Estimation of nonstationary spatial covariance structure

By DAVID J. NOTT AND WILLIAM T. M. DUNSMUIR

Department of Statistics, University of New South Wales, Sydney 2052, Australia

djn@maths.unsw.edu.au   w.dunsmuir@unsw.edu.au

Summary

We introduce a method for estimating nonstationary spatial covariance structure from space-time data and apply the method to an analysis of Sydney wind patterns. Our method constructs a process honouring a given spatial covariance matrix at observing stations and uses one or more stationary processes to describe conditional behaviour given observing site values. The stationary processes give a localised description of the spatial covariance structure. The method is computationally attractive, and can be extended to the assessment of covariance for multivariate processes. The technique is illustrated for data describing the east-west component of Sydney winds. For this example, our own methods are contrasted with a geometrically appealing though computationally intensive technique which describes spatial correlation via an isotropic process and a deformation of the geographical space.

Some key words: Covariance estimation; Nonstationarity; Spatio-temporal modelling.

1. Introduction

Assessment of spatial covariance structure is an important aspect of many spatio-temporal modelling problems arising in meteorological and environmental applications. In these applications, assumptions such as stationarity and isotropy of the spatial covariance are often unrealistic, and may not be necessary where we have spatial measurements replicated in time.

In this paper we describe methods for estimating nonstationary spatial covariance structure that preserve a given nonstationary spatial covariance matrix at observing stations and use one or more stationary process models to extend the covariance structure to unobserved locations in the region of interest. The method is computationally attractive, produces a valid nonnegative definite nonstationary covariance function and is readily extended to the assessment of covariance for multivariate processes. The technique is illustrated for data describing the two orthogonal components of Sydney winds. The method may be useful in applications to spatial prediction, network design and as a tool for visualising spatial covariance structure.

The problem of assessing spatially nonstationary covariance behaviour for space-time data has received much recent attention. Guttorp & Sampson (1994) summarise the early literature and Sampson et al. (2001) describe more recent developments. Existing methods include those based on empirical orthogonal functions (Obled & Creutin, 1986), or on a process convolution approach (Higdon, 1998), those which use basis function expansions to define nonstationary processes in a similar way to the empirical orthogonal functions

approach (Nychka & Saltzman, 1998), Bayesian hierarchical models (Brown et al., 1994), empirical Bayes techniques (Loader & Switzer, 1992), kernel smoothing (Oehlert, 1993) and spatial deformation (Sampson & Guttorp, 1992). For the spatial deformation approach, recent attention has concentrated on a Bayesian implementation; see Damian et al. (2001) and a University of Sheffield technical report by A. Schmidt and A. O'Hagan.

Section 2 describes our method for covariance estimation, which is based on the idea of honouring a given spatial covariance matrix for observing sites, which is then extended to the whole region of interest by one or more representative stationary processes. Section 3 describes estimation of parameters used in the stationary extension models. Section 4 briefly describes the Sampson–Guttorp approach, as well as the method of Loader & Switzer and its connection with our own method. Section 5 compares our method with the spatial deformation method for modelling the covariance structure of wind fields in the Sydney area, and discussion and conclusions are given in § 6.

2. Models for nonstationary covariance structures

Let $Z = \{Z(s), s \in \mathbb{R}^d\}$ be a spatial process which can be observed at the sites $\mathcal{T} = \{s_1, \ldots, s_n\}$. The covariance matrix for $(Z(s_1), \ldots, Z(s_n))^T$ is denoted by Γ. For the present we assume that this is known. Estimation of Γ will be discussed in § 3.

Our method takes as a starting point this possibly nonstationary covariance matrix and extends it to a valid nonstationary covariance function for the whole region using one or more stationary processes with covariance functions chosen to approximate, regionally or subregionally, that of the nonstationary field at the observing sites.

Consider a collection of independent, zero mean, stationary processes $W_i(s)$ $(i = 1, \ldots, I)$, with known covariance functions $R_i(h)$ $(h \in \mathbb{R}^d)$. The processes $W_i(s)$ describe behaviour locally about a collection of locations z_i $(1 \leqslant i \leqslant I)$ chosen on a grid or based on a knowledge of likely sources of nonstationary behaviour for Z.

For the collection of sites $\mathcal{T}$, let $W_i = (W_i(s_1), \ldots, W_i(s_n))^T$, let $C_i = [R_i(s_j - s_k)]_{j,k=1}^n$ be the covariance matrix of W_i and let $c_i(s) = (R_i(s - s_1), \ldots, R_i(s - s_n))^T$ be the vector of cross-covariances between $W_i(s)$ and W_i. The random field $W_i(s)$ can be represented as

$$W_i(s) = c_i(s)^T C_i^{-1} W_i + \delta_i(s), \tag{1}$$

where $\delta_i(s)$ is a zero-mean nonstationary random field with covariance function

$$R_{\delta_i}(s, t) = R_i(t - s) - c_i(s)^T C_i^{-1} c_i(t).$$

The function $c_i(s)^T C_i^{-1} w$ is the simple kriging predictor of $W_i(s)$ given $W_i = w$ and, since the process $\delta_i(s)$ describes the residual variation, $\delta_i(s)$ is zero at the sites $s_1, \ldots, s_n$.

We now define a nonstationary process by a weighted sum of the processes $\mu_i(s) = c_i(s)^T C_i^{-1} W^*$ and $\delta_i(s)$ $(i = 1, \ldots, I)$, where W^* is a zero-mean $n \times 1$ random vector with covariance matrix Γ uncorrelated with each $\delta_i(s)$. Consider

$$W^*(s) = \sum_i v_i(s)\mu_i(s) + \sum_i v_i^{1/2}(s)\delta_i(s). \tag{2}$$

Note that W^* is common to all the $\mu_i(s)$. The covariance function of $W^*(s)$ is

$$R^*(s, t) = \sum_{i,j} v_i(s)v_j(t)c_i(s)^T C_i^{-1} \Gamma C_j^{-1} c_j(t) + \sum_i v_i(s)^{\frac{1}{2}} v_i(t)^{\frac{1}{2}} R_{\delta_i}(s, t). \tag{3}$$

We will use a covariance function of the form (3) as an estimator of the covariance function of $Z(s)$.

The weight function used in (2) is chosen so that $v_i(s) \geqslant 0$, $\sum_i v_i(s) = 1$, $v_i(s)$ has a maximum at s_i and decays smoothly to zero as $\|s - s_i\| \to \infty$. One choice of the $v_i(s)$ uses a kernel function such that

$$f_\eta(t) = \exp\left(-\frac{\|t\|^2}{\eta}\right),$$

where η is a smoothing parameter, and then

$$v_i(s) = \frac{f_\eta(s - z_i)}{\sum_j f_\eta(s - z_j)}.$$

The construction of the above process can be motivated by the following consideration. Suppose we have a zero-mean process with the covariance function (3), and that the vector $W^* = w$ is observed at the monitored sites. Also, suppose that $\{z_1, \ldots, z_I\}$ is the set of observing sites $\mathscr{T}$. Then the simple kriging predictor for the process at s is easily shown to be

$$\sum_i v_i(s) c_i(s)^{\mathrm{T}} C_i^{-1} w,$$

and the simple kriging variance is

$$\sum_i v_i(s) R_{\delta_i}(s, s).$$

What these relationships tell us is that simple kriging predictor and variance can be obtained as a weighted mean of simple kriging predictors and variances for the stationary processes $W_i(.)$. The weights $v_i(s)$ vary spatially so that the contribution from $W_i(.)$ increases as s becomes close to s_i. The extension is similar in spirit to a method of spatial prediction called moving window kriging (Haas, 1990a, b, 1995, 1996, 1998) in which covariance nonstationarity is modelled implicitly by local fitting of stationary models in windows centred on prediction locations.

It is easy to check that, with $s = s_k$ and $t = s_l$, (3) is Γ_{kl} since $c_i(s_k)$ and $c_j(s_l)$ are the kth column of C_i and lth column of C_j, respectively, so that, for instance, $c_i(s_k)^{\mathrm{T}} C_i^{-1} \Gamma C_j^{-1} c_j(s_l) = \Gamma_{kl}$ in (3). It is obvious from these remarks, or from the construction (2) of the nonstationary process, that the nonstationary covariance function (3) is a valid nonnegative definite covariance function which reproduces the covariance matrix Γ at the monitored sites regardless of what stationary covariance functions $R_i(h)$ are used in the construction. Note that, if s and t are far away from monitored sites, (3) is approximately $\sum_i v_i(s)^{\frac{1}{2}} v_i(t)^{\frac{1}{2}} R_i(t - s)$ if $R_i(h) \to 0$ as $\|h\| \to \infty$.

3. Estimation of nonstationary covariance structures

In the above construction we have assumed that the covariance matrix of the observed spatial process is known. In practice this will need to be estimated. We assume that we have replicates of the process through time. Let $Z_k = (Z_k(s_1), \ldots, Z_k(s_n))^{\mathrm{T}}$ be the observation of the spatial process at time k $(1 \leqslant k \leqslant M)$, write z_k for the corresponding realisation of Z_k, and let

$$\bar{z} = \frac{1}{M} \sum_{k=1}^{M} z_k$$

be the estimated spatial mean vector at the monitored sites. Write $R_Z(s, t)$ for the spatial

covariance function of Z,

$$R_Z(s, t) = \text{cov}\{Z(s), Z(t)\},$$

let Γ be the covariance matrix of Z_k, and let $\hat{\Gamma}$ be an estimate of Γ, the covariance matrix of Z_k, obtained by averaging over time:

$$\hat{\Gamma} = \frac{1}{M} \sum_{i=1}^{M} (z_i - \bar{z})(z_i - \bar{z})^{\text{T}}.$$

If there are missing values then a different estimator for Γ may be used; see §4.

It will sometimes be convenient to describe the second-order spatial structure of Z via the spatial dispersion function or variogram,

$$D_Z(s, t) = \text{var}\{Z(s) - Z(t)\} = R_Z(s, s) + R_Z(t, t) - 2R_Z(s, t).$$

We write $D = [d_{ij}]$ for the dispersion matrix of Z_k, and $\hat{D} = [\hat{d}_{ij}]$ for an estimated dispersion matrix.

Estimation of the stationary covariance functions $R_i(h)$ can be done parametrically within windows centred on the points z_i. Write $\gamma(h; \theta_i)$ for the corresponding dispersion function parameterised by θ_i; see for instance Cressie (1993, pp. 85–6) for a list of parametric forms which could be adopted for this dispersion function. For any given form, various methods of estimation can be used. For example, writing $\partial i(l)$ for the set of sites lying in the smallest disc centred on z_i and containing at least l points, we obtain θ_i by a local nonlinear least squares fit minimising

$$\sum_{s_j, s_k \in \partial i(l)} \left\{ \frac{\hat{d}_{jk} - \gamma(s_j - s_k; \theta_i)}{\gamma(s_j - s_k; \theta_i)} \right\}^2 \tag{4}$$

with respect to θ_i.

After obtaining an estimate $\hat{\theta}_i$ of θ_i, we obtain values $\hat{c}_i(s)$, $\hat{c}_j(t)$, $\hat{C}_i$, $\hat{C}_j$ and $\hat{R}_{\delta_i}(s, t)$ in the above construction and our final estimated spatial covariance function is

$$\hat{R}^*(s, t) = \sum_{i,j} v_i(s) v_j(t) \hat{c}_i(s)^{\text{T}} \hat{C}_i^{-1} \hat{\Gamma} \hat{C}_j^{-1} \hat{c}_j(t) + \sum_i v_i(s)^{\frac{1}{2}} v_i(t)^{\frac{1}{2}} \hat{R}_{\delta_i}(s, t). \tag{5}$$

We suggest either choosing I and $z_1, \ldots, z_I$ as a set of grid points covering the design region, or choosing these parameters based on knowledge of likely sources of nonstationary behaviour in the problem at hand. In the examples we have looked at, performance of the method does not seem to be very sensitive to the parameter η, although η may be chosen objectively using some form of crossvalidation. For choosing the number of sites in the local fitting of the stationary models, we recommend using at least 10 sites in each window, although this can be increased if any numerical problems are experienced in fitting. Haas (1990a, b) considers adaptive ways of choosing a window size for fitting stationary models locally in some purely spatial problems.

Once a covariance function for W has been determined, computation of (5) is straightforward. The computational cost in evaluating the covariance function is dominated by the need to solve two linear systems in $\hat{C}_i$. If (5) is to be evaluated for a large number of pairs of sites, as is usually the case, then initial Cholesky decompositions of the $\hat{C}_i$'s could be obtained, enabling the required solution of the linear systems to be done efficiently. The initial decomposition is feasible for up to several hundred sites, which makes the method computationally attractive relative to alternative approaches.

As an illustration of our method we present a simple one-dimensional example where the results are easily visualised. An application of our method for real data is given in §4.

Let $Y_i(t)$ ($t \in \mathbb{R}$, $i = 1, 2$) be a pair of independent one-dimensional stationary Gaussian processes having mean zero and covariance functions

$$R_{Y_i}(h) = \exp(-\theta_i |h|),$$

where $\theta_1 = 2$ and $\theta_2 = 0.5$. We construct a nonstationary process $Y(t)$ as a weighted sum of $Y_1(t)$ and $Y_2(t)$ as follows:

$$Y(t) = \Phi(t)^{\frac{1}{2}} Y_1(t) + \{1 - \Phi(t)\}^{\frac{1}{2}} Y_2(t),$$

where $\Phi(t)$ denotes the standard normal distribution function. The covariance function of this process is

$$\mathrm{cov}\{Y(s), Y(t)\} = \Phi(s)^{\frac{1}{2}}\Phi(t)^{\frac{1}{2}} R_{Y_1}(t-s) + \{1-\Phi(s)\}^{\frac{1}{2}}\{1-\Phi(t)\}^{\frac{1}{2}} R_{Y_2}(t-s).$$

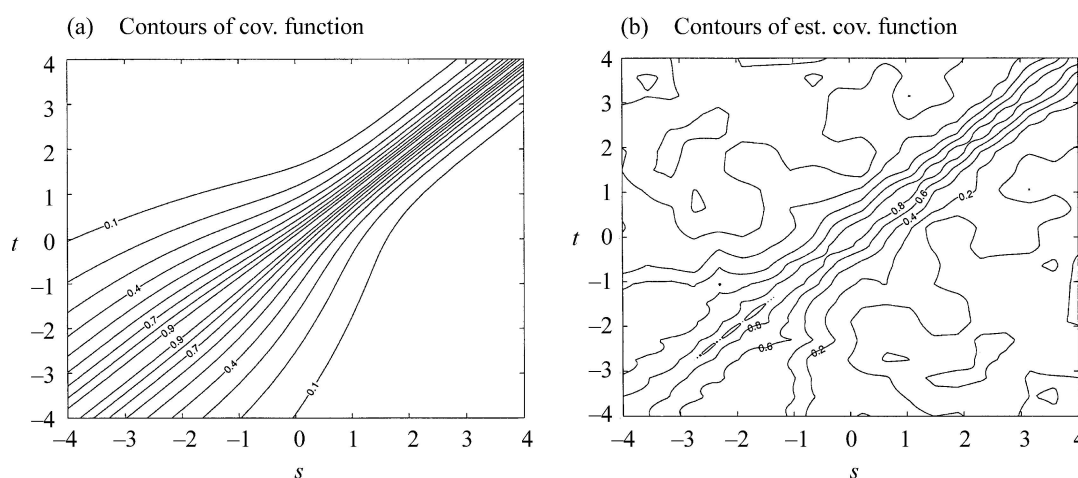A contour plot of this covariance function is shown in Fig. 1(a).

Fig. 1. (a) shows contour-plot of covariance function, cov{$Y(s)$, $Y(t)$}, of one-dimensional nonstationary process $Y(t)$; (b) shows contour-plot of estimate of covariance function cov{$Y(s)$, $Y(t)$}.

In Fig. 1(b) is shown a contour plot of an estimate of the covariance function obtained using our proposed method with $m = 2$ component stationary processes, $z_1 = -2$ and $z_2 = 2$. A stationary covariance function of the same form as for $Y_1(t)$, $Y_2(t)$ was used for the component stationary processes. The sites $s_1, \ldots, s_n$ here are taken to be 20 equally spaced points between -4 and 4. Our estimate $\hat{\Gamma}$ of the covariance matrix for $Y(t)$ at these sites was based on 50 simulated replicates. In the parametric local stationary fits the 10 nearest sites were used. The smoothing parameter was chosen to be $\eta = 1$.

4. OTHER APPROACHES

4·1. *Spatial deformation approach*

The idea of the Sampson–Guttorp approach to covariance estimation is to estimate a change of coordinates which results in isotropic spatial correlation. Let $\gamma_\theta(.)$ denote a valid isotropic dispersion function in $\mathbb{R}^2$ depending on a finite-dimensional parameter θ, and let f denote a coordinate transformation

$$f = (f_1, f_2) : \mathbb{R}^2 \to \mathbb{R}^2.$$

This transformation takes the monitoring sites in the geographical or G-space to points in a transformed space, dispersion or D-space. Sampson & Guttorp (1992) model $f_1(.)$ and $f_2(.)$ as thin-plate spline mappings, and after variance standardisation of the empirical dispersions estimate θ and f in such a way that

$$d_{ij} = \gamma_\theta \{ \| f(s_i) - f(s_j) \| \}. \tag{6}$$

Our implementation of this method follows the penalised weighted least squares approach to parameter estimation of Meiring, Guttorp & Sampson (1997). This requires selection of a smoothing parameter which controls a trade-off between smoothness of the spatial deformation and fidelity to the empirical dispersions.

This method's description of the covariance in terms of an isotropic model and spatial deformation mapping provides an elegant way of describing and visualising nonstationary spatial covariance structure. However, the method is not without its difficulties in practice as we indicate in the analysis of § 5. In particular, choice of smoothing parameter is not always straightforward; see Meiring, Monestiez et al. (1997) for a discussion of the difficulties here. Estimation of the spatial deformation also requires optimisation of a possibly multimodal objective function in a high-dimensional space. This makes it difficult to apply the method for large datasets. Another possible concern is that it is not possible to write any covariance function in the form (6). It is also not clear how multivariate covariance structure can be estimated in a natural way with this method. This question has been considered by Mardia & Goodall (1993), who suggest using the same spatial deformation for each component of a multivariate process.

4·2. *Loader & Switzer method*

The method of Loader & Switzer (1992) is used in § 5 for estimating the covariance matrix at observing sites. Loader & Switzer estimate the true spatial covariance matrix Γ by

$$\hat{\Gamma}_{LS} = \lambda \hat{\Gamma} + (1 - \lambda)\hat{C},$$

where $0 < \lambda < 1$ is a shrinkage parameter, and $\hat{C}$ is a covariance matrix obtained by fitting some parametric stationary covariance function model. The shrinkage parameter λ is chosen by an empirical Bayes argument: an inverse Wishart prior distribution is assumed for Γ, $\Gamma \sim IW\{(m - n - 1)C, m\}$. With this prior it can be shown that the posterior mean of Γ is

$$\lambda \hat{\Gamma} + (1 - \lambda)\hat{C},$$

where $\lambda = M/(M + m - n - 1)$. The marginal likelihood for the parameter m can be obtained by integrating out Γ and then an estimate of m, and hence of λ, can be obtained by maximum likelihood; see Loader & Switzer (1992) for further details.

For extension to a new site, $s*$, Loader & Switzer suggest estimating the vector of cross-covariances

$$[\text{cov}\{Z(s_1), Z(s*)\}, \ldots, \text{cov}\{Z(s_n), Z(s*)\}]^{\text{T}}$$

by

$$\hat{\Gamma}_{LS} \hat{C}^{-1} \hat{c}(s*), \tag{7}$$

where $\hat{c}(s*)$ is the vector of cross-covariances between $s*$ and the sites in $\mathscr{T}$ computed

with the stationary model used in constructing $\hat{C}$. To get the final estimated spatial covariance matrix it only remains to specify $\text{var}\{Z(s*)\}$. One simple estimate that can be used is the average of the diagonal elements of $\hat{\Gamma}_{LS}$. However, this choice may result in a final estimated covariance matrix which is not nonnegative definite. Furthermore, when the construction above is applied to add multiple sites sequentially the estimated covariance between pairs of sites not in $\mathscr{T}$ may not be invariant to the order in which those sites are added.

It is straightforward to show that, when our method is applied with Loader & Switzer's estimate for the sites in $\mathscr{T}$ and a single stationary process, that is $I = 1$, our estimate of the cross-covariances between sites in $\mathscr{T}$ and sites not in $\mathscr{T}$ agrees with that of Loader & Switzer. However, our method does not suffer from the potential lack of both positive definiteness and order invariance that can result from their proposed extension method.

5. Sydney wind patterns

We apply our techniques to some data on the east-west component of Sydney wind patterns. There are 45 observing stations, 38 maintained by the Australian Bureau of Meteorology, 5 by the Sydney Water Board and the remaining 2 by the New South Wales Department of Public Works. We do not consider the difficult problems of comparing the various networks, but assume that roughly comparable measurements are taken at all sites.

The work reported here was part of an effort to improve understanding of the behaviour of Sydney wind patterns. Of particular interest are sea-breeze effects driven by differential heating of the land and ocean. The behaviour of the sea breeze varies seasonally and throughout the day, and in what follows we consider only the east-west component of 3 p.m. wind measurements for September and October, 1997, so that there are 61 days of data in all. Our data consist of daily 3 p.m. measurements at the 45 stations. There are different numbers of missing observations at each of the stations. We have obtained a positive definite estimate of the spatial covariance matrix using the EM algorithm (Little & Rubin, 1987, Ch. 7) with an initial estimate of the covariance obtained by filling in missing values with site means. We then applied the technique of Loader & Switzer (1992) to shrink this raw estimate to a fitted stationary covariance matrix.

We fitted an exponential covariance function model

$$R(t-s) = R(0) \exp(-\theta|t-s|)$$

by estimating the variance as the average of the empirical variances across sites, giving an estimated variance of 58·04, and estimated the spatial dependence parameter θ by an ordinary least squares fit to the correlation matrix, giving an estimated θ of 1·115. We do not include a so-called nugget effect, see for instance Cressie (1993, pp. 112–9) for discussion, although this could be incorporated by preliminary adjustment of $\hat{\Gamma}$ if necessary. After fitting the stationary model, we computed the shrinkage parameter for Loader & Switzer's technique as $\hat{\lambda} = 0.88$. The final covariance matrix was then standardised to obtain a correlation matrix. The correlation matrix used for the computations below is available from statlib at http://lib.stat.cmu.edu.

The contour plots of Fig. 2 show points which have equal estimated correlation with two different fixed sites. One station was chosen to be on the coast, location 33·85°S,
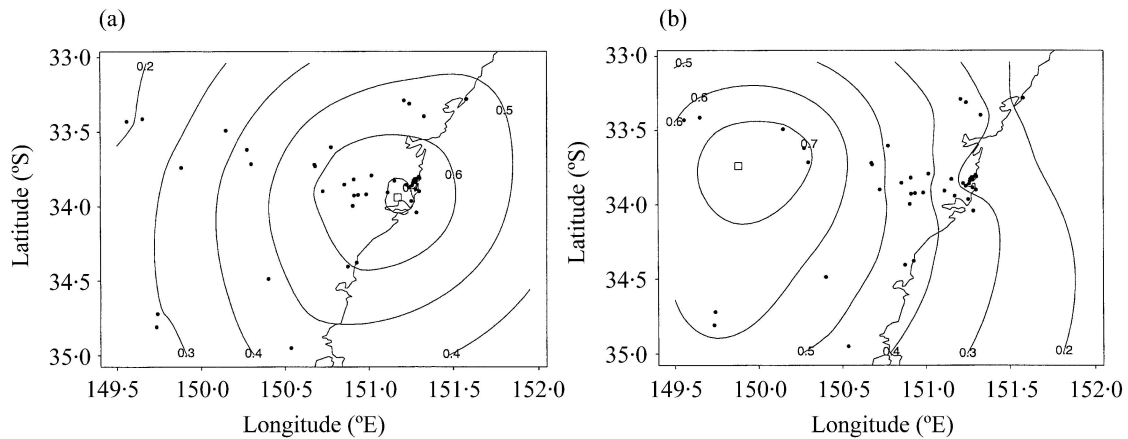
Fig. 2. Sydney wind pattern data. Contours of equal estimated correlation with two different fixed sites, shown by open squares: (a) location 33·85°S, 151·22°E, and (b) location 33·74°S, 149·88°E. The sites marked by dots show locations of the 45 monitored sites.
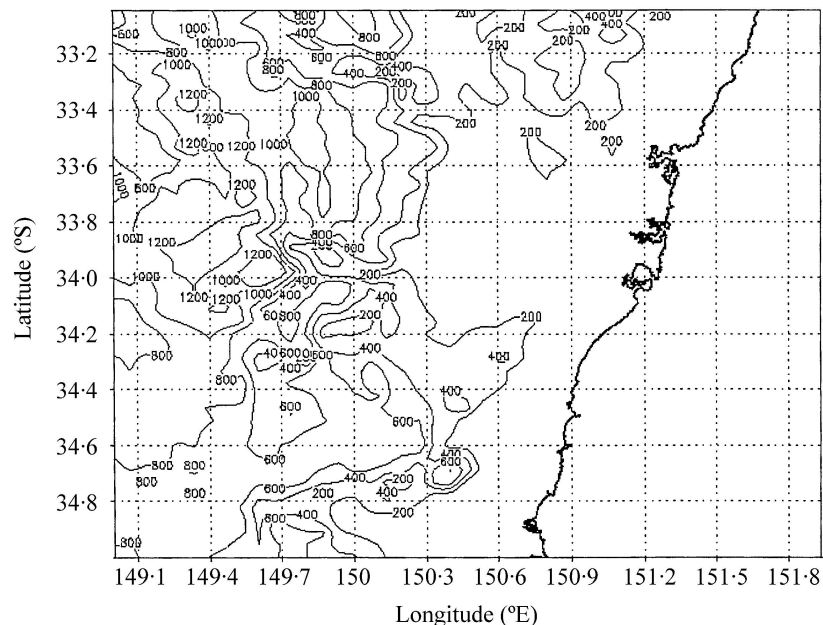

Fig. 3. Topography map of the Sydney area; relief contours show heights in metres.

151·22°E, and the other to be in the western half of the region under study, location 33·74°S, 149·88°E. The contour plots for the coastal station show some anisotropy that is thought to reflect sea breeze effects, with stations the same distance from the coast influenced to a similar extent. For the western reference station, correlation seems to decay more slowly with distance, and the shape of contours of high correlation can be interpreted in terms of the local topography with locations south of the reference site having a similar elevation. Figure 3 shows a topography map for an area which includes the study region.

For comparative purposes we have also implemented the penalised weighted least squares version of the spatial deformation method reviewed in § 4. We experienced some

difficulties in choosing the smoothing parameter λ controlling the trade-off between smoothness of the spatial deformation and fidelity to the empirical dispersions. If λ is chosen too small, then the spatial deformation may 'fold up', taking more than one point in the G-space to the same point in D-space, whereas as $\lambda \to \infty$ we obtain a stationary model in which contours of equal dispersion are elliptical. A number of sites, which were spatially anomalous with respect to their neighbours, appeared to have a large influence on the fitted spatial deformation. These were excluded from the spatial deformation analysis.

Figure 4 shows the results of applying the Sampson–Guttorp method to our data for two different values of λ, $\lambda = 7000$ and $\lambda = 10\,000$. The plots on the left show the fitted
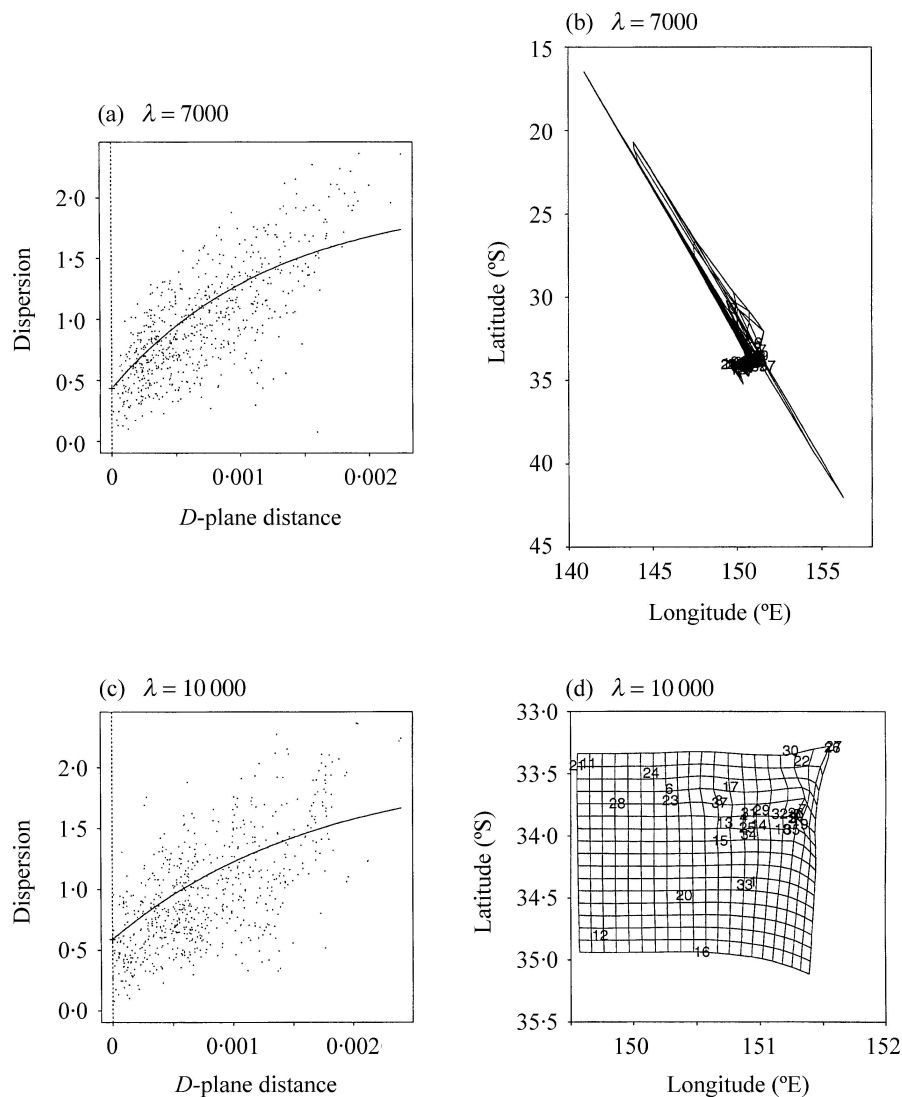

Fig. 4. Sydney wind pattern data. (a) and (c) show fitted exponential isotropic D-space dispersion functions plotted against inter-site distance in D-space. (b) and (d) show spatial deformations of regular grid for Sampson–Guttorp method; D-plane locations of monitoring sites are indicated by numbers. (a) and (b) are for smoothing parameters, λ, of 7000; (c) and (d) are for $\lambda = 10\,000$.

isotropic dispersion functions in D-space and the empirical dispersions plotted against inter-site distance in D-space. The plots on the right show the action of the spatial deformation on a regular grid. There is a very sharp transition as λ increases between a deformation which is severely folded and a deformation in which the sites in D-space are not moved from their geographic locations. In this latter case the resulting estimated field is very nearly stationary.

6. Discussion

The method we have described is appealing as a simple way of obtaining a valid nonnegative definite covariance function respecting a given covariance matrix. We anticipate that it may be useful for visualising covariance structure for large datasets. Extension of our methods to assessment of covariance for multivariate processes is formally straightfoward, as shown in a technical report, from the Department of Statistics, University of New South Wales, by D. J. Nott, W. T. M. Dunsmuir, M. Speer and T. Glowacki, who describe an application to bivariate components of wind in the Sydney region.

Another question to be addressed in future work is the application of our methods to problems of spatial prediction. It is hoped that more realistic estimators of nonstationary spatial covariance will result in improved spatial prediction methods and in improved assessments of predictive uncertainty.

References

Adler, R. J. (1982). *The Geometry of Random Fields*. New York: Wiley.

Brown, P. J., Le, N. D. & Zidek, J. V. (1994). Multivariate spatial interpolation and exposure to air pollutants. *Can. J. Statist.* **22**, 489–505.

Cressie, N. A. C. (1993). *Statistics for Spatial Data*, 2nd ed. New York: Wiley.

Damian, D., Sampson, P. D. & Guttorp, P. (2001). Bayesian estimation of semi-parametric non-stationary spatial covariance structures. *Environmetrics* **12**, 161–78.

Guttorp, P. & Sampson, P. D. (1994). Methods for estimating heterogeneous spatial covariance functions with environmental applications. In *Handbook of Statistics XII: Environmental Statistics*, Ed. G. P. Patil and C. R. Rao, pp. 663–90. New York: Elsevier/North Holland.

Haas, T. C. (1990a). Kriging and automated variogram modeling within a moving window. *Atmosph. Envir.* **24A**, 1759–69.

Haas, T. C. (1990b). Lognormal and moving window methods of estimating acid deposition. *J. Am. Statist. Assoc.* **85**, 950–63.

Hass, T. C. (1995). Local prediction of a spatio-temporal process with an application to wet sulfate deposition. *J. Am. Statist. Assoc.* **90**, 1189–99.

Haas, T. C. (1996). Multivariate spatial prediction in the presence of nonlinear trend and covariance nonstationarity. *Environmetrics* **7**, 145–65.

Haas, T. C. (1998). Statistical assessment of spatio-temporal pollutant trends and meteorological transport models. *Atmosph. Envir.* **32**, 1865–79.

Higdon, D. (1998). A process-convolution approach to modeling temperatures in the North-Atlantic Ocean. *Envir. Ecol. Statist.* **5**, 173–90.

LITTLE, R. J. A. & RUBIN, D. B. (1987). *Statistical Analysis with Missing Data.* New York: Wiley.

LOADER, C. & SWITZER, P. (1992). Spatial covariance estimation for monitoring data. In *Statistics in Environmental and Earth Sciences*, Ed. A. Walden and P. Guttorp, pp. 52–70. London: Edward Arnold.

MARDIA, K. V. & GOODALL, C. R. (1993). Spatial-temporal analysis of multivariate environmental monitoring data. In *Multivariate Environmental Statistics*, Ed. G. P. Patil and C. R. Rao, pp. 347–86. London: Elsevier.

MEIRING, W., GUTTORP, P. & SAMPSON, P. D. (1997). Computational issues in fitting spatial deformation models for heterogeneous spatial correlation. In *Proceedings of the 29th Symposium on the Interface: Computing Science and Statistics*, **29**(1), Ed. D. W. Scott, pp. 409–17. Fairfax Station, VA: Interface Foundation of North America, Inc.

MEIRING, W., MONESTIEZ, P., SAMPSON, P. D. & GUTTORP, P. (1997). Developments in the modelling of nonstationary spatial covariance structure from space-time monitoring data. In *Geostatistics Wollongong '96*, **1**, Ed. E. Y. Baafi and N. Schofield, pp. 162–73. Dordrecht: Kluwer Academic Publishers.

NYCHKA, D. & SALTZMAN, N. (1998). Design of air quality networks. In *Case Studies in Environmental Statistics*, Ed. D. Nychka, W. Piegorsch and L. Cox, pp. 51–76. New York: Springer-Verlag.

OBLED, C. & CREUTIN, J. D. (1986). Some developments in the use of empirical orthogonal functions for mapping meteorological fields. *J. Appl. Meteor.* **25**, 1189–204.

OEHLERT, G. W. (1993). Regional trends in sulfate wet deposition. *J. Am. Statist. Assoc.* **88**, 390–9.

SAMPSON, P. D. & GUTTORP, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *J. Am. Statist. Assoc.* **87**, 108–19.

SAMPSON, P. D., DAMIAN, D. & GUTTORP, P. (2001). Advances in modelling and inference for environmental processes with nonstationary spatial covariance. In *GeoENV 2000: Geostatistics for Environmental Applications*, Ed. P. Monestiez, D. Allard and R. Froidvaux, pp. 17–32. Dordrecht: Kluwer Academic Publishers.

[*Received November* 2000. *Revised May* 2002]