis (1994) consider the spatio-temporal modelling of winter temperature data but their approach was to carry out separate spatial analyses in each year using a Gaussian random field model. Guttorp et al. (1994) modelled the spatial covariances of hourly ozone levels using the Sampson and Guttorp (1992) spatial covariance approach, and allowed the parameters of the model to vary as a function of time of day. Huang and Cressie (1996) modelled snow water in time and space using a separable dynamic model. A spatio-temporal model for hourly ozone measurements was developed by Carroll et al. (1997). The model combined trend terms incorporating temperature and hourly/monthly effects, and an error model in which the correlation in the residuals was a nonlinear function of time and space, in particular the spatial structure was a function of the lag between observations. Unfortunately, as Cressie (1997) pointed out, this correlation function is not positive definite. Mardia et al. (1998) propose what they term a ‘kriged Kalman filter’ and outline a likelihood-based estimation strategy. A more general model and an (approximate) Bayesian estimation approach appears in Wilke and Cressie (1999), with a fully Bayesian version being described in Wilke et al. (1998). Brown et al. (2001) considered the spatio-temporal modelling of rainfall data, using a non-separable model in which the spatial field at a specific time is obtained by ‘blurring’ the field at the previous time point. Recently, independent research by Tonellato (2001) presented a similar modelling approach to that presented here, using a univariate auto-regressive process measured with error, for hourly measurements of a single pollutant (CO) from a small number of sites.

The structure of this paper is as follows. In Section 2 the general spatio-temporal model for multiple pollutants is described and in Section 3 a number of simplified models are considered in order to examine different aspects of the data, in particular to examine model fit. Section 4 describes the analysis of the London data using the full model, and Section 5 contains a concluding discussion.

2. The Model

2.1. Specification

The general model allows for a temporal-pollutant interaction, and a spatial-pollutant interaction, with the spatial model being constant across time, isotropic and stationary. The hierarchical dynamic linear model is described in three stages.

Stage One, Observed Data Model:

Let Y_{spt} denote the observed level of pollutant p at spatial location s on day t and assume

$$Y_{spt} = X_{spt}\beta + \theta_{pt} + m_s + v_{spt}, \quad (1)$$

In this model:

- v_{spt} represents the *measurement errors* which are assumed i.i.d. $N(0, \sigma_{sp}^2)$.
- m_s represents the spatial effect of being at site s .
- θ_{pt} is the error term that will induce temporal and pollutant dependence at Stage 2.
- The $q \times 1$ parameter vector β is a vector of regression coefficients.

- X_{spt} represents a $1 \times q$ vector of regressors that may change temporally (for example, temperature), and spatially. The latter may represent, for example, spatial characteristics of the site that may be constant across time such as latitude and longitude (which could be used to remove any trend), or characteristics of the monitor, for example roadside or elevation. The subscript p allows these effects to be pollutant specific.

At the first stage, the daily ($t = 1, \dots, T$) measurements of each pollutant ($p = 1, \dots, P$) at each monitoring site ($s = 1, \dots, S$) are modelled as a function of the true underlying level of the pollutant with a site adjustment and a pollutant-site specific error term.

Stage Two (a), Spatial/Pollutant Model:

Assuming that the collection of random effects $m_p = (m_{p1}, \dots, m_{pS})'$, $p = 1, \dots, P$, arise from the multivariate normal distribution

$$m_p \sim MVN(0_S, \sigma_{pm}^2 \Sigma_{pm}), \quad (2)$$

where 0_S is an $S \times 1$ vector of zeros, σ_{pm}^2 the between-site variance for pollutant p and Σ_{pm} is the $S \times S$ correlation matrix, in which element (s, s') represents the correlation between sites s and s' , $s, s' = 1, \dots, S$, for pollutant p . This model is stationary, a point discussed later in more detail, and an isotropic covariance model is assumed in which the correlation between sites s and s' is assumed to be a function of the distance between them, i.e. $f(d_{ss'}, \phi_p)$, where $d_{ss'}$ represents the distance between sites s and s' (in km). Such models are frequently used in geostatistics (e.g. Cressie (1993)), and the form of the relationship assumed here is,

$$f(d_{ss'}, \phi_p) = \exp(-\phi_p d_{ss'}) \quad (3)$$

where $\phi_p > 0$ describes the strength of the correlation. Note that the log correlation is linear in distance for this model and so the appropriateness of this form may be assessed using plots such as Figure 2. The use of a stationary and isotropic model with a single parameter is restrictive, but with only eight sites it is difficult to specify a more general model. In London, this assumption is likely to be more realistic than in other locations, since the meteorology and topography are relatively spatially stable. For more remarks concerning the specification of a spatial model see the discussion of Diggle et al. (1998). With a large enough number of sites, a two-parameter isotropic spatial with the desirable properties of the Matérn class could be used (e.g. (Handcock and Wallis (1994))).

A simpler model which is clearly unrealistic but is used for comparison is given by:

$$m_{ps} \sim \text{i.i.d } N(0, \sigma_{pm}^2), \quad (4)$$

This model assumes that the site-specific levels are (conditionally) independent.

Stage Two (b), Temporal/Pollutant Model:

$$\theta_{pt} = \theta_{p,t-1} + w_{pt}, \quad (5)$$

for $p = 1, \dots, P$. Here $w_t = (w_{1t}, \dots, w_{Pt})'$ are i.i.d. multivariate normal random variables with zero mean and variance-covariance matrix Σ_P . This matrix contains variances σ_{wp}^2 thus allowing different pollutants to have different amounts of temporal dependence, and $P(P-1)/2$ covariance terms reflecting the dependence (more precisely the covariance) between each of the pollutants, conditional on the previous day's values.

This stage represents a first order smoothing model (e.g West et al. (1985); Fahrmeir and Knorr-Held (2000)) with the true levels on day t modelled as a function of those on the previous day. The model is a limiting form of the autoregressive first order model and provides a non-stationary temporal model. Such an approach has been widely used (see for example, (Pole et al. (1994))). In terms of DLMS, (1) is known as the *observation equation*, (5) is the *system equation*, and θ_{pt} , the *state*.

Stage Three, Hyperprior:

A normal prior $N(c, C)$ is assumed for β , where c is a $q \times 1$ vector and C a $q \times q$ variance-covariance matrix. Gamma priors are specified for the precisions, specifically $\sigma_{sp}^{-2} \sim Ga(a_v, b_v)$ (and for simple

univariate updating $\sigma_w^{-2} \sim Ga(a_w, b_w)$. These distributions are parameterised so that, for example, $E[\sigma_v^{-2}] = a_v/b_v$ and $\text{var}(\sigma_v^{-2}) = a_v/b_v^2$. In the case of multivariate updating, the variance-covariance matrix, $\Sigma_P^{-1} \sim W_P(D, d)$, where $W_P(D, d)$ denotes a P -dimensional Wishart distribution with mean D and precision parameter d . The precision is chosen to be $d = P$ which corresponds to the flattest ‘proper’ distribution. Given that the modelling is performed on the logarithmic scale, the variances σ_{wp}^2 are approximately equal to the coefficient of variation of the underlying states on the original scale (which is a more natural scale on which to specify priors). For the mean therefore values are chosen such that the diagonals of the expected value (D/d) relate to the coefficient of variation that might be expected. The off-diagonal elements of D/d are chosen to reflect the expected correlations between the pollution-specific states.

Unless there is specific information to the contrary, i.e. that a monitor with different characteristics is used at a particular site, it is assumed $\sigma_{vs}^{-2} \sim Ga(a_v, b_v)$, $s = 1, \dots, S$. A uniform prior is used for ϕ_p , with the limits being based on beliefs about the relationship between correlation and distance. For example, the distance, d , at which the correlation, ρ , between two sites might be expected to fall to a particular level would be $d = -\log(\rho)/\phi_p$.

Berger et al. (2000) provide an interesting discussion of the choice of priors in spatial models, and in particular show that the improper uniform prior on the positive real line leads to an improper posterior distribution. There is a lack of identifiability with this model since a constant pollution surface is consistent with both a non-zero mean process with zero correlation, and a zero-mean process with correlations of one, and the data alone cannot distinguish between these possibilities.

The assumptions of the model include the following:

- The measurement error variance σ_{sp}^2 does not depend on time. The model is easily extendable to situations in which the measurement error may change as a function of t , for example, when a monitor is replaced.
- The relationship between the pollutants is constant over time.
- The relationship between the pollutants is spatially constant.
- The temporal and spatial components are independent.

Discussing this last point in more detail, for notational convenience, the dependence on p is suppressed and a generic pollutant considered. If data is missing at a site s' and at a time t' , these will be ‘filled in’ using $m_{s'}$ and $\theta_{t'}$, where the former component does not change with time. In the spatio-temporal literature, *separable* models are often considered, these impose a particular type of independence between space and time components. Let $\rho_{s'}$ denote the correlation between observations Y_s and $Y_{s+s'}$, $\rho_{tt'}$ the correlation between Y_t and $Y_{t+t'}$, and $\rho_{s't'}$ the correlation between $Y_{s+s',t+t'}$ and $Y_{s,t}$. Then the correlation of a stationary model is separable if

$$\rho_{s't'} = \rho_{s'} \times \rho_{t'}.$$

The model presented here exhibits a different kind of independence, though since it is non-stationary, the marginal distribution of Y cannot be evaluated. Consider the model

$$\theta_t = \alpha\theta_{t-1} + w_t,$$

with $|\alpha| < 1$. Combining this stationary time series model with the spatial model gives

$$\rho_{s't'} = q \times \rho_{s'} + (1 - q) \times \rho_{t'},$$

where $q = \sigma_m^2 / (\sigma_m^2 + \sigma_w^2)$, $\rho_{t'} = \alpha^{|t'|} / (1 - \alpha^2)$, illustrating that the ‘joint’ correlation is a weighted sum of the spatial and temporal components, with the weights being independent of t and s . The model presented here is obtained with $\alpha = 1$ as the limiting form of this model.

The temporal model is now discussed in more detail, for ease of exposition, again consider a generic pollutant. From a Bayesian perspective, the second stage may be viewed as a prior distribution for

$\theta' = (\theta_1, \dots, \theta_T)$, with

$$\begin{aligned} p(\theta|\sigma_w^2) &\propto \prod_{t=2}^T p(\theta_t|\theta_{t-1}, \sigma_w^2) \\ &\propto \exp\left\{-\frac{1}{2\sigma_w^2} \sum_{t=2}^T (\theta_t - \theta_{t-1})^2\right\} \\ &\propto \exp\left\{-\frac{1}{2\sigma_w^2} \sum_{t=1}^T n_t \theta_t (\theta_t - \bar{\theta}_t)^2\right\} \end{aligned} \quad (6)$$

where n_t indicates the number of, and $\bar{\theta}$ the mean of, the neighbours of θ_t , i.e. θ_{t-1} and θ_{t+1} . The (autoregressive) prior distribution for θ used in (6), $p(\theta|\sigma_w^2)$, can therefore be expressed as

$$p(\theta_t|\theta_{-t}, \sigma_w^2) \sim \begin{cases} N(\theta_{t+1}, \sigma_w^2) & \text{for } t = 1, \\ N\left(\frac{\theta_{t-1} + \theta_{t+1}}{2}, \frac{\sigma_w^2}{2}\right) & \text{for } t = 2, \dots, T-1, \\ N(\theta_{t-1}, \sigma_w^2) & \text{for } t = T. \end{cases} \quad (7)$$

where θ_{-t} represents the vector of θ 's with θ_t removed. It is noted that σ_w^2 is a *conditional* variance and so it is not comparable to σ_v^2 . The joint distribution is improper and only expresses prior beliefs about differences in levels on neighbouring days, and no moments exist for the state process. Due to this impropriety, an intercept is not specified in (1). An intercept may be incorporated if an additional constraint is imposed, for example, $\sum \theta_t = 0$. Letting $y = (y_1, \dots, y_T)'$, the distribution $p(\theta|y)$ exists, however, and may be evaluated using the moments of the collection of $\theta_1, \dots, \theta_T$ empirically.

2.2. Inference

The posterior distribution is given by

$$p(\theta, \beta_1, \sigma_v^2, \sigma_w^2|y) = p(y)^{-1} \left\{ \prod_{t=1}^T p(y_t|\theta_t, \beta_1, \sigma_v^2) \right\} \left\{ \prod_{t=2}^T p(\theta_t|\theta_{t-1}, \sigma_w^2) \right\} p(\theta_1) p(\beta_1) p(\sigma_v^2) p(\sigma_w^2) \quad (8)$$

which is analytically intractable but samples from this distribution may be generated in a straightforward fashion using Markov chain Monte Carlo (e.g., Smith and Roberts (1993)). This was performed using the WinBUGS software (Spiegelhalter et al. (1998)), noting that dealing with the cyclical graph that arises at stage two requires some of the conditional distributions to be explicitly specified (Spiegelhalter et al. (1996)). If the variances σ_v^2 and σ_w^2 were known then the Kalman filter could be applied (Meinhold and Singpurwalla (1983); Fahrmeir and Tutz (1994)) for efficient estimation.

In this context, the values of the pollutants on unmonitored (i.e. missing) days is of great interest; it should also be noted that analytical computation is hampered by missing values. Such values may be treated as parameters and the posterior is then obtained over these values and the model parameters. This approach to dealing with missing values may be easily implemented within the WinBUGS software. Inference on the parameters of interest is then performed via averaging over the distribution of the missing values. If, for storage reasons for example, it is impractical to save the samples on-line, samples may be generated from the posterior over the missing values retrospectively. If y_o denotes the observed values, y_m the missing values, and $\lambda = (\theta, \beta_1, \sigma_v^2, \sigma_w^2)'$, then samples from the distribution of missing values may be generated via

$$\begin{aligned} p(y_m|y_o) &= \int p(y_m|\lambda, y_o) p(\lambda|y_o) d\lambda \\ &\approx \frac{1}{K} \sum_{k=1}^K p(y_m|\lambda^{(k)}, y_o), \end{aligned}$$

where $\theta^{(k)} \sim p(\lambda|y_o)$, $k = 1, \dots, K$ denote K realizations of the Markov chain. This formulation can also be used for predicting pollution levels on future days.

Using this model it is possible to estimate the site effects, and thus pollution levels, at locations where there is no monitoring site. Again, considering a generic pollutant; based on the posterior estimates of the site effects, m_s and the variance-covariance matrix $\sigma_m^2 \Sigma_m$, then for a site at a new location, m_{S+1} , the vector of levels at $S+1$ locations $(m_1, \dots, m_S, m_{S+1})$ follows a multivariate normal distribution with zero mean and $(S+1) \times (S+1)$ variance-covariance matrix with upper $S \times S$ component given by $\sigma_m^2 \Sigma_m$, element $(S+1) \times (S+1)$ given by σ_m^2 , and final row, without this element, given by the correlations $\sigma_m^2 f(d_{s,S+1}, \phi)$, $s = 1, \dots, S$. The latter is denoted by the $S \times 1$ vector $\sigma_m^2 \Omega$. Letting $m = (m_1, \dots, m_S)'$, the conditional distribution of $m_{S+1}|m$ is, by properties of the multivariate normal distribution, also normal with mean and variance given by

$$E[m_{S+1}|m] = \sigma_m^{-2} \Omega' \Sigma_m^{-1} m,$$

and

$$\text{var}(m_{S+1}|m) = \sigma_m^2 (1 - \Omega' \Sigma_m^{-1} \Omega)^{-1},$$

respectively. For exploratory purposes, the posterior medians (for example) of the parameters may be substituted into these expressions. A more accurate, though computationally expensive, strategy is to average the normal distributions that result from individual draws from the posterior distribution.

3. Initial Analyses

In general it is recommended that a number of initial analyses are performed to examine modelling assumptions. Here, each pollutant may be separately at each site to give, in this example, 32 analyses. The fits to individual sites/pollutants may be examined and the plausibility of the second stage modelling assumptions may also be assessed. The assumptions made in the model presented in Section 2 allow data from the separate sites to be combined, resulting in a ‘smoothing’ across the individual analyses. By considering a simplification of the general model, the spatial and temporal aspects can be combined to carry out four analyses, one for each pollutant. Similarly, for each site all four pollutants may be modelled simultaneously. Examination of changes in model parameters across these analyses allows an assessment of the impact of modelling assumptions.

3.1. Single pollutant, single monitoring site

The use of the above model is now demonstrated using the London data. At this stage, separate analyses are performed for each pollutant at each site. The analyses are performed with and without the use of daily temperature. In later sections the inference is refined as the data are linked across sites and pollutants via the hierarchy. Prior sensitivity is also addressed, and residuals examined to assess the assumptions of the model.

Initially the prior distributions are specified as $\sigma_v^{-2} \sim Ga(1, 0.01)$, $\sigma_w^{-2} \sim Ga(1, 0.01)$ and β_1, β_2 (the linear and quadratic effects of temperature) $\sim N(0, 1000)$, the latter corresponding to vague prior beliefs. The choices for the precisions give approximate values of the standard deviations of 0.1, i.e. on the original scale, a coefficient of variation of approximately 10% with a large spread. Two separate chains starting from different initial values were run for each model. Convergence was assessed by visual examination of the ‘time series’ plots of the samples for each chain, and computing the Gelman and Rubin statistic (Gelman and Rubin (1992)), which calculates the ratio of the between to within chain variability. The two Markov chains were run for 15,000 iterations, discarding the first 5,000 of each as ‘burn in’.

Table 4 contains posterior summaries for the results of two pollutants, PM₁₀ and SO₂, analysed at the two sites with the longest runs of data, Bloomsbury and Bexley. The estimates of the system variances for PM₁₀ from all eight sites were of similar magnitude, as were the measurement variances. The day-to-day variability, as measured by $\text{sd}(\theta_t|y)$, showed greater differences over the sites, but was, in general, a greater component of variability than the measurement error component. For example, for PM₁₀ at the two sites presented in Table 4, the variance of the measurement error accounts for around 25% of the total variability, i.e. $\sigma_v^2 + \text{var}(\theta|y)$. For prediction or inference for missing values, higher precision is obtained if the total variability is primarily from sources that are modelled via components that are smoothed, across time in this case. For SO₂, there was greater overall variability than seen with PM₁₀, with σ_v , σ_w and $\text{sd}(\theta_t|y)$ generally being higher, with lowest variability (as with PM₁₀) being observed

at the Bloomsbury site. In general, measurement error comprised a higher proportion of total variability than for PM_{10} . The variability observed in both NO and CO was greater with more heterogeneity over the eight sites than both PM_{10} and SO_2 , with NO showing the largest differences.

From the relationships observed in Table 2 and due to the mechanisms affecting the pollutant levels discussed in Section 1, a relationship with temperature would be expected. Examination of the scatterplots of PM_{10} versus temperature indicated that a simple linear relationship would not be an adequate description, due to the seasonal changes in the composition of PM_{10} , for example the larger contribution of ozone in the warmer months and from heating materials in the winter. Such a plot also showed large residual variability, indicating that the predictive effect of temperature would not be strong. The observation variances are relatively unchanged when temperature is added (apart from PM_{10} for Bloomsbury where the median increases), but the system variances are reduced, with the reductions being comparatively larger for PM_{10} than for the other pollutants.

The posterior medians presented in Table 4 appear to be relatively robust to the choice of prior distributions for σ_v^{-2} and σ_w^{-2} . A number of prior combinations were tried, including the choices $Ga(1, 0.01)$, $Ga(0.5, 0.0005)$ and $Ga(0.001, 0.001)$, with very little difference in the resulting posterior medians.

3.2. *Single pollutant, multiple monitoring site*

The vague normal priors $N(0, 1000)$ were assigned to β_1 and β_2 , while for the precisions $\sigma_{v_s}^{-2}$, $s = 1, \dots, S$, σ_w^{-2} and σ_m^2 , the distributions $Ga(1, 0.01)$ were assumed. With a small number of sites, there will be little information on ϕ , the parameter relating correlation with the distance between sites, and it will therefore be highly affected by the choice of prior. In this example, the limits of the uniform distribution were chosen to represent the cases where the correlation falls to 0.95 at a distance of 20km, reflecting very high correlation between all the sites, and 0.1 indicating far less spatial dependence.

The results of jointly modelling PM_{10} at the eight sites, using exchangeable site effects (4) and those modelled multivariately according to (3) are shown in Table 5 and may be compared with those described in Section 3.1 and shown in Table 4. It is noted that the medians are virtually unchanged under the two spatial models, since there are large amounts of data at each site. The intervals are slightly wider under the spatial model reflecting the loss of information with dependent observations. The posterior for ϕ is very diffuse due to the small number of sites, and it shows high correlation between sites measuring PM_{10} within a urban area, the median corresponds to the correlation at 30km falling to 0.85 ($-\exp \phi d$) with a 95% credible interval of (0.75, 0.94). The effect of distance on the correlations between site measurements of the other pollutants was greater, with SO_2 , NO and CO falling to approximately 0.75 (not shown).

In comparison to the results for pollutant-site specific analyses, there is a noticeable decrease in the observation variance at all of the sites. Although the system variances increase slightly for the Bloomsbury and Bexley sites, they decrease for the others, for which higher system variances were observed in the separate analyses. This reflects the fact that the underlying level, θ_t , is now 'responding' to eight series of data rather than just 'smoothing' one. This suggests that if the objective of the analysis is to produce a set of estimates for a single site for which there is a relatively stable series of data over the study period, i.e. Bloomsbury in this example, there is little to be gained from incorporating data from other sites with greater variability and/or proportion of missing values. The advantage in doing so however, would be the ability to improve the estimates for the other sites, by borrowing information.

The coefficients of temperature, linear and quadratic, are similar to those seen in the separate analyses, but their addition to the model now accounts for a smaller reduction in the system variance (3.5%, not shown). Examining the average of the $S = 8$ measurement errors, the temporal, spatial and measurement error components account for, respectively, 80%, 10% and 10% of the total variability. Hence the variability due to the temporal components dominates that from the measurement error and spatial variability, as measured by σ_m .

Estimates of the differences in levels recorded at the eight sites are also given, the posterior medians of the site effects, m_s , range from -0.1209 to +0.1341. The site effects for the Bloomsbury and Bexley sites are +0.1341 and -0.0696 respectively, indicating the increased levels observed at the Bloomsbury site, which is in the centre of London (defined as an urban background site), compared to those at Bexley, which is on the outskirts of the city and defined as a suburban site.

Similar results were observed for the other pollutants, with the the site effects being more pronounced for CO and NO, reflecting the greater differences observed in the individual site analyses.

3.3. Multiple pollutants, single monitoring site

This model was applied to data on four pollutants (PM₁₀, SO₂, NO and CO) for each site individually. Again the priors $\sigma_{vp}^{-2} \sim Ga(1, 0.01)$, $p = 1, \dots, P$, and $\beta_1, \beta_2 \sim N(0, 1000)$, were assigned. For the parameters of the Wishart distribution, d was chosen to be equal to four, the dimension of Σ_p ; D was then chosen so that the diagonals of the expected value (D/d) represent a 10% coefficient of variation. The off-diagonals were taken to be zero.

The results of the model applied to data from the Bloomsbury site can be seen in Tables 6 and 7. The results for PM₁₀ shown in Table 6 are similar to those seen in the univariate example (Table 4), with a increase in the system variance σ_w^2 , an effect also seen at the other sites. This increase in underlying variability is a consequence of the multivariate updating mechanism, in which the estimate for a particular pollutant on day t incorporates not only the previous estimate of that pollutant, $\theta_{p,t-1}$, but also, through the off-diagonal elements of Σ_p , the previous estimates of the other pollutants.

The posterior medians of the (conditional) correlations between the underlying levels of the pollutants are given in Table 7. The correlations seen in Table 2 are not comparable as the former are conditional on the previous day's level and the latter include the non-negligible effect of independent measurement errors. Strong correlations are observed between all the parameters, meaning that in the presence of missing values of one pollutant, inference can be made by borrowing information from the non-missing values of the others.

Decomposing the total variability into that due to measurement error and that due to day-to-day variability it is apparent that the latter again dominates with 77%, 60%, 79% and 79% of the total variability being explained by the temporal variance for PM₁₀, SO₂, NO and CO, respectively.

The effect of using different prior beliefs about the variation and correlation was examined by using Wishart distributions representing combinations of coefficients of variations ranging from 0.1 to 0.75 with correlations from 0.001 to 0.9. Very little effect on the resulting posterior medians, particularly for the correlation between the system parameters, was observed. This may be expected since there are abundant daily measurements.

Figure 4(b) and (c) show contour plots of the site effects and standard deviations calculated in this way, using the posterior median values from the model, on a 20×20 grid covering the study area. As expected the variability of the effects increases with distance from the actual monitoring sites, and despite the small number of sites used, spatial patterns can be observed in the site effects, with higher values in the centre of the city. The predictions should be viewed with caution as the positions of the original monitors have not been accounted for, e.g. roadside, and this can strongly influence the level of recorded pollution.

4. Analysis

The results from the final model are shown in Tables 8 and 9. Again there is very strong correlation between the system, or underlying, parameters (Table 9). The posterior medians of the system standard deviations, σ_{wp} , can be seen in Table 8 together with those for the site-pollutant random (measurement) errors and the coefficients for temperature. The results for PM₁₀ shown in section 3.1 (Table 5) and the other pollutants (not shown) when they were modelled independently of the other pollutants are very similar to the results from this multi-pollutant model. The site effects are also given, with slightly increased standard deviations, σ_{mp} , and show a pattern similar to those from the single pollutant models, again with those for NO and CO again being more pronounced. The temporal component, represented by $sd(\theta|y)$ is also very similar to that observed when PM₁₀ was jointly modelled with the other pollutants at a single site (Section 3.3, Table 6). The measurement error (and temperature effects) are also consistent. The proportions of variability associated with the temporal, spatial and measurement error components are also similar to those previously observed, at 75%, 15% and 10% respectively for PM₁₀.

The results for SO₂ are also comparable to those seen from both the individually (Section 3.1, Table 4) and jointly (Section 3.3, Table 6) modelled pollutant effects. Again, the temporal effect dominates

(65%), but with a larger effect of measurement error (29%) and a reduction in the spatial component (6%). In comparison to the results from the single site (multiple pollutant) models, the temporal variation associated with NO and CO is reduced, with a corresponding increase in the measurement error component.

Whilst the spatial relationships for the different pollutants seen in Figure 2 appear somewhat similar, fitting model was fit with a single spatial effect for all four pollutants resulted in an over simplification of the individual relationships (apparent from the different values of σ_{mp} and ϕ_p) and ignored the disparate scales of the pollution variables. The observed results from such a model appeared to be dominated by PM₁₀, which has the strongest correlation and for which there is most data available, at the expense of the others, notably NO and CO, which have more marked decreases in correlation with distance, but where the data is more sparse.

Substituting the posterior medians into the quadratic model to examine the relationship between temperature and (log) pollutant level again showed clear curvature for all four pollutants implying that a linear form is not sufficient (not shown).

5. Discussion

In this paper daily multiple pollution levels measured at a small set of spatial locations are modelled. The proposed model is relatively simple, particularly the isotropic, stationary spatial component; this simplicity was necessary because of the paucity of spatial information and the computational expense. In London such a structure may not be too poor an approximation due to the topology, but in other locations with a more irregular topology the model may well be inadequate. It may be possible to ease the computational burden using extensions of the algorithms used for state space models Carter and Kohn (1994). Alternatively, approximations in the spirit of Wilke and Cressie (1999) may be used. Flexibility may also be added using the approach of Schmidt and O’Hagan (2000), though the computational burden will be high.

It is noted that a number of other forms of data that commonly arise in studies such as these, may be incorporated within our framework. For example, 6-day pollution monitors, hourly data, and local traffic information could be incorporated if available.

As illustrated, estimates (and measures of uncertainty) can be made for locations at unmeasured sites though the accuracy of these estimates will crucially depend on the number of covariates (such as type of monitor) that have been incorporated. The width and symmetry of the intervals will also depend on the assumption of a Gaussian random field model. It has been seen here that spatial variability is much smaller than the temporal variability. The simple time series model used here may be extended to use more general, and stationary, AR processes, which have the advantage that the variance parameters would be directly comparable to the other variances of the other components in the model, but may compromise the model’s ability to model strong dependencies (for a discussion see Besag and Kooperberg (1995)). In similar approach to other authors, independent priors are assumed for ϕ and σ_m within the spatial model, this is not realistic and further research is required to find a more appropriate form. One possibility is to take priors such that $\sigma_m^2|\Sigma_m|$, a measure of overall variability, is constant across models.

When the health study that motivated this exposure modelling is performed, the health and exposure data could be modelled jointly. Strictly, from a Bayesian perspective, this is the correct approach, for example, it allows ‘feedback’ from the health data to inform the exposure modelling, but it could lead to the health/exposure relationship being compromised. If there was a problem with the exposure modelling, which, for example, may be due to the difficulties in estimating the spatial effects, it could distort the health/exposure relationship. Another drawback of this approach is that it is computationally intensive. An alternative approach is to plug-in the exposure levels, using for example, the median of the predictive distributions of observed pollutant levels and the predictive distributions themselves for unobserved levels. Such an approach ignores the variability and so will in general produce interval estimates that are too narrow. A refinement is to use an errors-in-variables approach in which the pollution estimates and standard errors are used to inform a measurement error model – such a model does not allow feedback between the exposure and exposure/health components, but does allow the uncertainty in the pollutants to be acknowledged.

Given confidence in the spatial model, it would be possible to combine health data at the individual

or small-area level, with exposure interpolations from this model. In the case of small-area data, a summary measure of all the pollutants may be used at, for example, the population weighted centroid. Strictly, as described in Wakefield and Salway (2001), the appropriate ecological model would integrate the pointwise model over the entire area, but this may not be possible in closed form.

Acknowledgements

The authors wish to thank David Lunn for invaluable assistance in implementing the models in BUGS and Leo Knorr-Held and David Briggs for useful comments. Special thanks is due to Peter Guttorp and the NRCSE for enabling Gavin Shaddick to work on this project at the University of Washington.

The Small Area Health Statistics Unit is funded by grants from the Department of Health, Department of the Environment, Food and Rural Affairs, Health and Safety Executive, Scottish Executive, National Assembly of Wales, and Northern Ireland Assembly.

Although the research described in this article has been funded in part by the United States Environmental Protection Agency through agreement CR825173-01-0 to the University of Washington, it has not been subjected to the Agency's required peer review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

References

- Berger, J., V. De Oliveira, and B. Sanso (2000). Objective Bayesian analysis of spatially correlated data. Technical report, ISDS, Duke University.
- Besag, J. and C. Kooperberg (1995). On conditional and intrinsic auto-regressions. *Biometrika* 82, 733–746.
- Brown, P., P. Diggle, M. Lord, and P. Young (2001). Space-time calibration of radar-rainfall data. *Journal of the Royal Statistical Society, Series C* 50(2), 221 – 241.
- Carroll, R., R. Chen, T. Li, H. Newton, H. Schmiediche, H. Wang, and E. George (1997). Modeling ozone exposure in Harris County, Texas. *Journal of the American Statistical Association* 92, 392–413.
- Carroll, R., D. Ruppert, and L. Stefanski (1995). *Measurement Error in Nonlinear Models*. Chapman and Hall, London.
- Carter, C. and R. Kohn (1994). On Gibbs sampling for state space models. *Biometrika* 81, 541–53.
- COMEAP (1995). *Committee on the Medical Effects of Air Pollutants. Asthma and Outdoor Air Pollution*. London, HMSO.
- Cressie, N. (1993). *Statistics for Spatial Data, Revised edition*. John Wiley, New York.
- Cressie, N. (1997). Discussion of Carroll, R.J. et al. 'Modeling ozone exposure in Harris county, Texas'. *Journal of the American Statistical Association* 92, 392–413.
- DETR (1998). Review and assessment: Monitoring air quality. Technical Paper LAQM.TG1(00), Department of the Environment, Transport and the Regions, HMSO, London.
- Diggle, P., J. Tawn, and R. Moyeed (1998). Model-based geostatistics. *Journal Royal Statistical Society C* 47(3), 299–350.
- Fahrmeir, L. and L. Knorr-Held (2000). Dynamic and semiparametric models. In M. Schimek (Ed.), *Smoothing and Regression: Approaches, Computation and Application*. Wiley.
- Fahrmeir, L. and G. Tutz (1994). *Multivariate statistical modelling based on generalized linear models*. Springer.
- Gamerman, D. and H. Migon (1993). Dynamic hierarchical models. *Journal Royal Statistical Society B* 55, 629–642.

- Gelman, A. and D. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* 7, 457–511.
- Guttorp, P., W. Meiring, and P. Sampson (1994). A space-time analysis of ground-level ozone data. *Environmetrics* 5, 241–254.
- Handcock, M. and J. Wallis (1994). An approach to statistical spatial-temporal modelling of meteorological fields (with discussion). *Journal of the American Statistical Association* 89, 368–390.
- Huang, H. and N. Cressie (1996). Spatio-temporal prediction of snow water equivalent using the Kalman filter. *Computational Statistics and Data Analysis* 22, 159–175.
- Mardia, K., C. Goodall, E. Redfern, and F. Alonso (1998). The kriged Kalman filter. *Test* 7, 217–276.
- Meinhold, R. and N. Singpurwalla (1983). Understanding the Kalman filter. *American Statistician* 37, 123–127.
- Onursal, B. and S. Gautan (1997). Vehicular air pollution. experiences from seven latin American urban centers. Technical Paper 373, World Bank, The World Bank, Washington, D.C.
- Pole, A., M. West, and J. Harrison (1994). *Bayesian Forecasting and Time Series Analysis*. Chapman and Hall.
- Pope, C. I., D. Dockery, and J. Schwartz (1995). Review of epidemiological evidence of health effects of particulate air pollution. *Inhalation Toxicology* 7, 1–18.
- Sampson, P. and P. Guttorp (1992). Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association* 87, 108–119.
- Schmidt, A. M. and A. O’Hagan (2000). Bayesian inference for nonstationary spatial covariance structure via spatial deformations. Research report, Department of Probability and Statistics, University of Sheffield.
- Schwartz, J. (1993). Air pollution and daily mortality in Birmingham, Alabama. *American Journal of Epidemiology* 137, 1136–1147.
- Schwartz, J. and D. Dockery (1992). Increased mortality in Philadelphia associated with daily air pollution concentrations. *Am Rev Respir Dis* 145, 600–604.
- Smith, A. and G. Roberts (1993). Bayesian computation via the Gibbs sampler and other related markov chain monte carlo methods. *Journal of the Royal Statistical Society, Series B* 55, 3–23.
- Spiegelhalter, D., A. Thomas, and N. Best (1996). Computation on bayesian graphical models. In J. Bernardo, J. Berger, A. Dawid, and A. Smith (Eds.), *Sixth Valencia international meeting on Bayesian statistics*, London. Oxford University Press.
- Spiegelhalter, D., A. Thomas, and N. Best (1998). WinBUGS user manual, version 1.1.1. Cambridge, UK.
- Tonellato, S. (2001). A multivariate time series model for the analysis and prediction of carbon monoxide atmospheric concentrations. *Journal of the Royal Statistical Society C* 50(2), 187–200.
- Wakefield, J. and R. Salway (2001). A statistical framework for ecological and aggregate studies. *Journal of the Royal Statistical Society, Series A* 164, 119–137.
- West, M., J. Harrison, and H. Migon (1985). Dynamic generalized linear models and bayesian forecasting. *Journal of the American Statistical Association* 80, 73–90.
- Wilke, C., L. Berliner, and N. Cressie (1998). Heirarchical Bayesian space-time models. *Environmental and Ecological Statistics* 5, 117–154.
- Wilke, C. and N. Cressie (1999). A dimension reduction approach to space-time Kalman filtering. *Biometrika* 86, 815–829.

Table 1. Summary of pollutants measured, and periods of operation, at eight sites in London, 1994–97. The total number of days of operation are given for each pollutant at each site together with the percentage of missing observations. The units are μgm^{-3} for PM_{10} , parts per billion for SO_2 and NO and parts per million for CO .

	Period	Total	Missing	%	Mean	Min.	25%	Med.	75%	Max.
Bexley										
PM_{10}	1994-97	1461	211	14.4	24.0	4.0	15.0	20.0	29.0	92.0
SO_2	1994, 1996-97	1095	178	16.3	6.9	1.0	3.0	4.0	8.0	76.0
NO	-	-	-	-	-	-	-	-	-	-
CO	1994-97	1461	192	13.1	0.5	0.1	0.3	0.4	0.5	4.4
Bloomsbury										
PM_{10}	1994-97	1461	61	4.2	28.0	7.0	19.0	24.0	34.0	103.0
SO_2	1994-97	1461	115	7.9	8.3	1.0	4.0	6.0	11.0	48.0
NO	1994-97	1461	44	3.0	42.4	4.0	19.0	30.0	50.0	467.0
CO	1994-97	1461	68	4.7	0.7	0.1	0.4	0.6	0.8	4.3
Brent										
PM_{10}	1996-97	731	120	16.4	20.8	6.0	14.0	18.0	25.0	82.0
SO_2	1996-97	731	33	4.5	4.4	1.0	2.0	3.0	5.2	20.0
NO	1996-97	731	57	7.8	23.8	1.0	5.0	8.0	22.5	414.0
CO	1996-97	366	15	4.1	0.5	0.1	0.2	0.3	0.7	5.0
Eltham										
PM_{10}	1996-97	731	166	22.7	21.2	8.0	15.0	18.0	25.0	81.0
SO_2	1996-97	731	91	12.4	4.6	1.0	2.0	3.0	5.0	40.0
NO	1996-97	731	95	13.0	21.7	1.0	5.0	9.0	20.0	339.0
CO	-	-	-	-	-	-	-	-	-	-
Harringey										
PM_{10}	1996-97	731	161	22.0	26.2	8.0	18.0	22.0	32.0	89.0
SO_2	-	-	-	-	-	-	-	-	-	-
NO	1996-97	731	139	19.0	63.3	5.0	28.0	43.0	68.6	562.0
CO	-	-	-	-	-	-	-	-	-	-
Hillingdon										
PM_{10}	1996-97	731	225	30.8	24.5	6.0	16.0	21.0	31.0	88.0
SO_2	1996-97	731	230	31.5	5.1	1.0	3.0	4.0	6.0	28.0
NO	1996-97	731	252	34.5	81.9	2.0	31.0	67.0	105.0	506.0
CO	1996-97	731	268	36.7	0.8	0.2	0.5	0.6	0.9	4.3
N. Kensington										
PM_{10}	1996-97	731	99	13.5	23.6	9.0	16.0	20.0	27.2	89.0
SO_2	1996-97	731	91	12.4	4.6	1.0	2.0	3.0	6.0	32.0
NO	1996-97	731	106	14.5	27.6	1.0	6.0	11.0	25.0	442.0
CO	1996-97	731	93	12.7	1.2	0.1	0.4	0.7	1.3	16.6
Sutton										
PM_{10}	1996-97	731	92	12.6	25.1	9.0	17.0	22.0	29.0	250.0
SO_2	1996-97	731	96	13.1	4.9	1.0	2.7	4.0	6.0	28.4
NO	1996-97	731	106	14.5	51.1	3.0	26.3	39.0	57.0	404.0
CO	1996-97	731	104	14.2	1.1	0.2	0.8	1.0	1.3	6.7

Table 2. Correlation/covariance matrix for (logged) pollutants and temperature, Bloomsbury, 1994–97. The variances lie on the diagonal, with covariances above and correlations below.

	PM_{10}	SO_2	NO	CO	Temp
PM_{10}	0.41 ²	0.15	0.14	0.091	0.23
SO_2	0.49	0.72 ²	0.30	0.14	-1.21
NO	0.45	0.57	0.73 ²	0.26	-2.09
CO	0.45	0.38	0.71	0.49 ²	-0.82
Temp	0.10	-0.29	-0.50	-0.29	5.71 ²

Table 3. Correlation/covariance matrix for (logged) values of PM₁₀, measured at eight sites, 1994–97. The variances lie on the diagonal, with covariances above and correlations below.

	Bexley	Bloomsbury	Brent	Eltham	Harringey	Hillingdon	N. Kensington	Sutton
Bexley	0.46 ²	0.16	0.19	0.18	0.16	0.17	0.18	0.17
Bloomsbury	0.92	0.39 ²	0.15	0.15	0.14	0.15	0.15	0.14
Brent	0.92	0.90	0.43 ²	0.16	0.16	0.17	0.17	0.15
Eltham	0.95	0.91	0.91	0.42 ²	0.14	0.15	0.16	0.16
Harringey	0.90	0.89	0.94	0.87	0.39 ²	0.16	0.15	0.13
Hillingdon	0.83	0.83	0.88	0.79	0.91	0.45 ²	0.17	0.14
N.Kensington	0.95	0.93	0.96	0.94	0.94	0.87	0.42 ²	0.15
Sutton	0.89	0.84	0.84	0.90	0.82	0.72	0.88	0.42 ²

Table 4. Observation and system variances, with coefficients for temperature, for two pollutants, PM₁₀ and SO₂, modelled individually at two sites, Bloomsbury and Bexley. The three sets of results show the effect of adding linear and quadratic effects of temperature to the model.

	Without Temp.			Linear Temp.			Quad. Temp		
	Median	2.5%	97.5%	Median	2.5%	97.5%	Median	2.5%	97.5%
PM₁₀, Bloomsbury									
σ_v	0.2057	0.1822	0.2278	0.2839	0.2694	0.2989	0.2774	0.2639	0.2916
σ_w	0.2393	0.2143	0.2661	0.1710	0.1480	0.1977	0.1528	0.1313	0.1753
sd(θy)	0.3639	0.3459	0.3842	0.3263	0.3038	0.3511	0.3068	0.2856	0.3291
Temp.	-	-	-	-0.0006	-0.0092	0.0076	-0.0899	-0.1071	-0.0721
Temp. ²	-	-	-	-	-	-	0.0042	0.0035	0.0049
PM₁₀, Bexley									
σ_v	0.3235	0.3057	0.3421	0.3229	0.3049	0.3415	0.3144	0.2977	0.3319
σ_w	0.2037	0.1743	0.2357	0.2045	0.1760	0.2387	0.1848	0.1577	0.2138
sd(θy)	0.5435	0.4007	1.0250	0.4916	0.3796	0.9152	0.4281	0.3649	0.5880
Temp.	-	-	-	-0.0038	-0.0142	0.0071	-0.1055	-0.1260	-0.0850
Temp. ²	-	-	-	-	-	-	0.0047	0.0039	0.0056
SO₂, Bloomsbury									
σ_v	0.5709	0.5437	0.5989	0.5688	0.5426	0.5962	0.5632	0.5383	0.5897
σ_w	0.1690	0.1335	0.2095	0.1567	0.1246	0.1926	0.1349	0.1060	0.1679
sd(θy)	0.4680	0.4257	0.5119	0.4170	0.3743	0.4621	0.3886	0.3469	0.4353
Temp.	-	-	-	-0.0411	-0.0552	-0.0270	-0.1590	-0.1893	-0.1297
Temp. ²	-	-	-	-	-	-	0.0055	0.0043	0.0067
SO₂, Bexley									
σ_v	0.6433	0.6041	0.6862	0.6334	0.5949	0.6745	0.6192	0.5811	0.6605
σ_w	0.2655	0.2081	0.3313	0.2745	0.2216	0.3377	0.2488	0.1894	0.3128
sd(θy)	1.2880	0.9008	2.0620	1.1000	0.7747	2.0190	1.0860	0.6726	1.7760
Temp.	-	-	-	-0.0395	-0.0609	-0.0181	-0.2083	-0.2546	-0.1620
Temp. ²	-	-	-	-	-	-	0.0077	0.0058	0.0095

Table 5. Observation and system variances, with coefficients for temperature, for PM_{10} modelled jointly at eight sites, but independently of measurements of other pollutants. Site effects, and variances are presented for both independent and multivariate spatial models. The labelling is: 1 = Bexley, 2 = Bloomsbury, 3 = Brent, 4 = Eltham, 5 = Harringey, 6 = Hillingdon, 7 = North Kensington, 8 = Sutton.

	Independent			Multivariate		
	Median	2.5%	97.5%	Median	2.5%	97.5%
σ_{v1}	0.1361	0.1290	0.1437	0.1361	0.1291	0.1438
σ_{v2}	0.1279	0.1211	0.1349	0.1278	0.1211	0.1348
σ_{v3}	0.1042	0.0976	0.1115	0.1042	0.0975	0.1115
σ_{v4}	0.1248	0.1169	0.1334	0.1248	0.1170	0.1335
σ_{v5}	0.1222	0.1145	0.1306	0.1222	0.1145	0.1307
σ_{v6}	0.2116	0.1987	0.2266	0.2114	0.1984	0.2259
σ_{v7}	0.0632	0.0573	0.0695	0.0632	0.0573	0.0693
σ_{v8}	0.2290	0.2166	0.2425	0.2290	0.2166	0.2429
σ_w	0.3462	0.3326	0.3605	0.3465	0.3328	0.3609
$sd(\theta y)$	0.4062	0.4007	0.4143	0.4067	0.4005	0.4164
m_1	-0.0696	-0.0785	-0.0605	-0.0696	-0.0785	-0.0607
m_2	0.1341	0.1256	0.1429	0.1341	0.1257	0.1426
m_3	-0.1209	-0.1292	-0.1120	-0.1210	-0.1294	-0.1125
m_4	-0.1103	-0.1203	-0.0998	-0.1105	-0.1205	-0.1005
m_5	0.1098	0.1000	0.1200	0.1098	0.0999	0.1195
m_6	0.0131	-0.0043	0.0299	0.0132	-0.0032	0.0300
m_7	0.0031	-0.0031	0.0094	0.0030	-0.0031	0.0090
m_8	0.0410	0.0250	0.0575	0.0410	0.0250	0.0572
σ_m	0.0955	0.0627	0.2202	0.1019	0.0668	0.1794
ϕ	-	-	-	0.005675	0.002158	0.009778
Temp.	-0.0905	-0.1033	-0.0741	-0.0872	-0.1091	-0.0609
Temp. ²	0.0037	0.0031	0.0043	0.0035	0.0026	0.0044

Table 6. Observation and system variances, with coefficients for temperature, for four pollutants modelled using the multivariate pollution model. Results are from the Bloomsbury site, modelled independently of measurements at other sites.

	Median	2.5%	97.5%
PM₁₀			
σ_{v1}	0.2258	0.2148	0.2375
σ_{w1}	0.3424	0.3124	0.3737
$sd(\theta_1 y)$	0.4099	0.3899	0.4299
Temp ₁	-0.0972	-0.1144	-0.0782
Temp ₁ ²	0.0043	0.0035	0.0050
SO₂			
σ_{v2}	0.4823	0.4607	0.5049
σ_{w2}	0.4715	0.4188	0.5260
$sd(\theta_2 y)$	0.5948	0.5503	0.6385
Temp ₂	-0.1663	-0.1981	-0.1357
Temp ₂ ²	0.0052	0.0038	0.0067
NO			
σ_{v3}	0.2502	0.2378	0.2633
σ_{w3}	0.3703	0.3357	0.4069
$sd(\theta_3 y)$	0.4795	0.4578	0.5015
Temp ₂	-0.0904	-0.1089	-0.0716
Temp ₃ ²	0.0031	0.0024	0.0037
CO			
σ_{v4}	0.3383	0.3220	0.3555
σ_{w4}	0.5511	0.5025	0.6020
$sd(\theta_4 y)$	0.6661	0.6350	0.6970
Temp ₄	-0.1385	-0.1652	-0.1122
Temp ₄ ²	0.0027	0.0016	0.0039

Table 7. Posterior medians of the Σ_p matrix between the underlying levels of four pollutants modelled using the multivariate pollution model at the Bloomsbury site, modelled independently of measurements at other sites. The variances lie on the diagonal, with covariances above and correlations below.

	PM ₁₀	SO ₂	NO	CO
PM ₁₀	0.3424 ²	0.1442	0.1039	0.1535
SO ₂	0.8806	0.4715 ²	0.1472	0.2315
NO	0.8192	0.8472	0.3703 ²	0.1866
CO	0.8134	0.9202	0.9146	0.5511 ²

Table 8. Posterior medians from the general model with dependencies across time and pollutants, with pollutant specific spatial effects. Pollutant specific observation and system variances, together with coefficients for temperature, are presented together with site effects and variances, for the multivariate spatial components. The labelling is: 1 = Bexley, 2 = Bloomsbury, 3 = Brent, 4 = Eltham, 5 = Harringey, 6 = Hillingdon, 7 = North Kensington, 8 = Sutton.

PM ₁₀							
	Median	2.5%	97.5%		Median	2.5%	97.5%
σ_{v11}	0.1380	0.1310	0.1456	m ₁₁	-0.0696	-0.0785	-0.0607
σ_{v21}	0.1273	0.1209	0.1343	m ₂₁	0.1344	0.1260	0.1427
σ_{v31}	0.1049	0.0983	0.1122	m ₃₁	-0.1209	-0.1293	-0.1124
σ_{v41}	0.1251	0.1172	0.1339	m ₄₁	-0.1105	-0.1204	-0.1005
σ_{v51}	0.1227	0.1150	0.1312	m ₅₁	0.1098	0.1000	0.1195
σ_{v61}	0.2112	0.1983	0.2255	m ₆₁	0.0128	-0.0039	0.0291
σ_{v71}	0.0623	0.0565	0.0684	m ₇₁	0.0030	-0.0031	0.0090
σ_{v81}	0.2291	0.2166	0.2429	m ₈₁	0.0411	0.0250	0.0570
σ_{w1}	0.3438	0.3302	0.3581	σ_{m1}	0.1297	0.0806	0.2597
sd($\theta_1 y$)	0.4060	0.4005	0.4138	ϕ_1	0.00752	0.00181	0.01130
Temp _{·1}	-0.0951	-0.1155	-0.0811	-	-	-	-
Temp _{·1} ²	0.0040	0.0034	0.0047	-	-	-	-
SO ₂							
	Median	2.5%	97.5%		Median	2.5%	97.5%
σ_{v12}	0.4811	0.4545	0.5099	m ₁₂	0.0064	-0.0607	0.0783
σ_{v22}	0.4211	0.4012	0.4422	m ₂₂	0.3084	0.2451	0.3756
σ_{v32}	0.4244	0.3992	0.4517	m ₃₂	-0.2797	-0.3470	-0.2131
σ_{v42}	0.4743	0.4459	0.5052	m ₄₂	-0.1915	-0.2601	-0.1223
σ_{v52}	0.1199	0.0519	0.6343	m ₅₂	0.0965	-0.3228	0.5017
σ_{v62}	0.3551	0.3295	0.3830	m ₆₂	0.1505	0.0842	0.2173
σ_{v72}	0.3822	0.3580	0.4085	m ₇₂	-0.1495	-0.2149	-0.0848
σ_{v82}	0.4135	0.3885	0.4410	m ₈₂	0.0594	-0.0057	0.1266
σ_{w2}	0.4710	0.4458	0.4973	σ_{m2}	0.2701	0.1651	0.5590
sd($\theta_2 y$)	0.5719	0.5559	0.5880	ϕ_2	0.00871	0.00262	0.01138
Temp _{·2}	-0.1429	-0.1663	-0.1174	-	-	-	-
Temp _{·2} ²	0.0048	0.0038	0.0058	-	-	-	-
NO							
	Median	2.5%	97.5%		Median	2.5%	97.5%
σ_{v13}	0.1196	0.0523	0.6354	m ₁₃	0.1626	-1.2820	2.2640
σ_{v23}	0.2668	0.2509	0.2837	m ₂₃	0.3293	0.0293	0.5361
σ_{v33}	0.6721	0.6349	0.7124	m ₃₃	-0.8548	-1.1590	-0.6441
σ_{v43}	0.6660	0.6278	0.7067	m ₄₃	-0.8725	-1.1730	-0.6616
σ_{v53}	0.3804	0.3557	0.4064	m ₅₃	0.6277	0.3269	0.8358
σ_{v63}	0.8761	0.8211	0.9370	m ₆₃	0.7097	0.4105	0.9264
σ_{v73}	0.5013	0.4712	0.5334	m ₇₃	-0.6188	-0.9207	-0.4092
σ_{v83}	0.4407	0.4136	0.4699	m ₈₃	0.5235	0.2223	0.7323
σ_{w3}	0.4994	0.4757	0.5241	σ_{m3}	0.8929	0.5204	2.0340
sd($\theta_3 y$)	0.6250	0.6105	0.6432	ϕ_3	0.00842	0.00236	0.01138
Temp _{·3}	-0.1366	-0.1615	-0.1124	-	-	-	-
Temp _{·3} ²	0.0026	0.0016	0.0036	-	-	-	-
CO							
	Median	2.5%	97.5%		Median	2.5%	97.5%
σ_{v14}	0.3580	0.3419	0.3748	m ₁₄	-1.1940	-1.7390	-0.5778
σ_{v24}	0.2725	0.2592	0.2866	m ₂₄	-0.6926	-1.2390	-0.0771
σ_{v34}	0.5987	0.5537	0.6491	m ₃₄	-1.2810	-1.8270	-0.6542
σ_{v44}	0.1209	0.0526	0.6176	m ₄₄	1.5530	-0.9085	5.4500
σ_{v54}	0.1202	0.0525	0.6330	m ₅₄	2.7910	0.3848	7.9660
σ_{v64}	0.3465	0.3223	0.3729	m ₆₄	-0.6946	-1.2390	-0.0825
σ_{v74}	0.6167	0.5825	0.6539	m ₇₄	-0.5123	-1.0560	0.1012
σ_{v84}	0.2753	0.2572	0.2951	m ₈₄	-0.1723	-0.7199	0.4416
σ_{w4}	0.2962	0.2806	0.3126	σ_{m4}	1.9380	0.6739	4.7520
sd($\theta_4 y$)	0.4073	0.3961	0.4185	ϕ_4	0.00894	0.00310	0.01139
Temp _{·4}	-0.0975	-0.1140	-0.0800	-	-	-	-
Temp _{·4} ²	0.0032	0.0025	0.0038	-	-	-	-

Table 9. Posterior medians of the Σ_p matrix between the underlying levels of four pollutants modelled using the multivariate pollution model with multivariate spatial effects for all eight monitoring sites. The variances lie on the diagonal, with covariances above and correlations below.

	PM ₁₀	SO ₂	NO	CO
PM ₁₀	0.3438 ²	0.1142	0.1143	0.0705
SO ₂	0.7055	0.4710 ²	0.1889	0.1022
NO	0.6655	0.8027	0.4994 ²	0.1417
CO	0.6924	0.7327	0.9578	0.2962 ²

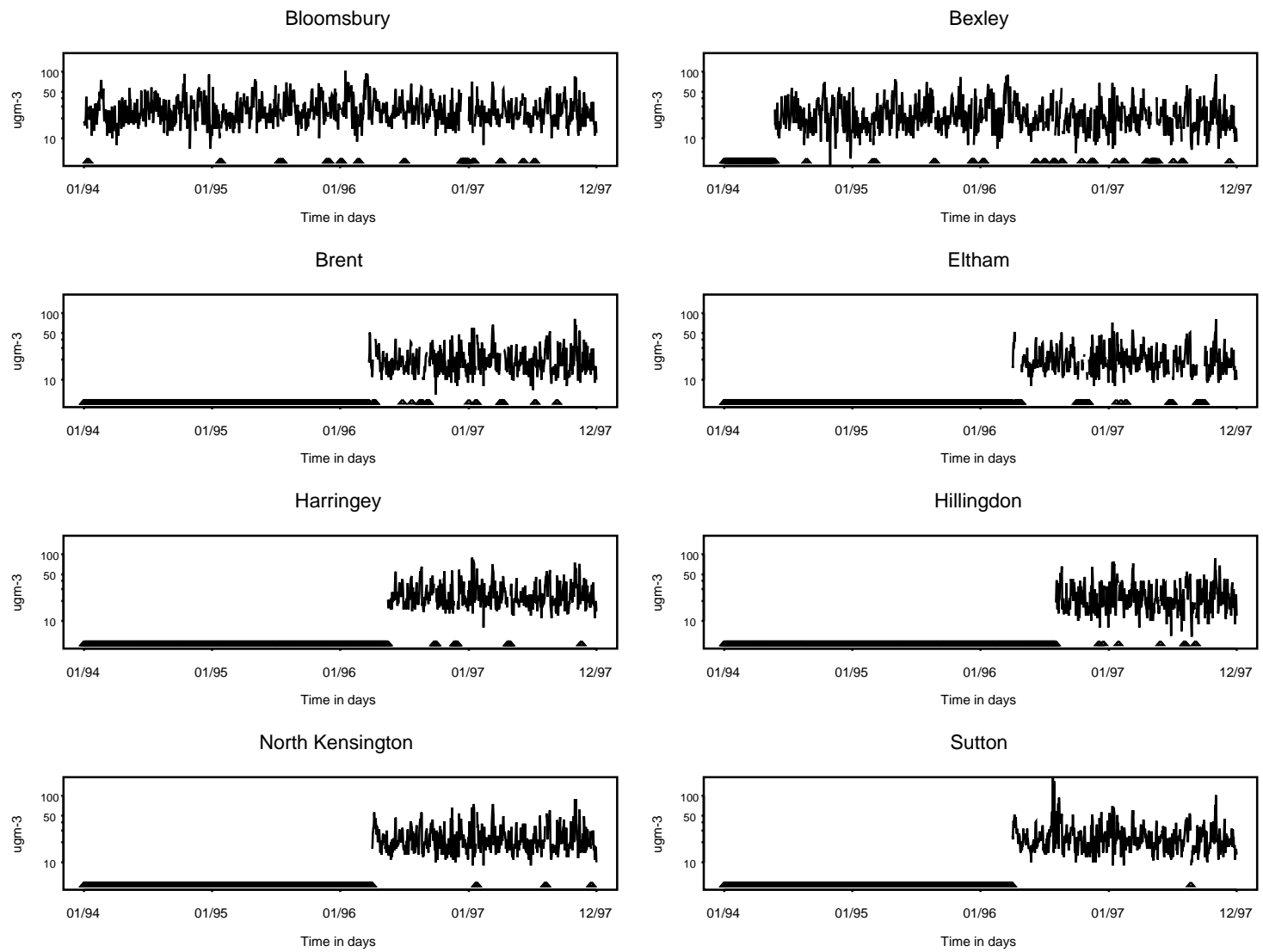


Fig. 1. Time series plots of logged values of PM_{10} for eight sites in London, 1994–97 (missing values are marked on the horizontal axis as triangles).

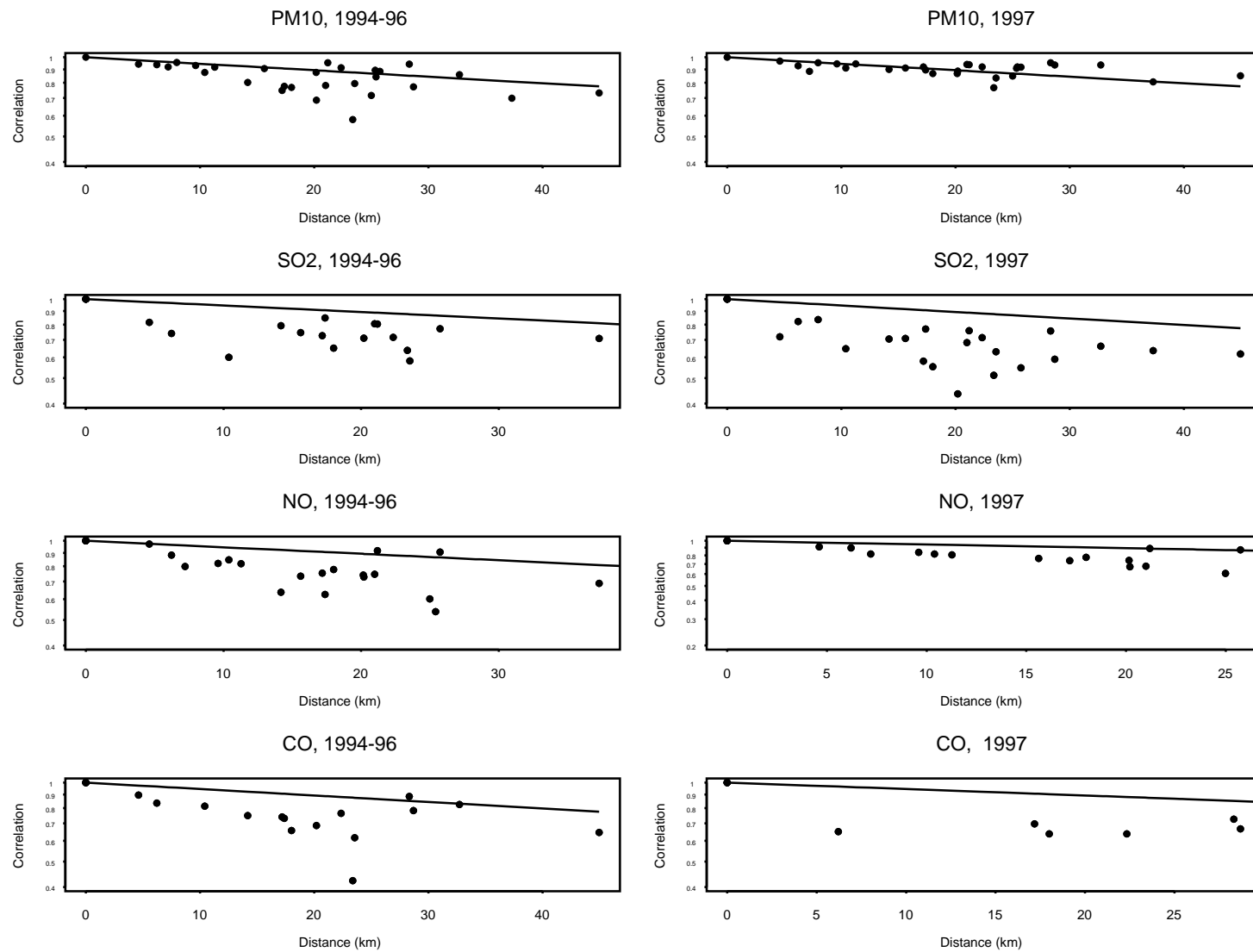


Fig. 2. The (log) correlation between logged values of pollutants and distance between sites. The reference line is $\log(\text{correlation}) = -\phi_p \text{ distance}$, using the posterior median for ϕ_p from the spatial model for PM₁₀

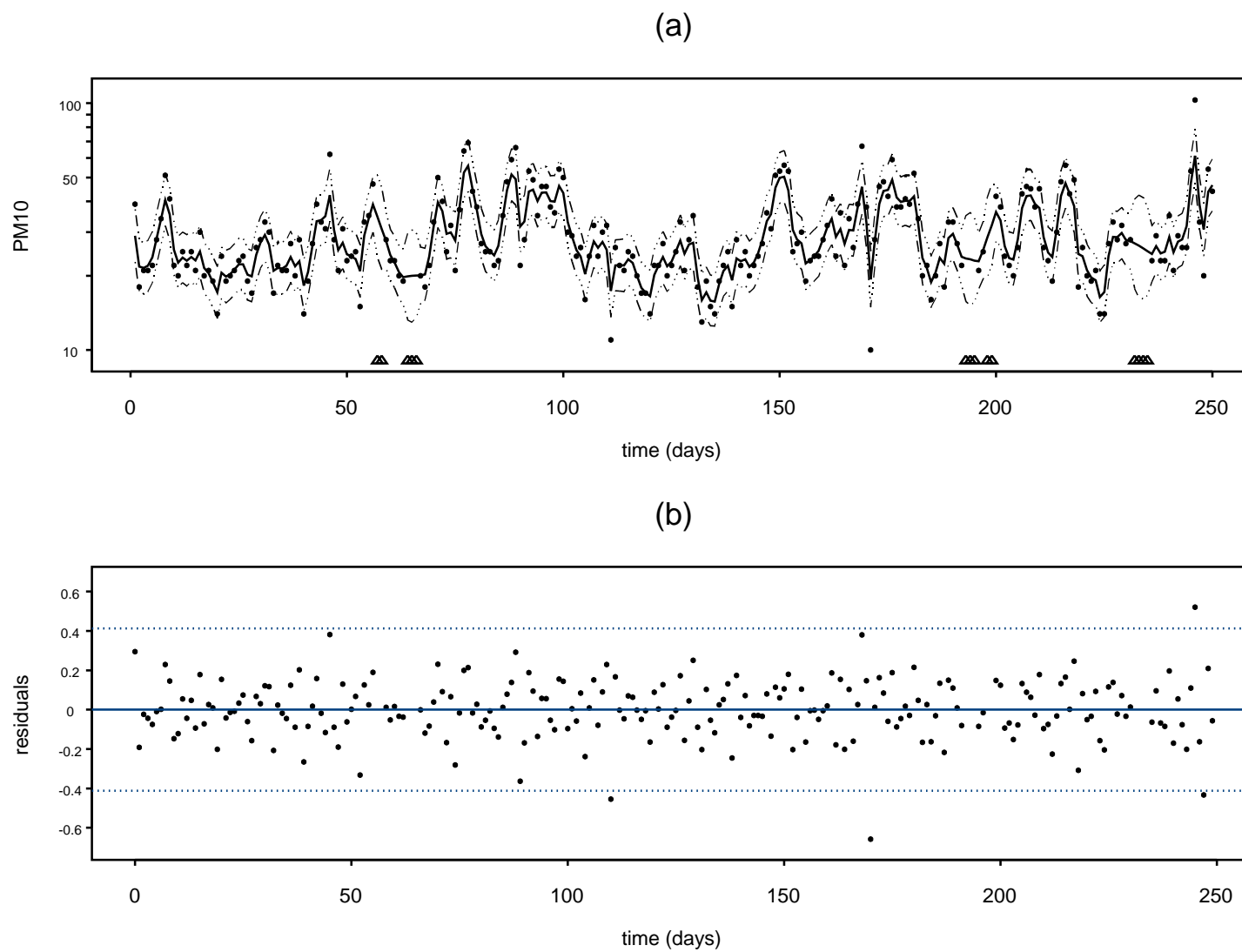


Fig. 3. Time series plot for a subset of 250 days: (a) recorded measurements and estimated $\hat{\theta}_t$ for (log) PM₁₀, Bloomsbury, triangles indicate missing values; (b) differences between recorded and estimated values. In (a) the dotted lines indicate approximate 90% intervals, in (b) the dotted lines are $\pm 2\hat{\sigma}_v$

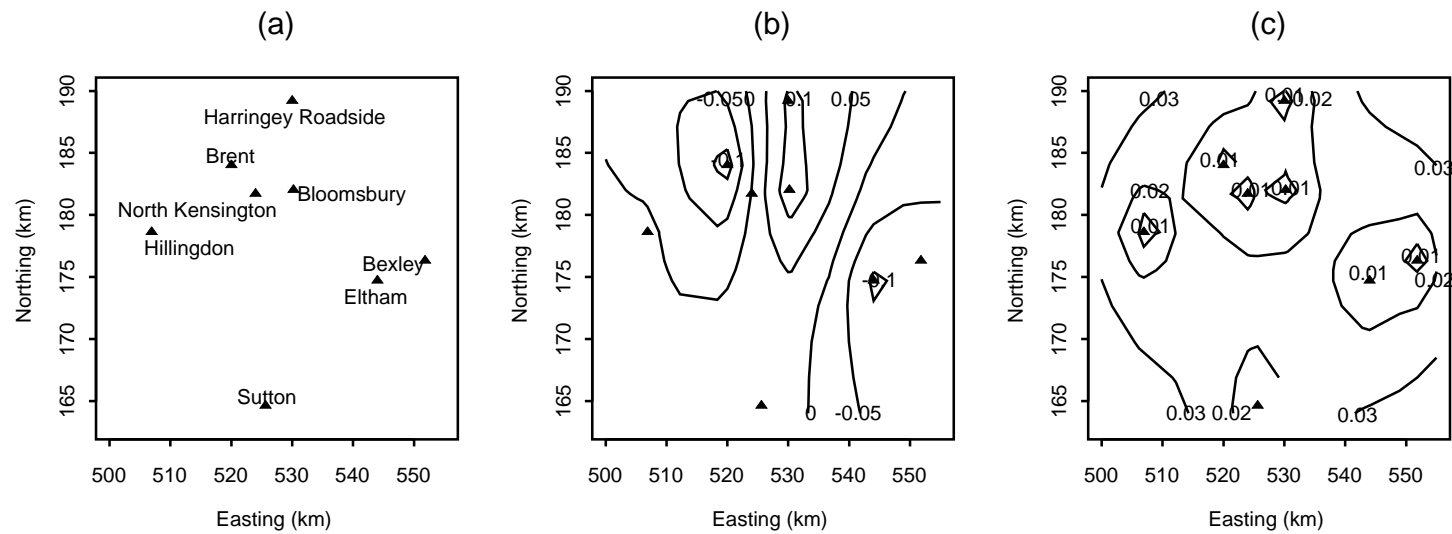


Fig. 4. Contour plot of (a) site effects based on a 20x20 grid of locations without a pollution monitor (b) the corresponding standard deviations of the site effects