

A REVIEW OF NON-STATIONARY SPATIAL METHODS FOR GEODETIC LEAST-SQUARES COLLOCATION

N. Darbeheshti (corresponding author)

Western Australian Centre for Geodesy and The Institute for Geoscience Research,
Curtin University of Technology, GPO Box U1987, Perth 6845, Australia

Fax: +61 (0)8 9266 2703; Email: neda.darbeheshti@gmail.com

W.E. Featherstone

Western Australian Centre for Geodesy and The Institute for Geoscience Research,
Curtin University of Technology, GPO Box U1987, Perth 6845, Australia

Fax: +61 (0)8 9266 2703; Email: W.Featherstone@curtin.edu.au

ABSTRACT

This paper reviews a field that is herein termed spatial “non-stationarity”, which is specifically concerned with non-stationarity in the geodetic theory of least-squares collocation (LSC). In practice, many geodesists rely on stationary assumptions in LSC, i.e., using a constant mean and isotropic and spatially invariant covariance for estimation and prediction of geodetic quantities. However, new theories in spatial statistics and geostatistics allow for better statistical methodologies to be used in geodesy. The aim of this paper is to introduce these methodologies and adapt them for dealing with non-stationarity in LSC.

Keywords: geodesy, least-squares collocation, geostatistics, spatial statistics, non-stationarity

INTRODUCTION

Stationarity is an in-built assumption in [standard] geodetic least-squares collocation (LSC), where the mean value of a dataset is taken as a constant and the covariance function is taken to be both isotropic (uniformly shaped in all directions) and spatially invariant (e.g., Moritz, 1980). Non-stationarity, on the other hand, means that the mean value of a dataset is not necessarily constant and/or the covariance is anisotropic (varies with direction) and spatially variant. In the discipline of spatial statistics, for instance, this is called weak or second-order non-stationarity (Armstrong, 1998).

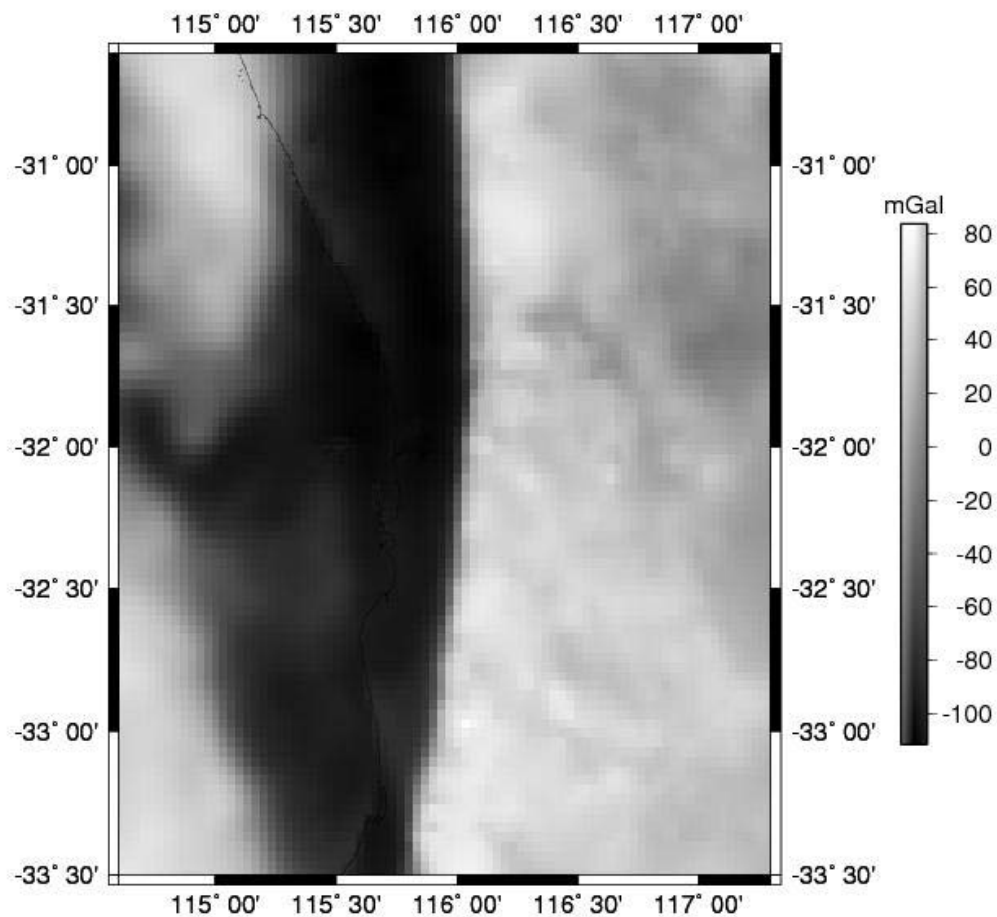


Figure1. Gravity anomalies over the Darling Fault in Western Australia: to a first approximation, the region is divided between low gravity anomalies to the west, and high gravity anomalies to the east, which shows the spatial non-stationarity in the dataset.

Gravity anomalies (differences between observed and model values) over the Darling Fault in Western Australia (Lambeck, 1987) provide a nice illustration of non-stationarity in geodetic data (Figure 1). It is caused by the mass-density contrast of $-0.4 \text{ kg} / \text{m}^3$ (Boschetti et al., 1997) across the fault. Using a single covariance structure across the whole dataset, for instance to interpolate the spatial data, will not be representative. Thus, there is a necessity to develop non-stationary covariance functions (i.e., ones that are spatially variable) to better predict the spatial random field (SRF) among dissimilar areas (cf. Darbeheshti and Featherstone, 2009, 2010).

Research disciplines other than geodesy, mostly environmental studies (e.g., Nychka and Saltzman, 1998; Fuentes, 2001; Lloyd and Atkinson, 2002), have studied the problem of non-stationarity in spatial or time-series data. Among these, the disciplines of geostatistics and spatial statistics are more compatible with geodetic LSC theory (cf. Moritz, 1980). Regardless of the dataset involved, the fundamental principles and concepts in one research field have frequently been applied to problems in other research fields. This paper will introduce non-stationary approaches to best linear unbiased estimation (BLUE) problems, some from geodesy, but most from geostatistics and spatial statistics.

The SRF estimation and prediction scheme successfully implemented in geodesy is LSC, which can be used to: (i) account for systematic effects in the data (trends); (ii) predict the SRF between data points (interpolation); and estimate the SRF at other points (filtering). The generalised model of LSC in geodesy is given by

$$\mathbf{l} = \mathbf{A} \mathbf{x} + \mathbf{y} + \mathbf{n} \quad (1)$$

where \mathbf{l} is the vector of observations, \mathbf{A} is the design matrix with full column rank, \mathbf{x} is the vector of unknown parameters, \mathbf{y} is the predicted signal vector, and \mathbf{n} is the error vector of observations. The BLUE solution of Eq. (1) is obtained by (Moritz, 1980)

$$\hat{\mathbf{x}} = (\mathbf{A}^T (\mathbf{C}_{ll} + \mathbf{C}_{nn})^{-1} \mathbf{A})^{-1} \mathbf{A}^T (\mathbf{C}_{ll} + \mathbf{C}_{nn})^{-1} \mathbf{l} \quad (2)$$

$$\hat{\mathbf{y}} = \mathbf{C}_{yl}(\mathbf{C}_{ll} + \mathbf{C}_{nn})^{-1}(\mathbf{I} - \mathbf{A}\hat{\mathbf{x}}) \quad (3)$$

$$\hat{\mathbf{n}} = \mathbf{C}_{nn}(\mathbf{C}_{ll} + \mathbf{C}_{nn})^{-1}(\mathbf{I} - \mathbf{A}\hat{\mathbf{x}}) \quad (4)$$

where $\hat{\cdot}$ refers to the LSC-estimated quantity, \mathbf{C}_{nn} is the variance-covariance matrix of the noise (a diagonal matrix), \mathbf{C}_{ll} is the auto-covariance matrix of the vector of the observations \mathbf{l} , and \mathbf{C}_{yl} is the cross-covariance matrix between observation \mathbf{l} and prediction \mathbf{y} .

In standard LSC, it is assumed that the observation \mathbf{l} and prediction \mathbf{y} have an expected value equal to zero: i.e., $E\{\mathbf{l}\} = 0$ and $E\{\mathbf{y}\} = 0$, with the expectation $E\{\cdot\}$ being the average or mean value in the sense of probably theory. As can be seen from Eq. (3), the prediction is a function of not only the auto-covariances \mathbf{C}_{ll} of the observed quantities, but also the cross-covariances \mathbf{C}_{yl} of the predicted quantity with the observed quantity. As such, LSC can take observations with errors assigned, and then generate quantities at the same or other points with an error estimate for those predicted values.

This review paper covers the substantial literature dealing with non-stationarity in spatial data. The aim is to select, from a wide range of options, those that are most applicable to geodetic LSC. The sources are classified among three categories: geodesy, geostatistics and spatial statistics. The advantages and disadvantages of each method are presented, and the justification given as to why the kernel convolution method of Higdon et al. (1999) from spatial statistics should be chosen for non-stationary covariance modelling in geodetic LSC. A simulated numerical example is given to demonstrate the effectiveness of the non-stationary methods, when applied to non-stationary data.

NON-STATIONARY BLUE SOLUTIONS FROM GEODESY

Trend removal

Trend removal is the most common approach in geodetic LSC to deal with non-stationarity in the mean of geodetic data: “*The LSC solution is giving the minimum mean square error in a very specific sense, namely as the mean over all data-configurations which by a rotation of the Earth's centre may be mapped into each other. So if this should work locally, we must make all areas of the Earth look alike, seen from the gravity field standpoint*” (Tscherning, 1994).

In geodesy, this is often known and implemented as the remove-restore method, and involves removing as much as we know about the input data, and later adding it back to get the final result, thereby using a SRF that is statistically more homogeneous than before. In the LSC procedure for gravity field modelling, first, the contribution from a high degree (typically 360) spherical harmonic expansion is removed. Secondly, the effect of the local topography is reduced. This leaves a residual field, with a smoothness in terms of standard deviation of between 50% and 25% less than the original signal (Tscherning, 1994).

Other methods of trend removal have been tried in the geodetic literature, which vary depending upon the application; these are:

- Equation (2) can be used for modelling a trend in LSC (Moritz, 1980);
- Some authors (e.g., Tscherning, 1994) simply subtract the mean value as a trend from the data, which basically comes from geostatistics. [However, Cressie (1993) suggests using the median instead, because by removing the mean, there is a danger of adding a bias];
- Goad et al. (1984) use a linear polynomial (tilted plane) to remove the trend from Bouguer gravity anomalies over the continental United States;

- Stopar et al. (2006) present a method employing the artificial neural network approximation to obtain a trend surface in LSC for geoid determination, although Tscherning (2006) disagrees with this;
- In the case of coordinate transformations by LSC, the residual is regarded as a distortion that remains after the application of a conformal datum transformation (e.g., Ruffhead, 1987; Collier et al., 1997, 1998; You and Hwang, 2006).

Trend removal has some disadvantages, however:

- It adds extra steps to the computation: first the trend is removed from the input data and after the interpolation or prediction it is added back to the result;
- LSC is based on a zero-mean assumption in the input data (Moritz, 1980), but there is no guarantee that this condition is satisfied after trend removal;
- Trend removal itself may introduce errors and biases into the input data (Cressie, 1993).

Riesz representers

Tscherning (1999) proposed a non-stationary geodetic covariance model based on replacing stationary degree variances of the anomalous potential T in Eq. (5) with non-stationary degree variances. From Tscherning and Rapp (1974), the covariance function $K(P,Q)$ of the anomalous potential T at points P and Q has been chosen as the basic covariance function $K(T_P, T_Q)$ that has the solid scalar spherical-harmonic expansion of

$$K(T_P, T_Q) = \sum_{l=0}^{\infty} \sigma_l(T, T) \left(\frac{R^2}{r_P r_Q} \right)^{l+1} P_l(\cos \psi) \quad (5)$$

where σ_l are the degree variances of the anomalous potential, P_l are Legendre polynomials, R is the radius of the Bjerhammar (1964) sphere, and r_P, r_Q are the geocentric radii to points $P(\phi_P, \lambda_P)$ and $Q(\phi_Q, \lambda_Q)$, which are separated by the spherical distance ψ (Heiskanen and Moritz, 1967):

$$\psi = \cos^{-1} (\cos \phi_P \cos \phi_Q \cos(\lambda_P - \lambda_Q) + \sin \phi_P \sin \phi_Q) \quad (6)$$

This approach uses heterogeneous sets of basis functions associated with point masses, where the point masses are buried at varying depths covering the whole Earth (cf. Barthelmes et al., 1991; Vermeer, 1995).

The finite set of functions is linearly independent because they may be regarded as a set of Riesz representers (Tscherning, 1984) of the evaluation functionals associated with the point masses. Riesz representers exist for all linear functionals in separable Hilbert spaces. However, for the reproducing-kernel Hilbert space (RKHS), the inner product of the representer and an arbitrary function gives the value of the quantity represented by it. A RKHS of functions, which are harmonic (i.e., satisfy Laplace's equation) in the set outside a sphere with radius R_0 , having a reproducing kernel K_0 (equal to the covariance function), is considered. The degree variances of this kernel are denoted by σ_{0l} .

The set of Riesz representers associated with the evaluation functionals related to distinct points $P_l, l = 1, \dots, L$, on a 2D surface surrounding the bounding sphere will be linearly independent. These functions are used to define a new L -dimensional RKHS with kernel $a_{l>0}$

$$K_l(P, Q) = \sum_{l=1}^L K_0(P_l, P) \cdot K_0(P_l, Q) \cdot a_l \quad (7)$$

with P, Q and P_l being points in the set of harmonicity. If all the points are located on a concentric Bjerhammar (1964) sphere with radius $R_1 > R_0$, and form a net covering the sphere, and a_l are area elements (depending on L), then this kernel will converge towards an isotropic kernel with degree variances given by

$$\sigma_l^2 = (2l + 1) \sigma_{0l}^2 \cdot \left(\frac{R_0}{R_1}\right)^{2l+1} \cdot (\text{constant}) \quad (8)$$

If $K_l(P, Q)$ is required to represent an isotropic covariance function $C(P, Q)$, σ_{0l} can be selected so that σ_l is equal to the empirical degree variances. If the points are chosen at varying radial distances $R_1 > R_0$, then an anisotropic covariance function can be constructed.

This approach effectively introduces non-stationary covariance models on a global scale, but it first needs to find the optimum position of point masses to match the external gravity field, which introduces the complication of non-uniqueness (any set of point masses can be selected to generate the same external gravity field). Principally because of this, Riesz representers have never been applied practically in geodesy.

Wavelet approaches

The space-localisation properties of wavelets (e.g., Daubechies, 1992) make them an efficient and useful tool for spectral studies of non-stationary signals, and have attracted three principal studies in geodesy:

Kotsakis and Sideris (1999) show that the method of spatio-statistical (non-probabilistic) collocation, expressed by the optimal estimation criterion and the translation-invariance condition, leads to signal-approximation models similar to those encountered in multiresolution analysis (MRA) theory. The classical MRA formalism (Mallat, 1989) lies at the very core of some of the approximation principles traditionally used in geodetic problems.

Kotsakis (2000) shows that the use of a spatio-statistical minimum mean square error criterion, for linear estimation of deterministic signals, always gives a generalised MRA in the Hilbert space $L^2(R)$, under some mild constraints on the covariance function and the power spectrum of the SRF under consideration. Using the theory and approximation algorithms associated with statistical collocation, a constructive (frequency-domain-based) framework for building generalised MRA in $L^2(R)$ was presented, without the need of the usual dyadic restriction that exists in classical wavelet theory. Although Kotsakis's (2000) work introduced a wavelet framework for non-stationary LSC, there is no practical computation algorithm, so it has not been used in geodesy.

In a different wavelet approach, Keller (1998, 2000, 2002) gives a numerical solution based on the Haar (1910) wavelet's equivalence to the Wiener-Kolmogorov equations in stationary LSC (cf. Brovelli et al., 2003). Keller's wavelet approach solves the problem of filtering non-stationary errors from a stationary signal by LSC, i.e., when the variance of the data errors differs in different areas. Filtering of stationary errors is classically solved by Wiener-Kolmogorov equations using the discrete Fourier transform (DFT). For non-stationary errors, the equations cannot be transformed into the frequency domain and solved by DFT. As such, Keller's approach has a somewhat limited application, with it only being applied to the filtering of non-stationary errors from stationary signal by LSC.

NON-STATIONARY BLUE SOLUTIONS FROM GEOSTATISTICS

Locally adaptive Kriging

Non-stationary models of the spatial mean have been applied in geostatistics for many years (e.g., Wackernagel, 2003). One of the most useful methods is the adaption of ordinary Kriging (OK) to account for non-stationarity of the mean, which was introduced by Deutsch and Journel (1998) as part of the GSLIB software.

OK amounts to re-estimating, at each new location \mathbf{s} , the mean m as used in the simple Kriging (SK) expression. Since the only difference between SK and LSC is that SK assumes that the mean is known, while LSC is based on the zero-mean assumption of the observation vector (Moritz, 1980), the LSC Eq. (3) is recalled for SK:

$$\hat{\mathbf{y}} = \mathbf{C}_{y|l} \mathbf{C}_{ll}^{-1} \mathbf{l} \tag{9}$$

Because OK is most often applied within moving search neighbourhoods (Deutsch and Journel, 1998), i.e., using different datasets for different locations \mathbf{s} , the implicit re-estimated mean depends on the location \mathbf{s} . Thus, the OK estimator of

$$\hat{\mathbf{y}} = \mathbf{C}_{y|l} \mathbf{C}_{ll}^{-1} \mathbf{l} + (1 - \sum(\mathbf{C}_{y|l} \mathbf{C}_{ll}^{-1}))m(\mathbf{l}) \tag{10}$$

is SK, where the constant mean value m is replaced by the location-dependent estimate.

If $m(\mathbf{I})=0$, all formulations of LSC, SK and OK (Eqs. (3), (9) and (10)) degenerate to the same case. Hence, OK as applied within moving data neighbourhoods is a partly non-stationary algorithm, in the sense that it corresponds to a non-stationary SRF model with varying mean, but a stationary covariance. This ability to locally rescale the SRF model to a different mean explains the robustness of the OK algorithm (Chilès and Delfiner, 1999).

The idea of using neighbouring data is derived from Kriging. Most Kriging algorithms consider a limited number of nearby conditioning data. The first reason for this is to limit the CPU and computer memory requirements. Furthermore, adopting a global search neighbourhood would require knowledge of the covariance for the largest separation distance between data points. The covariance is typically poorly determined for distances beyond around one-half or one-third of the size of a study area. Another reason for a limited search neighbourhood is to allow for local rescaling of the covariance parameters for each computation point (Deutsch and Journel, 1998).

Segmentation

Non-stationary covariance modelling in Kriging is usually accounted for by developing local variograms and performing a piecewise interpolation (e.g., Atkinson and Lloyd, 2007). In other words, it involves dividing the region of interest into smaller segments within which the covariance function is assumed stationary. [Segmentation is also used indirectly in geodesy: Tscherning et al. (1987) considered segmentation for merging regional geoid models, and Knudsen (2005) used segmentation for satellite radar altimeter data processing.]

There are some problems with the segmentation approach, however. In some subdivisions, the data may not be sufficiently dense to properly estimate the local covariance functions. It is difficult to justify which covariance parameters should be chosen at the

boundary of two regions. Segmentation methods also need to subsequently patch the covariances or results at the boundaries of the sub-regions, which will cause edge effects resulting from discontinuities in the statistical parameters at the borders of neighbouring areas (cf. Knudsen, 2005; Tscherning et al. 1987). As such, segmentation is not always such an attractive option for use in non-stationary LSC.

NON-STATIONARY BLUE SOLUTIONS FROM SPATIAL STATISTICS

There is a huge body of literature in spatial statistics on methods for modelling non-stationarity, which has never been pointed out in geodesy before. A number of modelling and inference methods were introduced in the early 1990s, beginning with Sampson and Guttorp's (1992) spatial deformation approach. The majority of literature concerns methods that are semi-parametric: they are non-parametric with respect to the way that the spatial variation in covariance is described, but the local covariance structure is described by conventional parametric models. Much of this literature discusses Bayesian modelling strategies that enable the uncertainty in the covariance structure to be reflected in the estimation and prediction.

This section gives a review of the non-stationarity methods from spatial statistics, which should only be considered as a quick introduction to each, with the aim of opening up a new path in statistical geodesy. The pros and cons of each method with regards to their application to LSC in geodesy are discussed, to finally justify why the kernel convolution method of Higdon et al. (1999) should be chosen for non-stationary modelling of covariances in geodetic LSC (cf. Darbeheshti and Featherstone, 2009, 2010).

Spatial deformation models

Sampson and Guttorp (1992) introduced an approach to handling non-stationarity through spatial deformation. Instead of the original non-stationary space G , they define the correlation function by reference to a latent (transformed) space D , where stationarity holds.

Suppose that temporally independent samples $Z_{it} = Z(\mathbf{x}_i, t)$ are available at N sites, $\mathbf{x}_i, i = 1, \dots, N$, typically in \mathbb{R}^2 and at T points in time $t = 1, \dots, T$. $\mathbf{x} = [X_1 \ X_2]$ represents the matrix of locations. The underlying spatio-temporal process is written as

$$Z(\mathbf{x}, t) = \mu(\mathbf{x}, t) + v(\mathbf{x})^{1/2} E_t(\mathbf{x}) + \varepsilon(\mathbf{x}, t), \quad (11)$$

where $\mu(\mathbf{x}, t)$ is the mean field, $v(\mathbf{x})$ is a smooth function representing spatial variance, and $E_t(\mathbf{x})$ is a zero-mean, variance-one, second-order continuous Gaussian spatial process, i.e. $C(E_t(\mathbf{x}), E_t(\mathbf{y})) \rightarrow 1$ as $\mathbf{x} \rightarrow \mathbf{y}$. $\varepsilon(\mathbf{x}, t)$ represents measurement error and/or very short-scale spatial structure (in comparison with the spatio-temporal process being modelled), which is assumed to be both Gaussian and independent of E_t .

The correlation structure of the spatial process is expressed as a function of Euclidean distances between site locations after a bijective (“one-to-one” and “onto”) transformation of the geodetic coordinate system,

$$R(E_t(\mathbf{x}), E_t(\mathbf{y})) = R_\theta(\|f(\mathbf{x}) - f(\mathbf{y})\|), \quad (12)$$

where $f(\cdot)$ is the one-to-one transformation that expresses the non-stationarity, and R_θ belongs to a parametric family with unknown parameters θ .

For mappings from one \mathbb{R}^2 onto another \mathbb{R}^2 , the geodetic coordinate system has been called the “ G -plane” and the space representing the images of these coordinates under the transformation the “ D -plane”. Perrin and Meiring (1999) prove that this spatial deformation model is identifiable for mappings from one \mathbb{R}^k to another \mathbb{R}^k assuming only differentiability of the isotropic correlation function R_θ .

Thus, from Eq. (12) there are two unknown functions to estimate: f and R_θ . The latter is a parametric choice from a standard class of covariance functions. To determine the former is a challenging fitting problem of choosing a class of transformations and to obtain the best member from this class. Sampson and Guttorp (1992) employ thin plate splines and optimise a version of a 2D non-metric multi-dimensional scaling criterion, providing an algorithmic solution. However, this solution is generally not well behaved, in the sense that f will be bijective, often folding over itself. Smith (1996) embedded this approach within a likelihood setting, but worked instead with radial basis functions.

Damian et al. (2001) formulated a Bayesian approach to implement Eq. (11). They still work with thin plate splines, but place priors over an identifiable parameterisation, which depends upon the number of points, n , being transformed. They elected not to model f directly, but instead model the transformed locations. The set of n -transformed locations are modelled as n realisations from a bivariate Gaussian spatial process and a prior is placed on the process parameters. That is, $f(\mathbf{x})$ arises as a random realisation of a bivariate process at \mathbf{x} rather than the value at \mathbf{x} of a random bivariate transformation.

Figure 2 presents an illustration of the G -plane and D -plane for rainfall observation sites in southern France, as presented by Damian et al. (2001).

The limitations of these spatial deformation approaches are:

- The implementation requires independent replications of the process in order to obtain an estimated sample covariance matrix. In practice, such replications of a spatial process are rarely obtained. If a repeated measurement is obtained at a particular location, it is typically collected over time (e.g., satellite remote sensing);
- Spatial deformation techniques, like trend removal, allow for a reduction to a stationary covariance structure; i.e., they are not actually methods to specifically model non-stationarity.

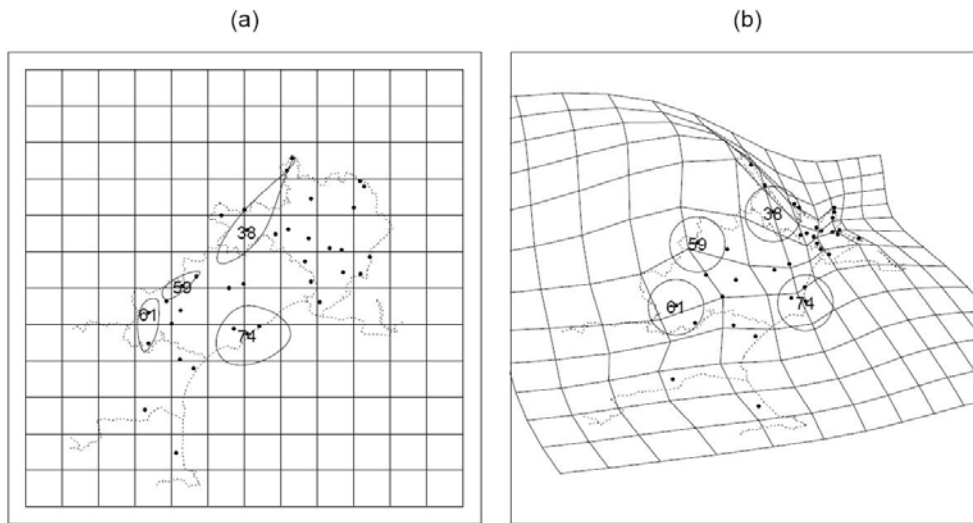


Figure 2: An illustration of (left) the G -plane (original surface) and (right) D -plane (transformed surface) for the spatial deformation method (from Damian et al. (2001))

Kernel smoothing of empirical covariance matrices

Perhaps the simplest approach to deal with non-stationarity begins either from the perspective of local stationary models, which are empirically smoothed over space, or from the perspective of the smoothing and/or interpolation of empirical covariances estimated among a finite number of observation sites. Fuentes (2001) and Nott and Dunsmuir (2002) propose approaches for representing non-stationarity in terms of spatially weighted combinations of stationary covariance functions to represent the local covariance structure in different regions.

There is a distinct difference between kernel smoothing and the segmentation method (described earlier under the geostatistics section). In kernel smoothing, the covariance function is constructed based on segmented covariances, and eventually one single LSC or Kriging operation is performed for the whole region.

Fuentes's approach: This involves dividing the spatial domain D into k subregions S_i , each with a sufficient number of points to estimate a stationary covariance function K_{θ_i} .

Fuentes (2001) represents the spatial process $Z(\mathbf{x})$ (defined over the entire region) as a weighted average of orthogonal local stationary processes $Z_i(\mathbf{x})$:

$$Z(\mathbf{x}) = \sum_{i=1}^k \omega_i(\mathbf{x}) Z_i(\mathbf{x}) \quad (13)$$

where $\omega_i(\mathbf{x})$ is the weight function, such as the inverse-squared distance between \mathbf{x} and the centre of subregion S_i , and $Z_i(\mathbf{x})$ denotes a spatial process with the covariance function K_{θ_i} .

The non-stationary spatial covariance structure is given by

$$\begin{aligned} C(Z(\mathbf{x}), Z(\mathbf{y})) &= \sum_{i=1}^k \omega_i(\mathbf{x}) \omega_i(\mathbf{y}) C(Z_i(\mathbf{x}), Z_i(\mathbf{y})) \\ &= \sum_{i=1}^k \omega_i(\mathbf{x}) \omega_i(\mathbf{y}) K_{\theta_i}(\mathbf{x} - \mathbf{y}) \end{aligned} \quad (14)$$

Fuentes (2001) chooses the number of subgrids, k , using a BIC (Bayes Information Criterion) as a weight function. The stationary processes $Z_i(\mathbf{x})$ are actually *local* only in the sense that their corresponding covariance functions, $K_{\theta_i}(\mathbf{x} - \mathbf{y})$, are estimated locally, and they are orthogonal by assumption only, so as to represent the overall non-stationary simply as a weighted sum of covariances. Fuentes (2001) estimates the parameters within the context of Bayesian estimation with predictive distributions accounting for uncertainty in the parameter estimates.

Fuentes and Smith (2001) extend the finite decomposition of $Z(\mathbf{x})$ of Fuentes (2001) to a continuous convolution of local stationary processes. It is an alternative kernel-based approach in which the process is taken to be the convolution of a fixed kernel over independent stationary processes, $Z_{\theta(u)}(\cdot)$

$$Z(\mathbf{x}) = \int K(\mathbf{x} - u) Z_{\theta(u)}(\mathbf{x}) du. \quad (15)$$

The resulting covariance is expressed as

$$C(\mathbf{x}_i, \mathbf{x}_j) = \int K(\mathbf{x}_i - u) K(\mathbf{x}_j - u) C_{\theta(u)}(\mathbf{x}_i - \mathbf{x}_j) du. \quad (16)$$

For each u , $C_{\theta(u)}(\cdot, \cdot)$ is a covariance function with parameters $\theta(u)$, where $\theta(u)$ is a multivariate spatial process that induces non-stationarity in $Z(\cdot)$. Estimation and prediction

will require that the SRF of parameter vectors $\theta(u)$, indexing the stationary Gaussian processes, be constrained to vary smoothly.

This method has the advantage of avoiding the need to parameterise smoothly varying positive-definite matrices, as required in the Higdon et al. (1999) Gaussian kernel approach (described later). One drawback is the lack of a general closed-form for $(\mathbf{x}_i, \mathbf{x}_j)$ and the need to compute covariances by Monte Carlo integration. The latter is of particular concern because of the numerical sensitivity of covariance matrices, where the inverse of the covariance matrices used in LSC could become ill-conditioned or even singular. In addition to Bayesian methods, Fuentes and Smith (2001) and Fuentes (2001) describe spectral methods for fitting models when the data are (nearly) on a grid; these may be much faster than likelihood methods (Chilès and Delfiner, 1999).

Nott and Dunsmuir's approach: This has the stated aim of reproducing an empirical covariance matrix at a set of observation sites and describing the conditional behaviour given observation site values with a collection of stationary processes. The same notation as above is used here, though for Nott and Dunsmuir (2002), i indexes the observation sites rather than a smaller number of subregions, and K_{θ_i} represents local residual covariance structure after conditioning from values at the observation sites. These are derived from locally fitted stationary models.

In their general case, Nott and Dunsmuir's (2002) representation of the spatial covariance structure can be written as

$$C(Z(\mathbf{x}), Z(\mathbf{y})) = \sum_{i=0}(\mathbf{x}, \mathbf{y}) + \sum_{i=1}^k \omega_i(\mathbf{x})\omega_i(\mathbf{y})K_{\theta_i}(\mathbf{x} - \mathbf{y}) \quad (17)$$

where $\sum_{i=0}(\mathbf{x}, \mathbf{y})$ is a function of the empirical covariance matrix at the observation sites, $C=[c_{ij}]$, and the local stationary models computed so that $C(Z(\mathbf{x}_i), Z(\mathbf{y}_i)) = c_{ij}$. This exact interpolation is relaxed by replacing the empirical covariance matrix C by the empirical Bayesian shrinkage estimator.

While the models introduced by Fuentes (2001) and Nott and Dunsmuir (2002) look similar (cf. Eqs. (14) and (17)), the details are substantially different. Nott and Dunsmuir (2002) use hypothetical conditional processes, and assume an empirical covariance matrix computed from spatio-temporal data. Fuentes (2001) uses unconditional processes and applies them to purely spatial data. It involves Bayesian analysis without having to resort to computationally intensive Markov Chain Monte Carlo (MCMC) methods.

While the kernel smoothing method is convenient for accommodating non-stationarity, certain key elements of the approach (such as the number of locally stationary component models, or the size of the neighbourhoods for fitting the local models, and the nature of the weight or kernel) must be determined by somewhat *ad hoc* means.

Basis-function models

Another of the strategies for modelling non-stationarity in spatial statistics is based on decompositions of spatial processes in terms of empirical orthogonal functions (EOFs) (Sampson et al., 2001). The original methodology received attention in Nychka et al. (2002) for using a wavelet basis for decomposition of the empirical covariance matrix.

While most attention in the spatial statistics literature has focused on smoothing fields based on a single set of spatial observations, in many cases, replicates of the field are available, e.g., with environmental data collected over time. This sort of data is becoming more common with the growing availability of remotely sensed data. In this situation, one has multiple replicates for estimating the spatial covariance structure, albeit with certain restrictions, such as modelling only non-negative covariances.

Briefly, using the same spatio-temporal notation as above, the $n \times n$ empirical covariance matrix may be written with a spectral decomposition as

$$C = \hat{F} \Lambda F = \sum_{k=1}^{n^T} \lambda_k F_k \hat{F}_k \quad (18)$$

where $nT = \min(n, T)$, and \hat{F} and F are eigenfunctions. The extension of this finite decomposition to the continuous spatial case represents the covariance function as

$$C(\mathbf{x}, \mathbf{y}) = \hat{F} \Lambda F = \sum_{k=1}^{nT} \lambda_k F_k(\mathbf{x}) \hat{F}_k(\mathbf{y}) \quad (19)$$

where the eigenfunctions $F_k(\mathbf{x})$ and $\hat{F}_k(\mathbf{y})$ represent solutions to the Fredholm integral equation and correspond to the Karhunen-Loève (KL) decomposition (originally used in pattern recognition (Fukunaga, 1990)) of the mean-centred field as

$$Z(\mathbf{x}, t) = \sum_{k=1}^{\infty} A_k(t) F_k(\mathbf{x}) \quad (20)$$

The modelling and computational task here is in a numerical approximation of the Fredholm integral equation, or equivalently, choosing a set of generating functions $e_1(\mathbf{x}), \dots, e_p(\mathbf{x})$ that are the basis for an extension of the finite eigenvectors F_k to eigenfunctions $F_k(\mathbf{x})$.

In Nychka and Saltzman (1998) and Holland et al. (1999), the covariance function is represented as the sum of a conventional stationary model and a finite decomposition in terms of EOFs. This corresponds to a decomposition of the spatial process as a sum of stationary processes and a linear combination of M additional basis functions with random coefficients, the latter representing the deviation of the SRF structure from stationarity.

Nychka et al. (2002) propose a method for smoothing the empirical covariance structure of replicated data by thresholding the decomposition of the empirical covariance matrix in a wavelet basis. This approach has the advantages of allowing for very general types of covariance structure and of being very fast by virtue of the use of the discrete wavelet transform, with a computational focus on large problems with observations discretised to the nodes of a (large) $N \times M$ grid. They use a W-wavelet basis with parent forms that are piecewise quadratic splines, which are neither orthogonal nor compactly supported. These were chosen because they can approximate the shape of common covariance models such as the exponential and Gaussian, depending on the sequence of variances of the basis functions in the decomposition.

A first drawback to the basis-function approach is that it is not clear how much or what type of thresholding to apply, since there is no explicit model for the data. Given the difficulties involved in modelling high-dimensional covariance structures, it is also not clear how well the resulting smoothed covariance approximates the true covariance in a multivariate sense, although Nychka et al. (2002) have shown - in simulations - that individual elements of the smoothed covariance matrix can closely approximate the elements of stationary covariance matrices.

Kernel convolution methods

Higdon et al. (1999) define a non-stationary covariance function based on the convolution of kernels centred on the sites of interest. They propose a non-stationary spatial covariance function, defined by

$$C(\mathbf{s}_i, \mathbf{s}_j) = \int_{\mathbb{R}^2} K_{\mathbf{s}_i}(u) K_{\mathbf{s}_j}(u) du, \quad (21)$$

where $\mathbf{s}_i, \mathbf{s}_j$, and u are locations in \mathbb{R}^2 , and $K_{\mathbf{s}}$ is a kernel function (not necessarily non-negative) centred at location \mathbf{s} , with a shape depending on \mathbf{s} . This covariance function is positive-definite for spatially varying kernels of any functional form.

Higdon et al. (1999) justify this construction as the covariance function of a process, $Z(\cdot)$, constructed by convolving a white noise process, $\omega(\cdot)$, with a spatially varying kernel, $K_{\mathbf{s}}$:

$$Z(\mathbf{s}) = \int_{\mathbb{R}^2} K_{\mathbf{s}}(u) \omega(u) du \quad (22)$$

The evolution of the kernels in space produces a non-stationary covariance, and the kernels are usually parameterised so that they vary smoothly in space, under the assumption that nearby locations will share a similar local covariance structure.

Higdon et al. (1999) use Gaussian kernels, which give a closed form for $C(\mathbf{s}_i, \mathbf{s}_j)$, the convolution in Eq. (21). They use a bivariate Gaussian density function with a 2×2 covariance matrix Σ , which results in processes with a non-stationary Gaussian correlation function with the principal axes of Σ determining the directions of the anisotropic structure.

Higdon et al. (1999) demonstrate the particular case where the $K_s(\cdot)$ are bivariate Gaussian distributions characterised by the shapes of ellipses underlying the 2×2 covariance matrices Σ . The kernels are constrained to evolve smoothly in space by estimating the local ellipses under a Bayesian paradigm that specifies a prior distribution on the parameters of the ellipse (the relative location of the foci) as a Gaussian random field with a smooth (in fact, Gaussian) covariance function. The form of the kernel determines the shape of the local spatial correlation function, with a Gaussian kernel corresponding to a Gaussian covariance function.

Figure 3, from the presentation in Swall (1999), illustrates the nature of a fitted model for analysis of the spatial distribution of Dioxin concentrations in their Piazza Road study area in Missouri, USA. In this purely spatial example, Dioxin was transported through a small stream channel, which follows a curving path generally along the path of greatest concentration from top to bottom, as indicated in Figure 3.

The solid ellipses in Figure 3 indicate the shape of the Gaussian kernels at the sampling sites, as given by the posterior distribution of the Bayesian analysis; the major axis of the ellipse indicates the direction of greater spatial correlation, which roughly parallels the direction of the stream channel. The dotted ellipses represent the spatially varying estimates of these local kernels on a regular grid, in accordance with the Gaussian SRF prior for their parameters.

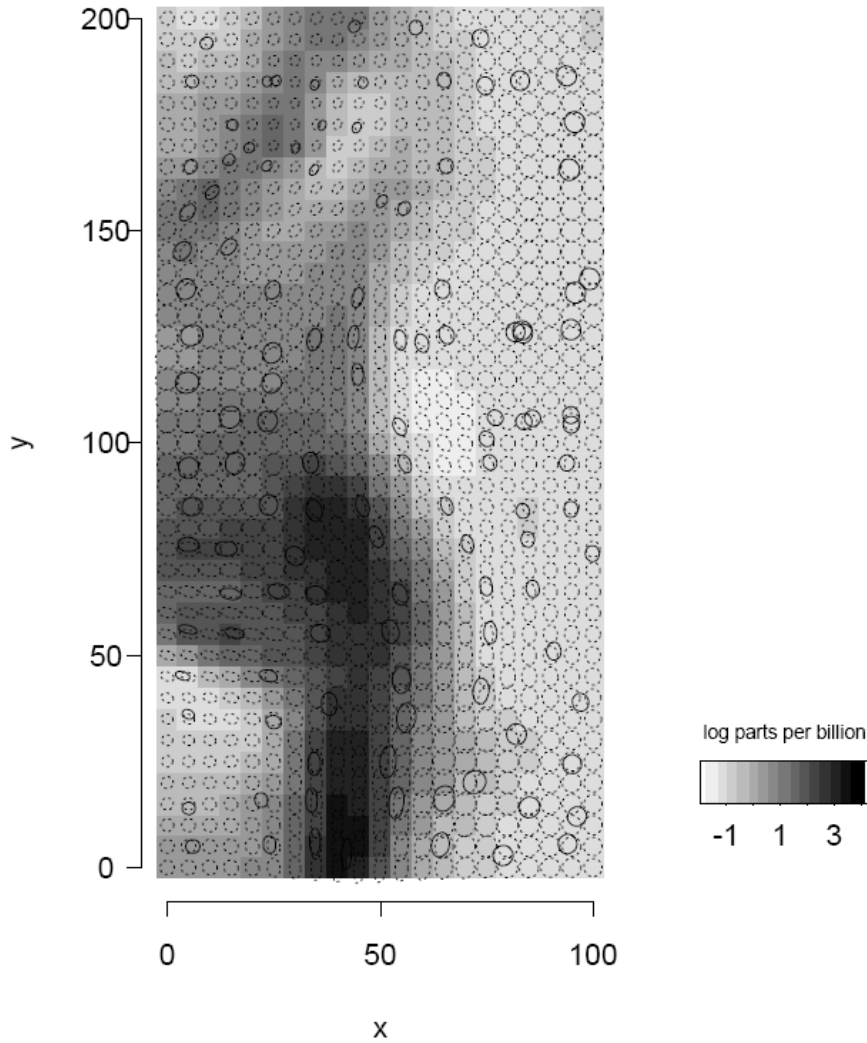


Figure 3: Estimated kernels of the process-convolution model for the Piazza Road data. Solid ellipses represent the kernels at the sampling sites and dotted ellipses the extension to a regular grid according to the random field prior model. The underlying image shows the corresponding posterior mean estimates for the Dioxin concentrations (from Swall (1999)).

Two key advantages of the non-stationary covariance model based on Higdon et al. (1999) is that it fully defines the covariance at unobserved as well as observed locations, and does not require a regular grid of observations. This stands in contrast to the approach of Nychka et al. (2002), although they briefly suggested an iterative approach to deal with irregularly spaced observations. This makes the Higdon et al. (1999) approach particularly

applicable to geodetic LSC, which is used for estimation and prediction at observed and unobserved locations.

NUMERICAL SIMULATION OF NON-STATIONARY VERSUS STATIONARY COVARIANCES

In this simulated case-study example, a small regular grid is provided to show how stationary and non-stationary covariances are computed in different ways and how they affect the predictions by LSC differently.

Suppose that we want to estimate a quantity at point **p** on a regular 20 m × 20 m grid using the first two rings of neighbouring data (a ring means the nearest neighbouring points in a square grid). To build the vector of observations, a set of 25 random numbers from a normal distribution with $\mu=0$ and $\sigma=1$ were selected. Figure 4 shows the data set.

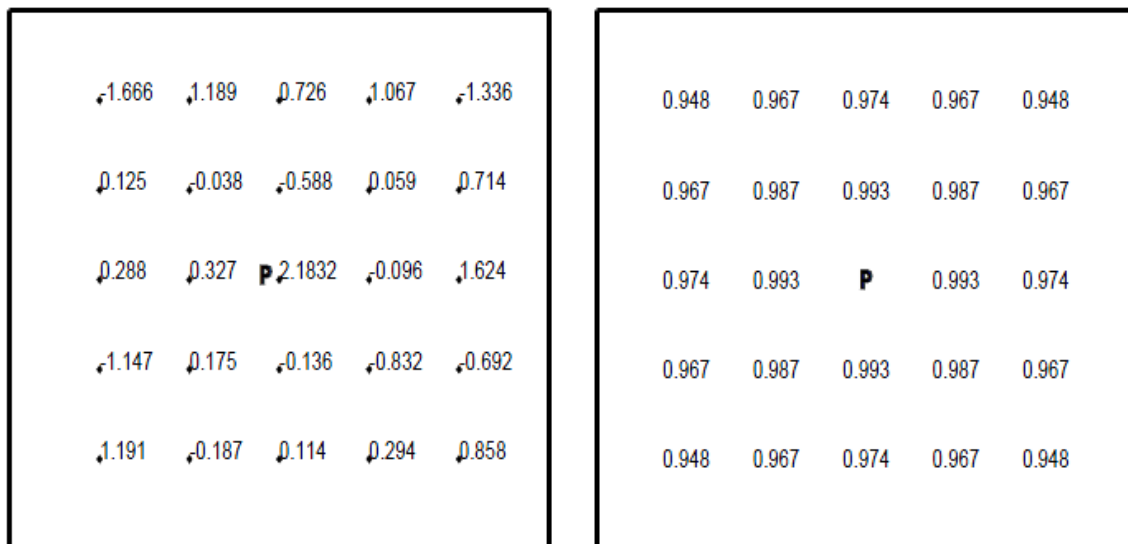


Figure 4: (left) Stationary configuration of a sample data set. (right) Gaussian stationary covariances between point **p** and stationary data with $d=245$ m

For computing covariances between point \mathbf{p} and each neighbouring point, we only need to know the two parameters of variance C_0 and correlation length d . In LSC practice, these two parameters are estimated by fitting empirical covariances to covariance models, but because of the randomness and limited extent of data here, this usual procedure is not possible. Instead, $C_0=\sigma^2=1$ is fixed and the value of d is changed, until the minimum absolute error of

$$e_{absolute} = P_{real} - P_{prediction} \quad (23)$$

for point \mathbf{p} is reached at $d=245$ m, or in other words $\lim_{d \rightarrow 245} (e_{absolute}) = minimum$.

The correlation length is rather large in this case, because a random set of data was used. Figure 5 shows stationary covariances for point \mathbf{p} based on a Gaussian stationary covariance model.

$$C_{Gaus}(r) = C_0 \exp\left(-\frac{r^2}{d^2}\right) \quad (24)$$

Note that stationary covariances change with the distance of the points from \mathbf{p} (a result of the symmetric characteristic of a stationary covariance function); points with equal distances from \mathbf{p} have the same covariances, and further points from \mathbf{p} have smaller covariances.

To simulate a non-stationary data configuration, two sets of random numbers from a normal distribution, 16 points with $\mu=0$ and $\sigma=1$ (dots in Figure 5) and nine points (circles in Figure 5) with $\mu=0$ and $\sigma=2$ were selected. Although the new grid is non-stationary (Figure 5), standard stationary covariance modelling was first applied to estimate covariances between point \mathbf{p} and the rest of the grid. For computing stationary covariances, the same logic of the previous stationary grid applies. The Gaussian stationary covariances for point \mathbf{p} with variance of $C_0=\sigma^2=1$ and correlation length of $d = 210$ m is seen in Figure 5.

For computing non-stationary covariances, according to Higdon et al. (1999), the parameters of ellipses (α , a and b) should first be defined at every point. There is no sign of anisotropy in this data set, so $\alpha=0$ was set for all ellipses. Considering two groups of

statistical parameters for this grid, two parameter groups should be defined, each attributed to points belonging to the two separate regions.

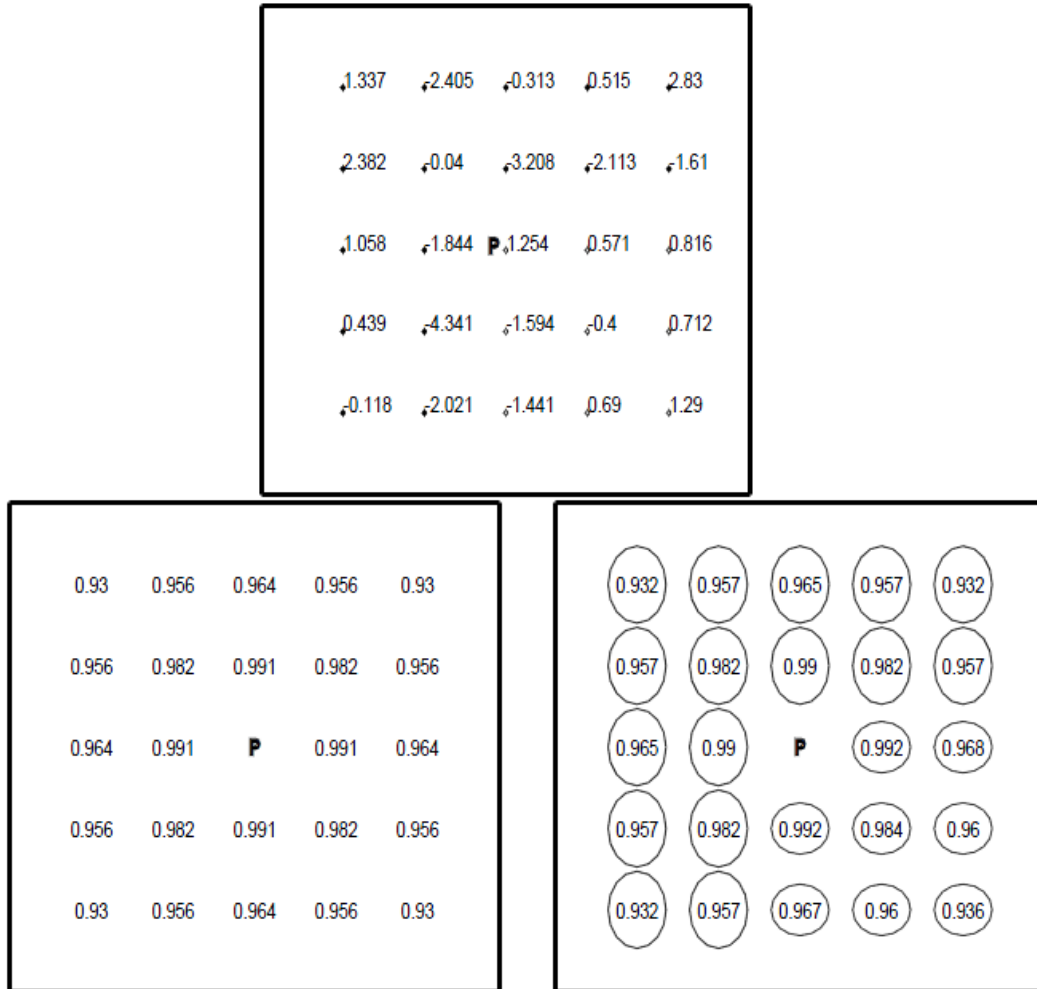


Figure 5: (top) Non-stationary configuration of a sample data set. (bottom left) Stationary covariances between point **p** and non-stationary data. (bottom right) Non-stationary covariances between point **p** and non-stationary data, elliptical kernels attributed to each location are used to construct the non-stationary covariances

Here, like the stationary case, it is not possible to estimate non-stationary parameters from empirical covariances, so again the minimum absolute error for point **p** is used:

$$\lim_{\substack{a \rightarrow 210.01m \\ b \rightarrow 210.01m \\ a_o \rightarrow 220.39m \\ b_o \rightarrow 220.01m}} (e_{absolute}) = \text{minimum} \quad (25)$$

In Figure 5, ellipses for each point are based on the parameters in Eq. (24). Accordingly, Gaussian non-stationary covariances of point \mathbf{p} with the rest of the grid are printed in the centre of each ellipse.

$$C_{Gaus}^{NS}(\mathbf{s}_i, \mathbf{s}_j) = C_0 |\Sigma_i|^{\frac{1}{4}} |\Sigma_j|^{\frac{1}{4}} \left| \frac{\Sigma_i + \Sigma_j}{2} \right|^{-\frac{1}{2}} \exp(-Q_{ij}) \quad (26)$$

where

$$Q_{ij} = (\mathbf{s}_i - \mathbf{s}_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2} \right)^{-1} (\mathbf{s}_i - \mathbf{s}_j) \quad (27)$$

$$\frac{1}{\Sigma^2} = \tau \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \quad (28)$$

For the stationary and non-stationary spatial data configurations, stationary covariances are decreased by a ratio of distance, which is characterised by the correlation length (Figures 4 and 5). For non-stationary covariance models, however, the covariances also depend on the location. Figure 5 shows that non-stationary covariances between point \mathbf{p} and points in the bottom-right corner are higher than rest of the grid, reflecting that the non-stationary covariance function accounts for non-stationary data. Table 1 shows the non-stationary covariance model has improved the prediction of point \mathbf{p} in the non-stationary configuration; $e_{absolute}$ decreased from 1.4153 to 0.0331.

Table 1: Results of the predictions for point **p** based
on stationary and non-stationary covariance models

Data property	Covariance model	Parameter (m)	$e_{absolute}$
Stationary	Stationary	$d=245$	0.1549
Non-stationary	Stationary	$d=210$	1.4153
Non-stationary	Non-stationary	$a.=210.01$ $b.=210.01$ $a_{\circ}=220.39$ $b_{\circ}=220.01$	0.0331

Standard geodetic LSC, regardless of the stationarity or non-stationarity of the data, uses stationary covariances for the prediction. This simulation (Table 1) has showed that using stationary covariances in LSC to predict values from non-stationary spatial data causes underestimation or overestimation in the predictions. On the other hand, using non-stationary covariances based on the Higdon et al. (1999) method in LSC improves the prediction when the data are non-stationary.

DISCUSSION AND SUMMARY

The primary goal of this review paper was to document the appropriate selection of non-stationary methods from the disciplines of geodesy, geostatistics and spatial statistics that can be applied to geodetic LSC. This is summarised in Table 2. This forms background for the non-stationary LSC application described in Darbeheshti and Featherstone (2009, 2010).

Of the methods reviewed from the very large body of literature, spatial deformation needs replicated data, which is not always the case in geodesy, and is better suited for spatial time-series analysis. Kernel smoothing models use Bayesian analysis and numerical approaches like Markov Chain Monte Carlo methods, which are very time consuming. Basis function models need gridded data, which are rarely available in geodesy, or require the prior

interpolation of irregularly spaced data. In such a case, non-stationary models would be needed in LSC to grid such datasets (cf. Darbeheshti and Featherstone, 2009).

On the other hand, the kernel convolution method of Higdon et al. (1999) has been tested in Kriging for topographic data by Paciorek and Schervish (2006), which is very similar to the LSC application to geodetic data. Therefore, this method does not have the disadvantages of other approaches from the geodetic perspective and the adaption of this method according to the similarity of LSC to Kriging is more likely (Dermanis, 1984). This is one reason why it was selected by Darbeheshti and Featherstone (2009, 2010).

Table 2: Summary and description of non- stationary methods

Discipline	Non-stationary method	Description
Geodesy	Trend removal	Adds extra steps to the computation and enforces biases to the input data
	Riesz representers	Based on iterative approach to find the optimum position of point masses to match the gravity field
	Wavelet approach	Limited application to the filtering of non-stationary errors from stationary signal
Geostatistics	Locally adaptive Kriging	A solution for non-stationary mean
	Segmentation	Results in edge effects from discontinuities in the statistical parameters at the borders of neighbouring areas
Spatial statistics	Spatial deformation	Requires independent replications of the process
	Kernel smoothing	No closed-form expression for covariance models and needs gridded data
	Basis-function models	Not clear how much or what type of thresholding to choose
	Kernel convolution	Chosen by Darbeheshti and Featherstone (2009, 2010) to model non-stationary covariance modelling in geodesy

Through this review, spatial statistics - specifically recent research in environmental applications - offers a variety of non-stationary covariance models that appear to be new to the discipline of geodesy. This has indicated that the kernel convolution of Higdon et al. (1999), among other methods in spatial statistics, has advantages so that it can be adapted for LSC in geodesy. As such, we have applied them in the following two case-study examples.

Firstly, non-stationary 2D LSC was used to interpolate gravity anomalies over the Darling Fault in Western Australia by Darbeheshti and Featherstone (2009), where this particular gravity field functional is highly anisotropic and non-stationary. From internal and external error estimates, the non-stationary covariance models in LSC were consistently better than stationary LSC for interpolation; they also gave more realistic error estimates in areas where the SRF varies rapidly.

Secondly, non-stationary covariance functions were used by Darbeheshti and Featherstone (2010) to create an iterative optimisation loop to tune a gravimetric quasigeoid model to a geometric quasigeoid at GPS-levelling points for the Perth region of Western Australia to within a user-prescribed level of tolerance.

LIST OF MAJOR SYMBOLS

a major axis of an ellipsoid	λ geodetic longitude
b minor axis of an ellipsoid	m arithmetic mean
α geodetic azimuth	μ expected value
C covariance	ϕ geodetic latitude
C_0 variance	Q Mahalanobis distance
d correlation length	R correlation
K convolution kernel	r Euclidean distance

ψ spherical distance

T anomalous potential of the Earth

s spatial position

Σ spatial kernel

ACKNOWLEDGEMENTS

ND thanks Curtin University of Technology for a CIRTS scholarship and the Department of Spatial Sciences for a living allowance scholarship. WEF thanks the Australian Research Council for Discovery Project grant DP0663020 and J Hicks for alerting him to the origins of Thomas Bayes.

REFERENCES

- Armstrong M (1998) Basic linear geostatistics, Springer, Berlin Heidelberg New York
- Atkinson, P.M. and Lloyd, C.D. (2007) Non-stationary variogram models for geostatistical sampling optimisation: an empirical investigation using elevation data, *Computers & Geosciences*, vol. 33, pp. 1285-1300, doi:10.1016/j.cageo.2007.05.011
- Barthelmes, F., Dietrich, R. and Lehmann, R. (1991) Use of point masses on optimised positions for the approximation of the gravity field, In Rapp, R.H. (ed) *Determination of the Geoid*, Springer, Berlin Heidelberg New York, pp. 484-493
- Bjerhammar, A. (1964) A new theory of geodetic gravity, Technical Report 243, Royal Institute of Technology, Stockholm.
- Boschetti, F., Dentith, M.C. and List, R. (1997) Inversion of potential field data by genetic algorithms, *Geophysical Prospecting*, vol. 45, no. 3, pp. 461-478.
- Brovelli, M.A., Sansò, F. and Venuti, G. (2003) A discussion on the Wiener-Kolmogorov prediction principle with easy-to-compute and robust variants, *Journal of Geodesy*, vol. 76, no. 11, pp. 673-683, doi: 10.1007/s00190.002.0292.3.
- Chilès, J.P. and Delfiner, P. (1999) *Geostatistics*, Wiley, New York
- Collier, P.A., Leahy, F.J. and Argeşeanu, V.S. (1997) Options and strategies for transition to GDA94, *The Australian Surveyor*, vol. 42, no. 3, pp. 116-125.

- Collier, P.A., Leahy, F.J. and Argeseanu, V.S. (1998) Distortion modelling and the transition to GDA94, *The Australian Surveyor*, vol. 43, no. 1, pp. 29-40.
- Cressie, N. (1993) *Statistics for spatial data*, Wiley, New York
- Damian, D., Sampson, P. and Guttorp, P. (2001) Bayesian estimation of semi-parametric non-stationary spatial covariance structures, *Environmetrics*, vol. 12, no. 2, pp. 161-178, doi: 10.1002/1099.095X (200103)12:2<161::AID.ENV452>3.0.CO;2.G.
- Darbeheshti, N. and Featherstone, W.E. (2009) Non-stationary covariance function modelling for least-squares collocation, *Journal of Geodesy*, vol. 83, no. 6, pp. 495-508, doi:10.1007/s00190-008-0267-0.
- Darbeheshti N. and Featherstone W.E. (2010) Tuning a gravimetric quasigeoid to GPS-levelling by non-stationary least-squares collocation, *Journal of Geodesy* (accepted on 09/03/10)
- Daubechies, I. (1992) *Ten lectures on wavelets*, SIAM Press, Philadelphia.
- Dermanis, A. (1984) Kriging and collocation - a comparison, *manuscripta geodaetica*, vol. 9, no. 3, pp. 159-167.
- Fuentes, M. (2001) A high frequency Kriging approach for non-stationary environmental processes, *Environmetrics*, vol. 12, no. 5, pp. 469-483, doi: 10.1002/env.473
- Fuentes, M. and Smith, R. (2001) A new class of non-stationary spatial models, Technical Report, Department of Statistics, North Carolina State University, Raleigh.
- Fukunaga, K. (1990) *Introduction to statistical pattern recognition*, Elsevier, New York.
- Deutsch, C.V. and Journel, A.G. (1998) *GSLIB*, Oxford University Press, Oxford.
- Goad, C.C., Tscherning, C.C. and Chin, M.M. (1984) Gravity empirical covariance values for the continental United States, *Journal of Geophysical Research*, vol. 89, no. B9, pp. 7962-7968.
- Haar, A. (1910) Zur Theorie der orthogonalen Funktionensysteme, *Mathematische Annalen*, vol. 69, no. 3, pp. 331-371.
- Heiskanen, W. and Moritz, H. (1967) *Physical Geodesy*, Freeman, San Francisco
- Higdon, D., Swall J. and Kern, J. (1999) Non-stationary spatial modelling, In: Bernardo J.M., Berger J.O., Dawid A.P. and Smith A.F.M. (eds) *Bayesian Statistics 6*, Oxford University Press, Oxford, pp.761-768.

- Holland, D., Saltzman, N., Cox, L. and Nychka, D. (1999) Spatial prediction of sulfur dioxide in the eastern United States. In *Geostatistics for Environmental Applications*, Volume 10, Springer, Berlin, pp. 65-76.
- Keller, W. (1998) Collocation in reproducing kernel Hilbert spaces of a multiscale analysis, *Physics Chemistry and Earth*, vol. 23, no. 1, pp. 25-29.
- Keller, W. (2000) A wavelet approach to non-stationary collocation. In: Schwarz K.P. (ed), *Geodesy Beyond 2000*, Springer, Berlin Heidelberg New York, pp. 208-214.
- Keller, W. (2002) A wavelet solution to 1D non-stationary collocation with extension to the 2D case. In: Sideris M.G. (ed), *Gravity, Geoid and Geodynamics 2000*, Springer, Berlin Heidelberg New York, pp. 79-84.
- Knudsen, P. (2005) Patching local empirical covariance functions – a problem in altimeter data processing. In: Sansò F (ed), *A Window on the Future of Geodesy*, Springer, Berlin Heidelberg New York, pp. 483-487.
- Kotsakis, C. and Sideris, M.G. (1999) The long road from deterministic collocation to multiresolution approximation, Report 1999.5, Department of Geodesy and Geoinformatics, Stuttgart University, Stuttgart
- Kotsakis, C. (2000) The multiresolution character of collocation, *Journal of Geodesy*, vol. 74, nos. 3-4, pp. 275-290, doi: 10.1007/s001900050286
- Lambeck, K. (1987) The Perth Basin: A possible framework for its formation and evolution, *Exploration Geophysics*, vol. 18, nos. 1&2, pp. 124-128.
- Lloyd, C.D. and Atkinson, P.M. (2002) Non-stationary approaches for mapping terrain and assessing prediction uncertainty, *Transactions in GIS*, vol. 6, no. 1, pp. 17-30.
- Mallat, S.G. (1989) A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693.
- Moritz, H. (1980) *Advanced Physical Geodesy*, Abacus, Tunbridge Wells.
- Nott, D. and Dunsmuir, W. (2002) Estimation of non-stationary spatial covariance structure, *Biometrika*, vol. 89, no. 4, pp. 819-829, doi: 10.1093/biomet/89.4.819.

- Nychka, D. and Saltzman, N. (1998) Design of air quality monitoring networks. In: Nychka D., Piegorsch W. and Cox L. (eds) Case Studies in Environmental Statistics, Lecture Notes in Statistics 132, Springer, New York, pp. 51-75.
- Nychka, D., Wikle, C. and Royle J. (2002) Multiresolution models for non-stationary spatial covariance functions, *Statistical Modelling* vol. 2, no. 4, pp. 315-331, doi: 10.1191/1471082x02st037oa.
- Perrin, O. and Meiring, W. (1999) Identifiability for non-stationary spatial structure, *Journal of Applied Probability*, vol. 36, no. 4, pp. 1244-1250, doi: 10.1239/jap/1032374771.
- Ruffhead, A. (1987) An introduction to least-squares collocation, *Survey Review*, vol. 29, no. 224, pp. 85-94.
- Paciorek, C.J. and Schervish, M.J. (2006) Spatial modelling using a new class of non-stationary covariance functions, *Environmetrics* vol. 17, no. 5, pp. 483-506, doi: 10.1002/env.785
- Sampson, P.D. and Guttorp, P. (1992) Nonparametric estimation of non-stationary spatial covariance structure, *Journal of the American Statistical Association*, vol. 87, no. 417, pp. 108-119.
- Sampson, P., Damian, D. and Guttorp, P. (2001) Advances in modeling and inference for environmental processes with non-stationary spatial covariance. In *Geostatistics for Environmental Applications*, Volume 11, Springer, Berlin, pp. 17-32.
- Smith, R.L. (1996) Estimating nonstationary spatial correlations, Technical report, Department of Statistics, Cambridge University, Cambridge
- Stopar, B., Kuhar, A.T.M. and Turk, G. (2006) GPS-derived geoid using artificial neural network and least squares collocation, *Survey Review*, vol. 38, no. 300, pp. 513-524.
- Swall, J.L. (1999) Non-stationary spatial modelling using a process convolution approach, PhD thesis, Institute of Statistics and Decision Sciences, Duke University, Durham
- Tscherning, C.C. and Rapp, R.H. (1974) Closed covariance expressions for gravity anomalies, geoid undulations, and deflections of the vertical implied by anomaly degree variance models, Report 208, Department of Geodetic Science, The Ohio State University, Columbus.

- Tscherning, C.C. (1984) Local approximation of gravity potential by least squares collocation. In: Schwarz, K-P. (ed) Local Gravity Field Approximation, Division of Surveying Engineering, University of Calgary, Calgary, pp. 277-362.
- Tscherning, C.C., Sansò, F. and Arabelos, D. (1987) Merging regional geoids – preliminary considerations and experiences, *Bollettino di Geodesia e Scienze Affini*, vol. XLVI, no. 3, pp. 191-206.
- Tscherning, C.C. (1994) Geoid determination by least-squares collocation using GRAVSOFT, Lectures Notes for the International School for the Determination and Use of the Geoid, DIIAR - Politecnico di Milano, Milano
- Tscherning, C.C. (1999) Construction of anisotropic covariance functions using Riesz-representers, *Journal of Geodesy*, vol. 73, no. 6, pp. 332-336, doi: 10.1007/s001900050250
- Tscherning, C.C. (2006) Comments on Stopar et al., *Survey Review*, vol. 38, no. 302, pp. 716.
- Vermeer, M. (1995) Mass point geopotential modelling using fast spectral techniques; historical overview, toolbox description, and numerical experiment, *manuscripta geodaetica*, vol. 20, no. 5, pp. 362-378.
- You, R.J. (2006) Local geoid improvement using GPS and levelling data: case study, *Journal of Surveying Engineering*, vol. 132, no. 3, pp. 101-107.
- You, R.J. and Hwang, H.W. (2006) Coordinate transformation between two geodetic datums of Taiwan by least-squares collocation, *Journal of Surveying Engineering*, vol. 132, no. 2, pp. 64-70.
- Wackernagel, H. (2003) *Multivariate Geostatistics: An Introduction with Applications*, Springer, Berlin.