Journal of Geodesy manuscript No.
(will be inserted by the editor)

Non-stationary covariance function modelling in 2D least-squares collocation

N. Darbeheshti · W.E. Featherstone

Received: date / Accepted: date

Abstract Standard least-squares collocation (LSC) assumes 2D stationarity and 3D isotropy, and relies on a covariance function to account for spatial dependence in the observed data. However, the assumption that the spatial dependence is constant throughout the region of interest may sometimes be violated. Assuming a stationary covariance structure can result in over-smoothing of, e.g., the gravity field in mountains and under-smoothing in great plains. We introduce the kernel convolution method from spatial statistics for non-stationary covariance structures, and demonstrate its advantage for dealing with non-stationarity in geodetic data. We then compared stationary and non-stationary covariance functions in 2D LSC to the empirical example of gravity anomaly interpolation near the Darling Fault, Western Australia, where the field is anisotropic and non-stationary. The results with non-stationary covariance functions are better than standard LSC in terms of formal errors and cross-validation against data not used in the interpolation, demonstrating that the use of non-stationary covariance functions can improve upon standard (stationary) LSC.

N. Darbeheshti
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; E-mail: neda.darbehesti@postgrad.curtin.edu.au

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; E-mail: W.Featherstone@curtin.edu.au

Keywords Least squares collocation (LSC) · non-stationary covariance function modelling · elliptical kernel convolution · gravity field interpolation · Darling Fault, Australia

1 Introduction and Motivation

Modelling the spatial dependence structure of observed data is fundamental to theory of spatial estimation, which is commonly called least-squares collocation (LSC) in geodesy (e.g., Krarup 1969; Moritz 1980) and Kriging in geostatistics (e.g., Krige 1951; Matheron 1962). However, they are not exactly the same, so for details on the differences and similarities between LSC and Kriging, please refer to, e.g., Dermanis (1984) and Schaffrin and Felus (2005).

Ever since the conception of LSC in geodesy, 2D stationarity and 3D isotropy have been routinely assumed (e.g., Tschernning and Rapp 1974). Stationarity in this context means that the mean is constant and the covariance is position invariant, which is called weak or second-order stationarity in the discipline of spatial statistics (e.g., Armstrong 1998). Isotropy means that the spatial dependence is independent of direction.

Remark 1: The terminology of non-homogenous covariance modelling, used by Tscherning (1999), is the same as the non-stationarity terminology used in this paper. We choose this because it is used much more widely in spatial statistics (e.g., Cressie 1999) and physical geodesy (e.g., Keller 1998, 2000, 2002, Kotsakis 2000).

Non-stationarity models of the spatial mean have been applied in spatial statistics for many years (e.g., Deutsch and Journel 1998). Ordinary Kriging (OK) often involves an unknown, but constant, mean (Chilès and Delfiner 1999). OK, as applied within moving data neighbourhoods, is already a non-stationary algorithm in that it corresponds to a non-stationary model with varying mean but stationary covariance (Deutsch and Journel 1998). Reguzzoni et al. (2005) extended OK to general Kriging (GK) in order to deal with the case where the data are not point values but rather values of linear functionals of the unknown random field. More commonly, where the mean is known to vary with location, it can be modelled with a trend, resulting in Kriging with a trend (KT), also known as universal Kriging (UK) (Cressie 1993).

This paper investigates non-stationarity of covariances in standard 2D LSC. Stationary covariances previously led to almost exclusive reliance on models of the form

$Cov(s_i, s_j) = C(r)$, which expresses the covariance of between locations s_i and s_j in $\mathbb{R}^2$ (2D) as a function of Euclidean distance r. However, empirical covariances of gravity anomalies (e.g., Goad et al. 1984; Schwarz and Lachapelle 1980; Knudsen 2005) show large differences between the parameters of the covariance function for the whole region (based on a global covariance function) and sub-regions (based on local covariance functions), which shows the inadequacies of using only stationary covariance functions.

On the other hand, there is a growing body of literature in spatial statistics on methods for dealing with non-stationarity. The more common approaches are:

- *Spatial deformation of the data*: Sampson et al. (2001) review these deformation approaches. They act to distort the surface such that a stationary covariance function results from the data in the transformed space. They need repeated measurements that are suitable for temporally independent replications of space-time monitoring.
- *Trend removal*: The mean or median of the data and a linear polynomial are subtracted from the data (Cressie 1993).

  Remark 2: Deformation and trend removal techniques allow for a transformation to a stationary covariance structure; they are not actually methods to model the non-stationarity (Sampson and Guttorp 1992).
- *Locally adaptive Kriging*: This involves predicting and modelling the local covariances and using the coefficients of the locally fitted model in local Kriging (Goovaerts 1997).
- *Segmentation*: Segment-based covariances (Atkinson and Lloyd 2007) divide the region of interest into smaller sub-regions within which the covariance function can be considered stationary.
- *Kernel convolution*: This provides an attractive way of practically introducing non-stationarity, while retaining clear interpretation and permitting analytic calculation. There are two distinct approaches in the spatial statistics literature, one due to Higdon et al. (1999) and the other to Fuentes (2001) and Fuentes and Smith (2003). Kernel convolution provides a continuous way of non-stationary covariance function modelling. Unlike locally adaptive Kriging or segmentation, it avoids the need for patching at the borders. This is the method that we shall use in this paper.

Nevertheless, the issue of non-stationarity in geodetic data has already been recognised, as follows:

- *Trend removal*: One example is the remove-compute-restore technique in LSC-based gravity field determination, where gravitational effects of a global geopotential model and the topography are removed from the data and subsequently restored. Another example is the prior removal of a bias and tilt before LSC is used, e.g., to fit gravimetric quasi/geoid models to GPS-levelling.
- *Segmentation*: This was studied in Tscherning et al. (1987) for merging regional geoids. However, segmentation methods need to subsequently patch the covariances or results at the boundaries of the sub-regions (cf. Knudsen 2005).
- *Wavelets*: A multiresolution/wavelet framework for LSC has been suggested (Kotsakis and Sideris 1999; Kotsakis 2000; Keller 1998, 2000, 2002) because of the space-localisation properties of wavelets, which makes them useful for spectral studies of spatially varying (non-stationary) signals. Although the wavelet framework for LSC has had a purely theoretical character in geodesy, it has been practically applied in spatial statistics over independent replications of data in time (e.g., Nychka et al. 2002; Paciorek 2003).
- *Other examples*: Tscherning (1991) proposed an approach based on mass-density anomaly considerations, but it was not pursued due to the limited capability of computers in the early 1990s. Tscherning (1999) suggested another possibility to use Riesz representers, but which was not practically applicable because of the lack of a sufficient numerical algorithm.

  Moreaux et al. (1999) proposed three methods to construct positive-definite functions with compact support for the approximation of harmonic covariance functions on the sphere. While one of these non-harmonic locally supported functions was not isotropic, similar results were obtained for different configurations of points. Nevertheless, Moreaux et al. (1999) concluded that the most relevant method was not a non-isotropic function.

The apparent lack of interest in non-stationary covariance modelling for LSC, in comparison with other Earth-science-related disciplines, is possibly due to the insensitivity of the LSC solution to the covariance parameters (e.g., Sansò et al. 2000 and Xu 1991). Sansò et al. (2000) showed that in the limit of increasing data density, the dependency of the LSC solution on the type of covariance function decreased.

However, in support of non-stationary covariance modelling for LSC:

– the dependency of the error covariance matrices in the covariance function is evident from the LSC formulation and literature (e.g., Tscherning 1975);
– stationary covariance functions that depend on only variance and correlation length are not a good device with which to control the LSC results.

In this paper, we first demonstrate non-stationarity in gravity anomaly covariances across Australia. We then borrow and apply the idea of kernel convolution from Higdon et al. (1999), extended upon by Paciorek (2003), to create a class of non-stationary covariance functions for use in standard 2D LSC. These convolution kernels are specified at each (horizontal) location by an ellipse that is oriented and scaled depending on the non-stationarity of the data.

As an example of the application of non-stationary 2D LSC, we consider the problem of gravity anomaly interpolation around the near-north-south-trending Darling Fault in Western Australia (Lambeck 1987), where the anomalous field is non-stationary from the west versus the east. A comparison is made between LSC prediction errors using stationary and non-stationary covariance functions, showing that the use of non-stationary covariance functions can improve upon standard LSC for gravity anomaly gridding.

2 Invalidation of stationary covariances

2.1 Previous work by others

The assumption of a stationary gravity field over the Earth has been queried for as long as covariance techniques have been discussed and used (e.g., Kearsley 1977). Rapp (1964) investigated three different regions of the USA, and found significantly different covariance functions for each. Gaposhkin (1973) analyzed subsets of data of oceanic and continental gravity assuming isotropy, and found the differences between covariances to be significant. They concluded that "$\cdots$ gravity is not stationary". Similarly, in a statistical analysis of gravity over the then Czechoslovakia, Vyskočil (1970) investigated the covariance functions for gravity anomaly Δg means of $10' \times 15'$ blocks and concluded "$\cdots$ the theoretical assumption of homogeneity and isotropy$\cdots$ is basically only forced and in reality it will probably not always be satisfied."

Schwarz and Lachapelle (1980) showed that some common characteristics exist for a local covariance function over an area as large as Canada, excluding the Rocky Mountains, which shows a compromise. In similar study by Goad et al. (1984), variances and correlation distances were determined for complete/refined Bouguer anomalies in 3030 arc-minute blocks covering the continental USA. Their analysis showed that the correlation length is smallest in areas of high topography and large variances are associated with the mid-continental gravity high and tectonic features on the USA Pacific coast. Flury (2006) did the same analysis on covariances of topography-reduced gravity anomalies, but in the frequency domain over 13 test regions in Europe, showing that the covariance parameters are different in mountainous versus flat terrain.

Analogously, Atkinson and Lloyd (2007) studied local and global variograms in Kriging for height data, showing them to vary markedly across their region of study.

2.2 Empirical proof across Australia

Gravity anomalies over the Australian continent offer a (convenient to us) sample to test for non-stationarity in geodetic data. We analyzed free-air gravity anomalies to determine to what extent the empirical covariances vary spatially. An irregularly spaced data set of 1,106,662 point observations on land with average separation of mostly 11 km, but with some denser areas down to 30 m, were provided by Geoscience Australia (Fig. 1). They were used to estimate local gravity anomaly covariance functions following the approach in Schwarz and Lachapelle (1980).

Fig. 1 near here

Empirical covariance functions were computed in 39 $5^\circ \times 5^\circ$ sub-regions covering the whole onshore gravity data set in Australia (Fig. 1). The terrestrial free-air gravity anomalies are residual to the degree/order 360/360 EGM96 global geopotential model (Lemoine et al. 1998). The parameters of the local covariance functions (variance C_0 and correlation length d) for each sub-region were computed using Tscherning's (1994) EMPCOV routine, and are summarised in Table 1.

Table 1 near here

In Table 1, C_0 is the prior variance and $\hat{C}_0$ is the variance estimated from EMPCOV. It computes an empirical covariance function of scalar quantities on a sphere by

taking the mean of product-sums of samples of scalar values (here Δg):

$$cov(\psi_j) = \frac{1}{N_j} \sum_{k,i}^{N_j} \Delta g(\varphi_k, \lambda_k) \Delta g(\varphi_i, \lambda_i) \tag{1}$$

where ψj is the spherical distance and N_j is the number of pairs for each interval.

The difference between C_0 and $\hat{C}_0$ in Table 1 is due the zero-mean assumption implicit in LSC (Moritz 1980) and implemented in EMPCOV, but which is not satisfied for most of the sub-regions over Australia (Table 1, third column). Ideally, it should be the same if the mean value (m) of each subregion was included in Eq. (1):

$$cov(\psi_j) = \frac{1}{N_j} \sum_{k,i}^{N_j} (\Delta g(\varphi_k, \lambda_k) - m)(\Delta g(\varphi_i, \lambda_i) - m) \tag{2}$$

From Table 1, there is a significant difference in the $\hat{C}_0$ and $\hat{d}$ values between areas with rugged topography, like Tasmania (area 40), and those from flat regions in central Australia (areas 16, 17 and 22). The areas with rugged topography show smaller correlation lengths and larger variances, which agree with the findings in other parts of the world (cf. Gaposchkin 1973; Schwarz and Lachapelle 1980; Goad et al. 1984; Flury 2006).

The different covariance parameters for different sub-regions shows that one single stationary covariance function is an inadequate representation of localised non-stationarities in Australia. Consequently, using a stationary LSC covariance model would underestimate values at some points and overestimate others. Here, we have essentially used the segmentation method (Atkinson and Lloyd 2007), but this is still not satisfactory because the boundaries would still have to be patched (cf. Knudsen 2005, Tscherning et al. 1987). Therefore non-stationary covariance modelling becomes an attractive alternative.

3 Non-stationary covariance functions using kernel convolutions

First, the formulation here is in Euclidean space when the vector locations of s_i have the Euclidean coordinates of (x_i, y_i, z_i). Although much of the LSC literature deals with spherical covariance functions, this Euclidean formulation could be applied elsewhere in physical geodesy if $3D$ coordinates are available or if a planar approximation is

sufficient in $2D$ cases. In this paper, we consider the simplest 2D case of interpolation under a planar approximation as a first start.

We describe the approach of Higdon et al. (1999) (henceforth HSK) for defining non-stationary covariance functions in a geodetic context, which will later be implemented in standard 2D LSC. HSK propose a non-stationary spatial covariance function, defined by

$$C(\mathbf{s}_i, \mathbf{s}_j) = \int_{R^2} K_{\mathbf{s}_i}(u) K_{\mathbf{s}_j}(\mathbf{u}) d\mathbf{u}, \tag{3}$$

where $\mathbf{s}_i$, $\mathbf{s}_j$, and $\mathbf{s}$ are locations in $\mathbb{R}^2$, and $K_{\mathbf{s}}$ is a kernel function centered at $\mathbf{s}$. This construction is the covariance function of a white noise process $\omega(.)$ (Appendix B), convolved with the kernel function to produce the process, $Z(.)$, defined by

$$Z(\mathbf{s}) = \int_{R^2} K(\mathbf{s}-\mathbf{u})\omega(\mathbf{u}) d\mathbf{u} = \int_{R^2} K_{\mathbf{s}}(\mathbf{u})\omega(\mathbf{u}) d\mathbf{u}. \tag{4}$$

One can avoid technical difficulties involved in defining such a white noise process by using the definition of positive-definiteness (e.g., Cressie 1993; Moritz 1980) to show that the covariance function is positive-definite in every Euclidean space, $\mathbb{R}^p$; $p = 1, 2, \cdots, n$; this gives:

$$\begin{aligned} \textstyle\sum_{i=1}^n \sum_{j=1}^n a_i a_j C(\mathbf{s}_i, \mathbf{s}_j) &= \textstyle\sum_{i=1}^n \sum_{j=1}^n a_i a_j \int_{R^2} K_{\mathbf{s}_i}(\mathbf{u}) K_{\mathbf{s}_j}(\mathbf{u}) d\mathbf{u} \\ &= \textstyle\int_{R^2} \sum_{i=1}^n \sum_{j=1}^n a_i K_{\mathbf{s}_i}(\mathbf{u}) a_j K_{\mathbf{s}_j}(\mathbf{u}) d\mathbf{u} \\ &= \textstyle\int_{R^2} \sum_{i=1}^n a_i K_{\mathbf{s}_i}(\mathbf{u}) \sum_{j=1}^n a_j K_{\mathbf{s}_j}(\mathbf{u}) d\mathbf{u} \\ &= \textstyle\int_{R^2} (\sum_{i=1}^n a_i K_{\mathbf{s}_i}(\mathbf{u}))^2 d\mathbf{u} \\ &\geq 0. \end{aligned} \tag{5}$$

The key to achieving positive-definiteness is that each kernel K is solely a function of its own location. Apart from this condition, the exact structure of the kernel can be arbitrary.

Paciorek (2003) proved the closed form of the HSK covariance for Gaussian kernels is:

$$C(\mathbf{s}_i, \mathbf{s}_j) = \frac{1}{(2\pi)^{\frac{P}{2}} |\Sigma_i + \Sigma_j|^{\frac{1}{2}}} \exp(-\frac{1}{2}(\mathbf{s}_i - \mathbf{s}_j)^T (\Sigma_i + \Sigma_j)^{-1} (\mathbf{s}_i - \mathbf{s}_j) \tag{6}$$

Examining the exponential and its quadratic form in Eq. (6), it is a Gaussian stationary covariance function, but in place of the squared Mahalanobis (1936) distance

$(\mathbf{s}_i - \mathbf{s}_j)^T \Sigma^{-1} (\mathbf{s}_i - \mathbf{s}_j)$ for an arbitrarily fixed positive-definite matrix Σ, a quadratic form with the average of the kernel matrices for the two locations is used.

In spatial statistics, the Mahalanobis (1936) distance is a distance measure based on correlations between variables by which different patterns can be identified and analysed. It is a useful way of determining similarity of an unknown sample set to a known one. It differs from the Euclidean distance in that it takes into account the correlations of the data set and is scale-invariant.

Formally, the Mahalanobis distance from a group of values with mean $\mu = (\mu_1, \mu_2, \mu_3, \cdots, \mu_p)^T$ and covariance matrix Σ for a multivariate vector $\mathbf{s} = (s_1, s_2, s_3, \cdots, s_p)^T$ is defined as:

$$D_M(\mathbf{s}) = \sqrt{(\mathbf{s} - \mu)^T \Sigma^{-1} (\mathbf{s} - \mu)} \tag{7}$$

The Mahalanobis distance can also be defined as a dissimilarity measure between two random vectors $\mathbf{s}_i$ and $\mathbf{s}_j$ of the same distribution with the covariance matrix Σ :

$$D_M(\mathbf{s}_i, \mathbf{s}_j) = \sqrt{(\mathbf{s}_i - \mathbf{s}_j)^T \Sigma^{-1} (\mathbf{s}_i - \mathbf{s}_j)}. \tag{8}$$

If the covariance matrix is the identity matrix $\Sigma = I$, the Mahalanobis D_M distance reduces to the Euclidean distance D:

$$D(\mathbf{s}_i, \mathbf{s}_j) = \sqrt{(\mathbf{s}_i - \mathbf{s}_j)^T (\mathbf{s}_i - \mathbf{s}_j)}. \tag{9}$$

If the kernel matrices are constant, we recover the special case of the Gaussian covariance function based on the Mahalanobis distance. If they are not constant with respect to $\mathbf{s}$, the evolution of the kernel covariance matrices in $\mathbb{R}^2$ space produces a non-stationary covariance function.

Remark 3: This generalised kernel convolution covariance model can be extended for use on the sphere, S_2, and other non-Euclidean spaces. On the sphere, the equivalence of translation and rotation causes difficulty in defining kernels that produce correlation behaviour varying with direction, i.e., modelling anisotropy (Paciorek 2003).

Now we consider the quadratic form

$$Q_{ij} = (\mathbf{s}_i - \mathbf{s}_j)^T [\frac{1}{2}(\Sigma_i + \Sigma_j)]^{-1} (\mathbf{s}_i - \mathbf{s}_j), \tag{10}$$

which is at the heart of the covariance function in Eq. (6) constructed via the kernel convolution method. The HSK non-stationary covariance function is the Gaussian

covariance with this new quadratic form in place of the Mahalanobis distance. This relationship raises the possibility of producing a non-stationary version of an isotropic covariance function by using Q_{ij} in place of $\mathbf{r}^2 = \mathbf{r}^T\mathbf{r}$.

Remark 4: Stationary and isotropic covariance functions can be generalised to anisotropic covariance functions that account for directionality by using the Mahalanobis distance between locations $\mathbf{s}_i$ and $\mathbf{s}_j$,

$$r(\mathbf{s}_i, \mathbf{s}_j) = \sqrt{(\mathbf{s}_i - \mathbf{s}_j)^T \Sigma^{-1} (\mathbf{s}_i - \mathbf{s}_j)} \tag{11}$$

where Σ is a positive-definite matrix, rather than $\Sigma = I$, which gives the Euclidean distance and thus isotropy. For the anisotropic case,

$$\Sigma = \Theta(\alpha)\Delta(d_{min}, d_{max})\Theta(\alpha)^T \tag{12}$$

where $\Theta(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$ and $\Delta = diag(d_{min}^2, d_{max}^2)$.

The direction of anisotropy is specified by an azimuth angle, α. The correlation length in the horizontal maximum direction is specified by d_{max}. The range in the perpendicular direction, or the horizontal minimum direction, is specified by d_{min}. Anisotropic covariance functions have been used in Kriging software like GSLIB for many years (e.g., Deutsch and Journel 1998; Chilès and Delfiner 1999). Also, Kearsley (1977) and Duquenne et al. (2005) tried azimuth-dependent covariances in LSC.

Table 2 near here

In practice, one uses $\frac{\mathbf{r}}{d} = \sqrt{Q_{ij}}$ (cf. Table 2), since the correlation length d is redundant and can be absorbed into the kernel matrices, Σ_i and Σ_j, which are allowed to vary in size during the modelling. The non-stationary covariance is non-negative and decays with distance, but can decay at a different rate in different parts of space, thus allowing the degree of smoothing to vary (Paciorek 2003).

Table 3 near here

Table 3 compares distance measures for covariance functions in terms of stationarity. The following general result applies in particular to covariance functions that are positive-definite in the Euclidean space of $\mathbb{R}^n$, in particular the exponential and Gaussian covariance functions (Paciorek and Schervish 2006). Non-stationary versions of the most commonly used covariance functions in planar LSC theory are illustrated in Table 2.

4 Implementation of non-stationary covariances in LSC

A key point in the kernel convolution method is how to define the kernels at each point. Two different approaches have been proposed: one is eigendecomposition parameterisation (Paciorek 2003), and the other is to represent the Gaussian kernels as ellipses of constant probability density (Higdon et al. 1999). Both are based on a fully Bayesian model, with prior distributions on the kernel matrices that determine the non-stationary covariance structure. The eigendecomposition approach extends more readily to higher dimensions, which may be of interest for spatial data treatment in 3D and more general nonparametric regression problems (Paciorek and Schervish 2006).

Here, however, we describe the representation of kernels as ellipses, similarly to what HSK proposed. Instead of using Bayesian inference, which needs complicated and somewhat unfamiliar numerical approaches like the Monte Carlo Markov Chain method (e.g., Robert and Casella 2004), we specify parameters of the ellipses based on empirical covariance functions that are already used in standard LSC theory. We believe that this simpler strategy might make the approach more attractive to others for practical implementation in existing LSC software.

The idea of modelling kernels as ellipses comes from the visualisation of a bivariate Gaussian distribution in 3D. If we approach this distribution from above (from the positive z-direction), any $x - y$ plane that has constant z and intercepts the bivariate normal distribution at any point, except the apex, will form an ellipse in the plane of the bivariate kernel that it cuts. This allows us to use an ellipse (Fig. 2) as a way to more intuitively parameterise Σ (cf. Higdon et al. 1999; Swall 1999).

Figure 2 near here

The location and geometry of the ellipse adopted here can be defined through three parameters: its geometrical centre, major axis and minor axis. Since each ellipse represents a kernel for a certain location, the centre of the ellipse is at the point $\mathbf{s}$. In order to allow the ellipses to vary in size, shape and orientation, we can manipulate the major a and minor b axes and azimuth angle α;

Swall (1999) found over-fitting and mixing problems when the ellipse's area (i.e., πab) was allowed to vary, so HSK fixed the ellipse's area at a constant value common to all locations, but instead applied a scaling factor τ for each ellipse to vary at each

location:

$$\Sigma^{\frac{1}{2}} = \tau \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \tag{13}$$

where τa and τb are the semi-axes of the ellipse and α is the direction angle (measured anticlockwise from the x-axis) of the major axis of the ellipse.

Equation (13) provides the correspondence between ellipses and the bivariate Gaussian kernels that they aim to represent. The stationary and non-stationary covariance structures based on elliptical kernels are illustrated schematically in Fig. 3.

Figure 3 near here

Finally, for practical application of the kernel convolution method based on ellipses centred at each location using empirical covariance functions, we propose the following general sequence that can be to applied other cases.

- Partition the non-stationary study area into smaller parts that are as stationary as possible (e.g., east and west for the Darling Fault case)
- Estimate the empirical covariance functions in each partition, including empirical covariance map and directional empirical covariance functions;
- Set $\Sigma_i = \Sigma_{A(i)}$, where $A(i)$ denotes the region in which location i falls and $\Sigma_{A(i)}$ is constructed for each region from the parameters, $\{\alpha, a, b\}_{A(i)}$, of the anisotropic correlation structure estimated for the region;
- Having the parameters for the ellipses at each location, the kernels are then estimated from Eq. (13);
- Based on the kernels at each location, the covariances between any two points are estimated from the non-stationary covariance models in Table 2

In the next section, we will explain how to define the parameters of the elliptical kernels for each location based on empirical covariances for our chosen example of gravity anomaly interpolation by non-stationary 2D LSC.

5 Case-study: stationary versus non-stationary covariances

5.1 The Darling Fault's gravity field

This case-study refers to a standard interpolation problem, namely the construction of a gravity anomaly grid from point data. We have narrowed our focus to a reasonably small regular grid of data, where a planar approximation is sufficient (cf. Moritz 1980).

Gravity observations over the Darling Fault near Perth in Western Australia (cf. Lambeck 1987; Kirby 2003) provide a nice illustration of the application of non-stationary covariance functions in 2D LSC gridding. To a first approximation, this region is divided between low free-air gravity anomalies in the west, and high free-air gravity anomalies in the east (Fig. 4), which shows the non-stationarity in this case-study data set. These are residual free-air gravity anomalies based on the degree/order 360/360 expansion of EGM96 (Lemoine et al. 1998).

Figure 4 near here

We acknowledge that terrain-reduced gravity data are normally used for gridding because they are generally smoother (cf. Goos et al. 2003). We have deliberately chosen to use residual free-air anomalies here because they are clearly non-stationary, thus posing more of a challenge to LSC. However, applying terrain reductions to the Darling Fault gravity data will still not make them stationary. In this area, the Bouguer slab reduction is around 30 mGal over the fault and the terrain correction is less than 2 mGal east of the Fault (Kirby and Featherstone 2002). Compared to the 100 mGal difference in free air gravity anomalies (Fig. 4), terrain reductions will not remove the non-stationarity in this case.

Our residual free-air gravity anomaly prediction used the eight nearest neighbouring points with stationary and non-stationary covariance functions. This choice of nearest points is consistent with many other LSC-based predictions. However, the mean value of the LSC observation vector in each neighbourhood is different in this dataset (cf. Fig. 4). This violates the LSC assumption of a zero mean (Moritz 1980).

As mentioned in the Introduction, OK or GK consider non-stationarity of the mean, but this is not the subject of this paper, which instead focuses on non-stationarity via the covariance function. As such, we deliberately added 100 mGal to all the residual

free-air anomalies to make them positive over the whole region, giving an always-positive mean value of the observation vector in each neighbourhood.

Therefore, all points tested are affected approximately equally by non-stationarity of the mean. This strategy allows us to examine the non-stationarity of covariances in LSC, where the non-stationarity of mean is approximately equal, thus not contaminating our interpretations, as would be the case for spatially varying means for each neighbourhood. For a study of non-stationarity of the mean, see the GK approach (Reguzzoni et al. 2005).

5.2 Global stationary covariances

We start with an analysis of a stationary global empirical covariance function using the whole data-set (cf. Sect. 2) in Fig. 4. A stationary Cauchy covariance model was fitted to the empirical covariances using EMPCOV. Because of the large systematic difference between free-air gravity anomalies in the east and west, the correlation length is 0.3° (cf. Fig. 5 and Table 4). As such, a stationary global empirical covariance function is clearly inappropriate for this non-stationary test data-set.

Figure 5 near here

5.3 Segmented stationary covariances

Next, we used the segmentation approach, where the data were divided between eastern and western sections separated along the well-defined Darling Fault (cf. Fig. 4). Figure 6 shows the stationary Cauchy covariance models and empirical covariances in the east and west. The objective of presenting each covariance is to evaluate (and visualise) the extent to which the covariance function varies locally (i.e., non-stationarity of the data set versus when treated as a whole; Fig. 5). The key point is that the covariances vary markedly either side of the Fault (cf. Fig. 4). From Fig. 6, the western region has the greater variance and correlation length.

Figure 6 near here

These significant differences show further that a global stationary covariance function is an inadequate representation of the spatial variations across this case-study

region. Also, recall that the use of segmented stationary covariance functions will introduce the problem of how to subsequently patch them. While this segmented approach delivers some improvement over the use of global variances, it will be shown that the non-stationary approach is better still.

Table 4 near here

The parameters of the empirical Cauchy covariance functions for each of the global, and east and west segments are summarised in Table 4. We will use them later to help design the elliptical kernels at each location for the non-stationary covariance modelling.

5.4 Anisotropy

We now attempt to detect anisotropy in the case-study region using empirical covariances in different azimuths $(0^\circ, 45^\circ, 90^\circ, 135^\circ)$ (Fig. 7). The method developed by Forsberg (1986), using the Fourier transform and the so-called anisotropy index, could be used. Instead, however, we used the more simple and straightforward methods that are popular in geostatistics (e.g., Chilès and Delfiner 1999, Deutsch and Journel 1998, Goovaerts 1997).

Usually the covariances are calculated in several different directions, which are displayed separately. Sometimes the lines of constant covariation are graphed in 2D. This representation of the covariances in all directions can be a useful visual tool to highlight possible anisotropy in the data, which is called a covariance map (i.e., empirical covariances in 2D) (Fig.8). The principle is to define a grid such that the origin of the space is located at the centre of this grid. Each pair of samples corresponds to a distance and a direction, which can be attributed to a grid cell, and thus to a variability.

From Figure 8, there is anisotropy approximately in the north-south direction (azimuth 20°). Figure 7 also demonstrates that, except for $\alpha = 20^\circ$, the empirical covariances are largely similar in terms of variance and correlation length.

Figure 7 near here

Figure 8 near here

5.5 Non-stationary covariances

To construct the elliptical kernels at each location (cf. Fig. 2), we first fixed α parameter to 20° for the entire region because of the anisotropy detected from Fig. 8. From the estimated correlation lengths in Table 4), we chose $a = 0.3^\circ = \sqrt{d_{max}}$ and $b = 0.2^\circ = \sqrt{d_{min}}$ to elongate the elliptical kernels in the $\alpha = 20^\circ$ azimuth. Based on our numerous empirical tests (and Paciorek and Schervish 2006, Swall 1999), we shrink the size of elliptical kernels from east to west by decreasing the τ scale factor from 1 to 0.6. The resulting elliptical kernels are plotted in Fig. 9.

Figure 9 near here

For this non-stationary case, the Cauchy non-stationary model of Table 2 was applied. As seen in Table 2, unlike stationary covariance models that are a function of d and C_0, the non-stationary covariance model is a function of the elliptical kernels Σ at each location (Eq. 13). In our case, these kernels have the elliptical forms in Fig. 9. For estimation of covariances between two points i and j, we insert the ellipse parameters for each point from Fig. 9 into Eq.(13) and then the non-stationary Cauchy model from Table 2.

5.6 Interpolation error estimates

Two types of error were computed to assess the accuracy of the gravity anomaly interpolation using stationary and non-stationary covariances in 2D LSC.

1. The first is the formal standard (or internal) error of LSC as obtained from the error covariance matrix; this is:

$$(\sigma^F_{\Delta g})^2 = C_0 - C^{PQ}(C^{PP'})^{-1}(C^{PQ})' \tag{14}$$

where C_0 is the signal variance, C^{PQ} is the cross covariance of the predicted point Q and the observation point P, and $C^{PP'}$ is the auto-covariance of observation points (Moritz 1980).

2. The second is the prediction error estimated from cross-validation (Featherstone and Sproule 2006), where one point is omitted from each LSC prediction, then the predicted value at that point is compared with the observed value via:

$$\sigma^E_{\Delta g} = \Delta g_{obs} - \Delta g_{pred} \tag{15}$$

That is, each test point used is not used in each LSC prediction, and the statistics computed for all data points in the test area. While time-consuming, this does give an external error estimate.

Figures 10 and 11 compare the internal errors from LSC covariance propagation and the external errors from cross-validation for stationary and non-stationary LSC prediction, respectively, of the residual free-air gravity anomalies. Note the different scales in each panel. Tables 5 and 6 give a summary of the respective statistics for all points in the test data-set.

Figures 10 and 11 near here

Tables 5 and 6 near here

Firstly, Tables 5 and 6 and Figs. 10 and 11 together illustrate the advantage of using non-stationary covariances over the stationary covariances for our test region, which has already been shown to be anisotropic and non-stationary. Clearly, the application of non-stationary covariances to non-stationary data is more appropriate (and better) than assuming a stationary covariance.

Another key observation from Tables 5 and 6 and Figs. 10 and 11 is that the internal errors from LSC are smaller than the external errors from cross-validation. In particular, note the need to use different scales in the left and right panels of Figs. 10 and 11. This is of some concern, showing that the internal LSC error propagation may be far too optimistic by nearly an order of magnitude. However, there is a marginally better agreement between the internal and external errors for the non-stationary case.

In Fig. 11, the magnitude of errors (both internal and external) increase along the Fault (cf. Fig. 4), which shows that the non-stationary covariance modelling (especially for the internal errors) better reflects the difficulty in accurately interpolating data in areas of steep gradients. Basically, the internal error estimates, even though too optimistic in relation to the external error estimates, do reflect the worse prediction that is intuitively expected along the Darling Fault. On the other hand, the stationary covariance modelling tends to 'smear out' the error estimates, thus not reflecting the Fault's location.

Finally, the internal errors in both non-stationary and stationary cases are larger along the borders (Figs. 10 and 11), showing an edge effect in the latter that this is not reflected in the former. Therefore, LSC internally propagated errors, while also being too optimistic, are less reliable at the edges of the areas used. This could cause further

problems if trying to patch segmented stationary LSC solutions based on internally propagated errors.

6 Conclusions and Recommendations

We have introduced the kernel convolution method, borrowed from the discipline of spatial statistics (Higdon et al. 1999), to provide non-stationary covariance functions for geodetic 2D LSC. Conceptually, these are elliptical kernels that orient and scale according to the level of anisotropy and non-stationarity in the data.

As an example, non-stationary and stationary 2D LSC were used to interpolate residual (to EGM96) free-air gravity anomalies over the Darling Fault in Western Australia, 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 field varies rapidly.

While we have considered the rather simple case of 2D LSC interpolation and prediction, it is clear that non-stationary covariance modelling shows great promise, which might be beneficial for other geodetic applications. However, there is often the need to ensure that the covariance functions are harmonic for, e.g., geoid field modelling. There are numerous other LSC applications in geodesy where non-stationarity should not be ignored, such as the treatment of satellite altimeter data (e.g., Fieguth et al. 1995) and fitting quasi/geoids to GPS-levelling data.

We therefore recommend that non-stationarity in geodetic data is given more consideration than at present.

Although in Paciorek (2003) there is a possible development of kernel convolution approach for non-stationary covariances on the sphere, the harmonicity of this method should be studied further for expansion of non-stationary covariances for different quantities of gravity field for points having the spherical coordinates (φ, λ, h).

Also, non-stationary covariance functions provide a more flexible tool to study density-gravity covariance functions (Tscherning 1991), which leads to the solution of the inverse gravimetric problem in geophysical exploration because a non-stationary mass density distribution is more realistic than a stationary one.

Acknowledgements We wish to thank Dr Chris Paciorek for his many valuable comments and advice, and for providing his PhD thesis on non-stationary covariance functions. We also thank the reviewers C.C. Tscherning, F. Sansó and one anonymous for their constructively critical reviews. ND thanks Curtin University of Technology for a CIRTS scholarship. WEF thanks the Australian Research Council for grant DP0663020.

References

Armstrong M (1998) Basic linear geostatistics, Springer, Berlin Heidelberg New York

Atkinson PM, Lloyd CD (2007) Non-stationary variogram models for geostatistical sampling optimisation: an empirical investigation using elevation data, Comput & Geosci 33:1285-1300, doi: 10.1016/j.cageo.2007.05.011

Chilès JP, Delfiner P (1999) Geostatistics, Wiley, New York

Cressie N (1993) Statistics for spatial data, Wiley, New York

Dermanis A (1984) Kriging and Collocation - A Comparison, Manuscr Geod 9:159-167

Deutsch CV, Journel AG (1998) GSLIB, Oxford University Press, Oxford

Duquenne H, Everaerts M, Lambot P (2005) Merging a gravimetric model of the geoid with GPS/levelling data: an example in Belgium. In: Jekeli C, Bastos L, Fernandes J (eds), Gravity, Geoid and Space Missions, Springer, Berlin Heidelberg New York, pp 131-136

Featherstone WE, Sproule DM (2006) Fitting AusGeoid98 to the Australian Height Datum using GPS-levelling and least squares collocation: application of a cross-validation technique, Surv Rev 38(301):573-582

Fieguth PW, Karl WC, Willsky AS, Wunsch C (1995) Multiresolution optimal interpolation and statistical analysis of TOPEX/Poseidon satellite altimetry, IEEE Transactions on Geoscience and Remote Sensing 33(2):280-292, doi: 10.1109/36.377928

Flury J (2006) Short-wavelength spectral properties of the gravity field from a range of regional data sets, J Geod 79(10-11):624-640, doi: 10.1007/s00190-005-0011-y

Forsberg R (1986) Spectral Properties of the Gravity Field in the Nordic Countries, Boll Geodesia e Sc Aff XLV:361-384

Fuentes M (2001) A high frequency Kriging approach for non-stationary environmental processes, Environmetrics 12(5):469-483, doi: 10.1002/env.473

Fuentes M, Smith R (2003) A new class of non-stationary spatial models, Tech Rep, Department of Statistics, North Carolina State Univ

Gaposchkin EM (1973) Standard Earth III-1973, Special Rep 353, Smithsonian Astrophysical Observatory, Cambridge

Goad CC, Tscherning CC, Chin MM (1984) Gravity empirical covariance values for the continental United States, J Geophys Res 89(B9):7962-7968

Goos JM, Featherstone WE, Kirby JF, Holmes SA (2003) Experiments with two different approaches to gridding terrestrial gravity anomalies and their effect on regional geoid computation, Surv Rev 37(288):92-112

Goovaerts P (1997) Geostatistics for natural resources evaluation, Oxford University Press, Oxford

Higdon D, Swall J, Kern J (1999) Non-stationary spatial modelling. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds) Bayesian Statistics 6. Oxford University Press, Oxford, pp.761-768

Kearsley W (1977) Non-stationary estimation in gravity prediction problem, Rep 256, Dept of Geodetic Sci, The Ohio State Univ, Columbus

Keller W (1998) Collocation in reproducing kernel Hilbert spaces of a multiscale analysis. Phys Chem Earth 23(1):25-29

Keller W (2000) A wavelet approach to non-stationary collocation. In: Schwarz KP (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 MG (ed), Gravity, Geoid and Geodynamics 2000, Springer, Berlin Heidelberg New York, pp. 79-84

Kirby JF, Featherstone WE (2002) High-resolution grids of gravimetric terrain correction and complete Bouguer corrections over Australia, Explor Geophys 33: 161-165.

Kirby JF (2003) On the combination of gravity anomalies and gravity disturbances for geoid determination in Western Australia, J Geod 77(7-8):433-439, doi: 10.1007/s00190-003-0334-5

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, Sideris MG (1999) The long road from deterministic collocation to multiresolution approximation, Rep. 1999.5, Dept of Geodesy and Geoinformatics, Stuttgart Univ

Kotsakis C (2000) The multiresolution character of collocation. J Geod 74(3-4):275-290, doi: 10.1007/s001900050286

Krarup T (1969) A contribution to the mathematical foundation of physical geodesy, Rep 44, Danish Geodetic Institute, Copenhagen

Krige DG (1951) A statistical approach to some basic mine valuation problems on the Witwatersrand, J Chemical, Metallurgical and Mining Society of South Africa 52:119139

Lambeck K (1987) The Perth Basin: a possible framework for its formation and evolution, Exploration Geophysics 18(2):124-128, doi: 10.1071/EG987124

Lemoine FG, Kenyon SC, Factor JK, Trimmer RG, Pavlis NK, Chinn DS, Cox CM, Klosko SM, Luthcke SB, Torrence MH, Wang YM, Williamson RG, Pavlis EC, Rapp RH , Olson TR (1998) The development of the joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) geopotential model EGM96, Technical Report NASA/TP-1998- 206861, National Aeronautics and Space Administration, Greenbelt, 575 pp.

Mahalanobis PC (1936) On the generalized distance in statistics, Proc of the National Institute of Science of India 12:49-55

Matheron G (1962) Traité de Géostatistique Appliquée, Tome I, Mémoires du Bureau de Recherches Géologiques et Minières, No. 14, Editions Technip, Paris

Moreaux G, Tscherning CC, Sansò F (1999) Approximation of harmonic covariance functions on the sphere by non harmonic locally supported functions, J Geod 73: 555-567

Moritz H (1980) Advanced Physical Geodesy, Abacus, Tunbridge Wells

Nychka DW, Wikle C, Royle JA (2002) Multiresolution models for non-stationary spatial covariance functions. Statistical Modelling 2(4):315-331, doi: 10.1191/1471082x02st037oa

Paciorek CJ (2003) Non-stationary Gaussian processes for regression and spatial modelling, PhD thesis, Carnegie Mellon University, Pittsburgh

Paciorek CJ, Schervish MJ (2006) Spatial modelling using a new class of non-stationary covariance functions, Environmetrics 17(5): 483-506, DOI 10.1002/env.785

Rapp RH (1964) The prediction of point and mean gravity anomalies through the use of the digital computer, Rep 43, Dept of Geodetic Sci, The Ohio State Univ, Columbus

Reguzzoni M, Sansò F, Venuti G (2005) The theory of general Kriging, with application to the determination of a local geoid, Geophys J Int 162(4):303-314, doi: 10.1111/j.1365-246X.2005.02662.x

Robert CP, Casella G (2004) Monte Carlo Statistical Methods (second edition), Springer, Berlin Heidelberg New York

Sampson PD, Guttorp P (1992) Nonparametric estimation of non-stationary spatial covariance structure, J Amer Statist Assoc 87(417):108-119, doi: 10.2307/2290458

Sampson P, Damian D, Guttorp P (2001) Advances in modeling and inference for environmental processes with non-stationary spatial covariance. In: Monestiez P, Allard D, Froidvaux R (eds), GeoENV 2000: Geostatistics for Environmental Applications, Kluwer, Dordrecht, pp. 17-32

Sansò F, Venuti G and Tscherning CC (2000) A theorem of insensivity of the collocation solution to variations of the metric in the interpolation space. In: Schwarz KP (ed), International Association of Geodesy Symposia, Vol. 121, Springer Verlag, pp. 233-240

Schaffrin B, Felus Y (2005) On Total Least-Squares Adjustment with Constraints. In: Sansò F (ed), International Association of Geodesy symposia, Vol. 128, SpringerVerlag, pp. 417-421

Schwarz KP, Lachapelle G (1980) Local characteristics of the gravity anomaly covariance function, Bull Géod 54(1):21-36

Swall JL (1999) Non-stationary spatial modelling using a process convolution approach, PhD thesis, Institute of Statistics and Decision Sciences, Duke University, Durham

Tscherning CC, Rapp RH (1974) Closed covariance expressions for gravity anomalies, geoid undulations, and deflections of the vertical implied by anomaly degree variance models, Rep 208, Dept of Geodetic Sci, The Ohio State Univ, Columbus

Tscherning CC (1975) Application of collocation for the planning of gravity surveys, J Geod 49 (2): 183-198

Tscherning CC, Sansò F, Arabelos D (1987) Merging regional geoids - preliminary considerations and experiences, Boll di Geodesia e Sci Aff XLVI(3):191-206

Tscherning CC (1991) Density-gravity covariance functions produced by overlapping rectangular blocks of constant density, Geophys J Int 105(3):771-776, doi: 10.1111/j.1365-246X.1991.tb00811.x

Tscherning CC (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 CC (1999) Construction of anisotropic covariance functions using Riesz-representers, J Geod 73(6): 332-336, doi: 10.1007/s001900050250

Vyskočil V (1970) On the covariance and structure functions of anomalous gravity field, Stud Geophys et Geod 14(2):174-177, doi: 10.1007/BF02585616

Xu PL (1991) Least squares collocation with incorrect prior information, Z Verm 116:266-273

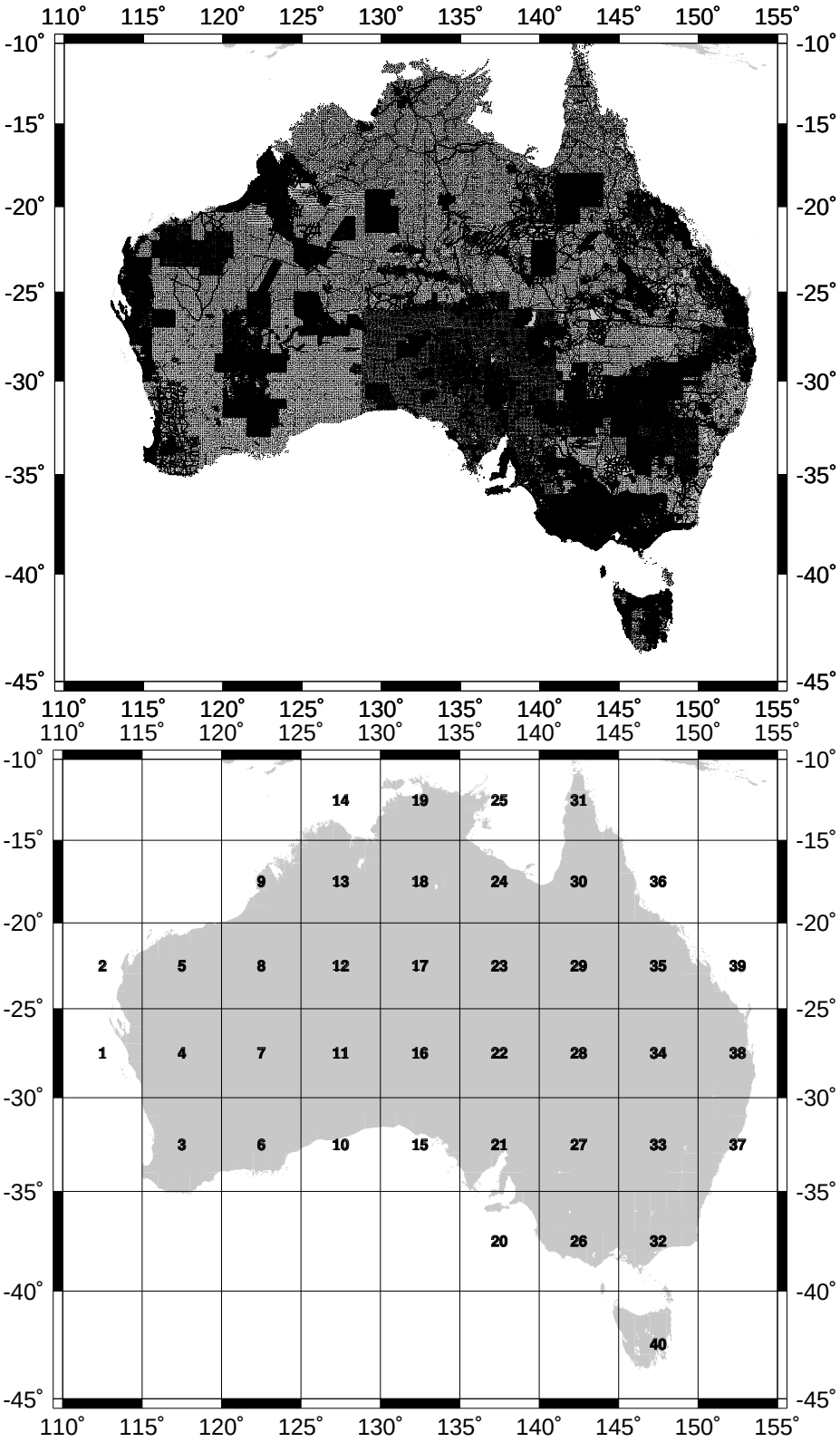

Fig. 1 Land gravity observation coverage over Australia (top); Thirty-nine (blocks 24 and 25 were merged) $5° \times 5°$ blocks used to estimate local empirical covariances (bottom) [Mercator projection]

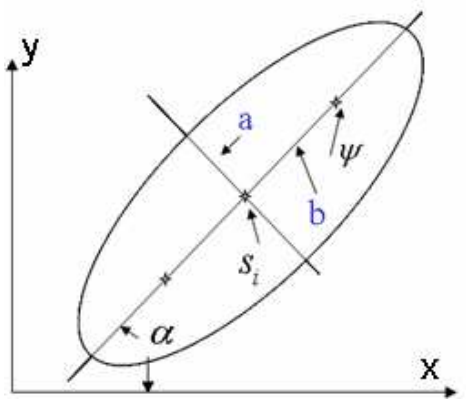

Fig. 2 Geometry and parameters of the ellipse used to parameterise the non-stationary kernels

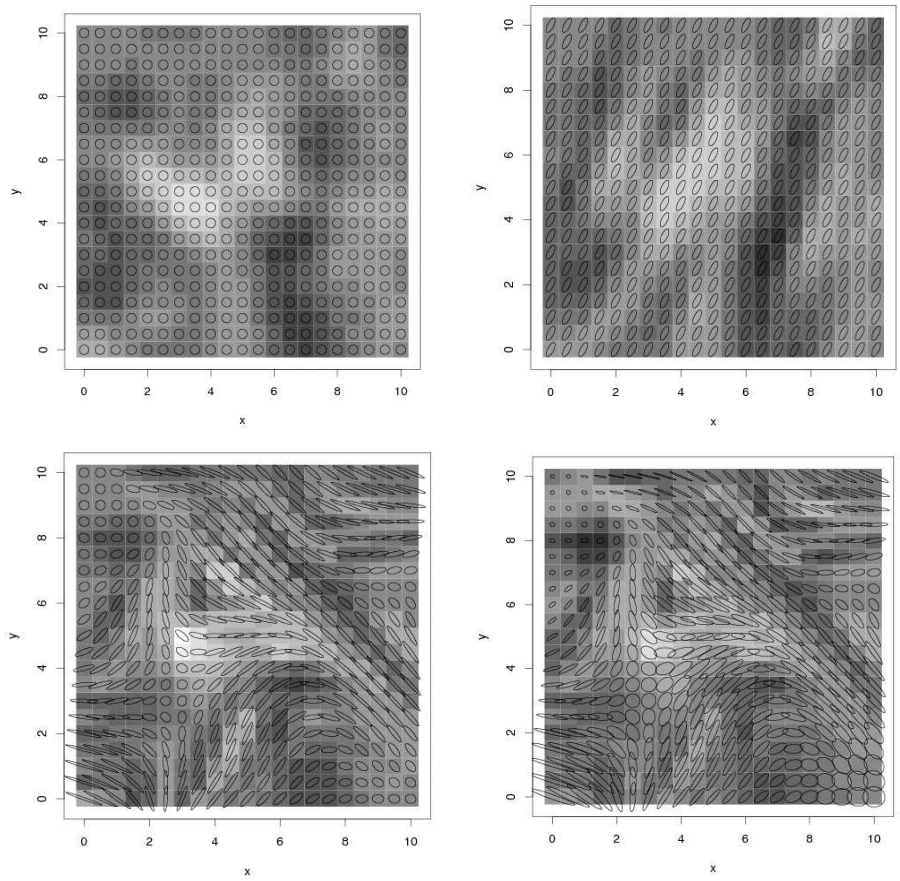

Fig. 3 Different forms of the kernels defined by Eq.(13): identical isotropic kernels representing stationary covariance functions (top-left); identical anisotropic kernels representing stationary anisotropic covariance functions (top-right); kernels varying in orientation representing non-stationary covariance functions (bottom-left); kernels varying in orientation and size representing non-stationary covariance functions (bottom-right) (adapted from Swall 1999)

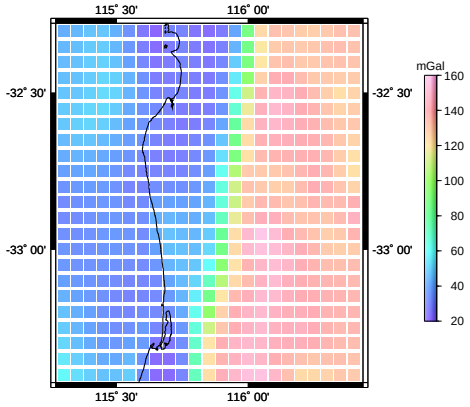

Fig. 4 Residual free-air gravity anomalies (in mGal) across the Darling Fault, Western Australia, showing non-stationarity where they are systematically lower to the west than the east [Mercator projection]

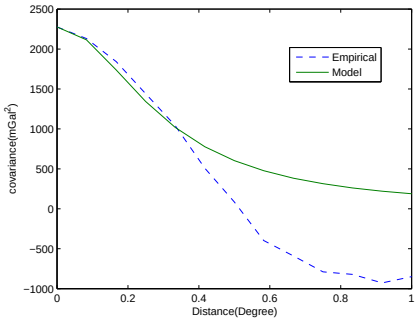

Fig. 5 Stationary Cauchy model and empirical covariances of residual free-air gravity anomalies for the whole Darling Fault test area

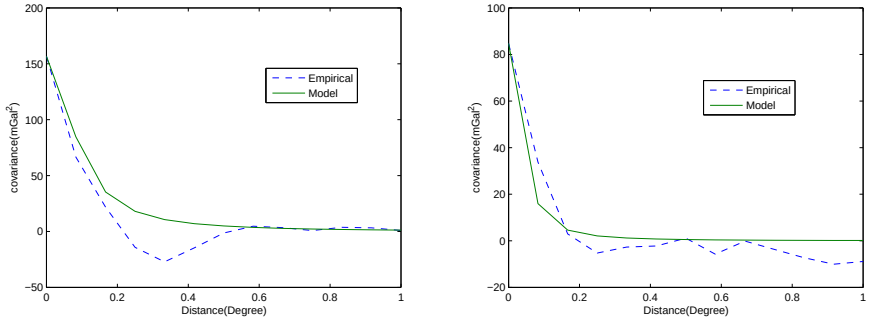

Fig. 6 Segmented stationary Cauchy empirical covariances for the eastern (left) and western (right) parts of the Darling Fault test area

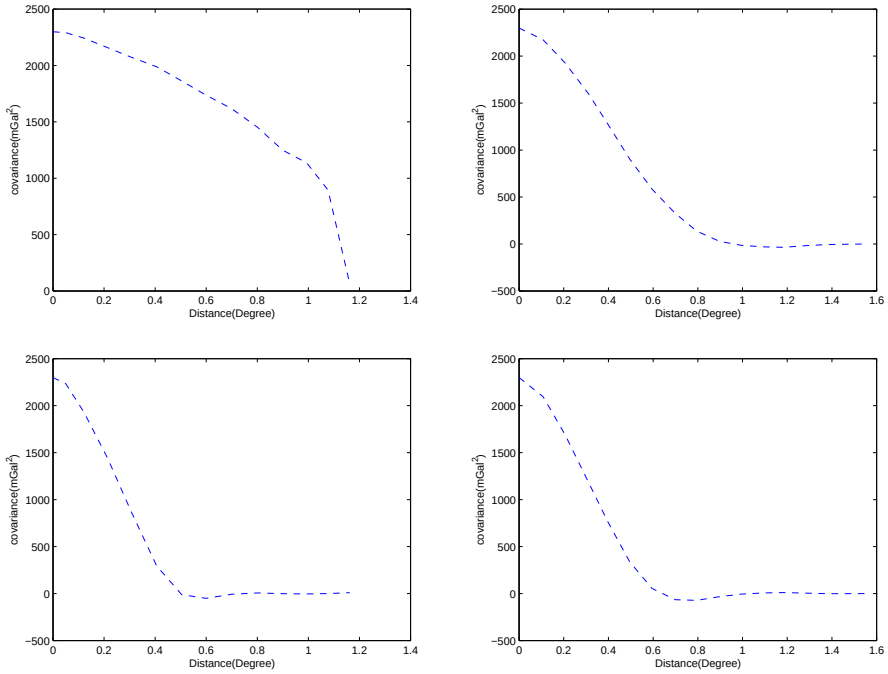

Fig. 7 Cauchy empirical covariances for azimuth 0° (upper left), 45° (upper right), 90° (lower left), 135° (lower right)

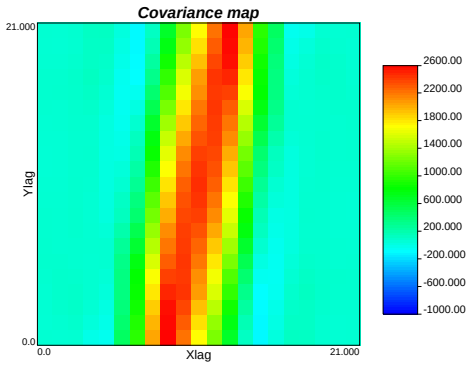

Fig. 8 Covariance map of the residual free-air gravity anomalies (in mGal^2) over the Darling Fault test area [linear projection]

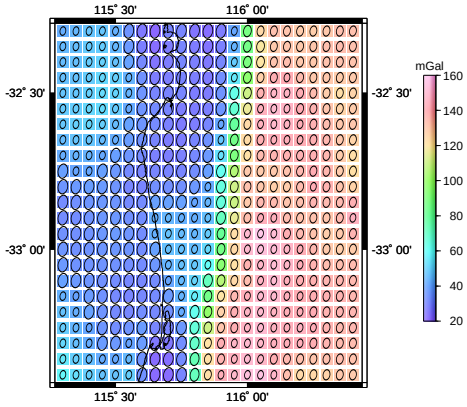

Fig. 9 Elliptical kernels attributed to each location and used to construct the non-stationary covariances for LSC interpolation. The underlying image shows the residual free-air gravity anomalies after adding 100 mGal [Mercator projection]

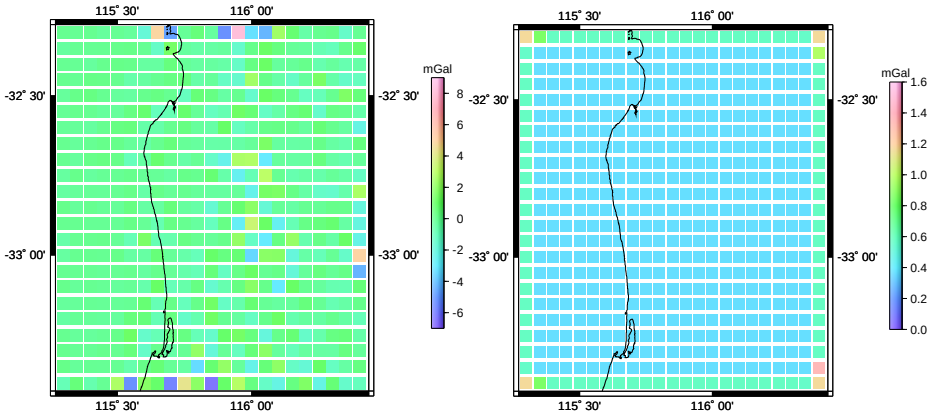

Fig. 10 Stationary LSC interpolation errors from external cross-validation (left) and from internal LSC covariance propagation (right) [Mercator projection]

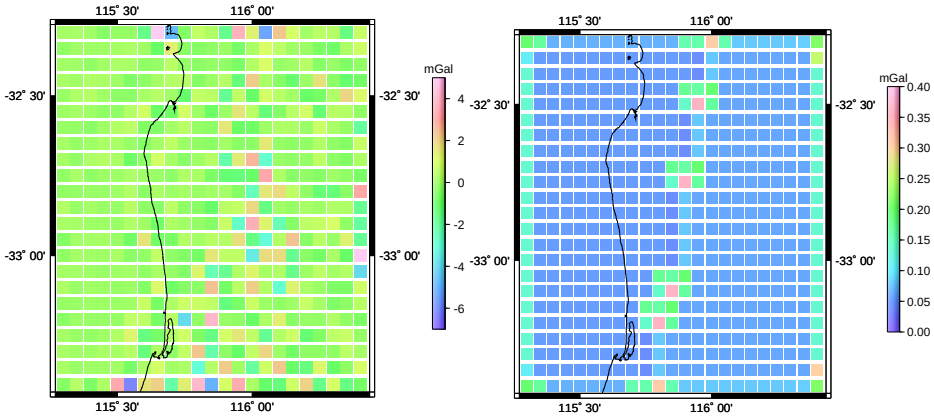

Fig. 11 Non-stationary LSC interpolation errors from external cross-validation (left) and from internal LSC covariance propagation (right) [Mercator projection]

Table 1 Local empirical covariance parameters from EMPCOV of free-air gravity anomalies over Australia residual to EGM96; cf. Fig. 1

Area	No. of data	Mean (mGal)	C_0 (mGal^2) prior variance	$\hat{C}_0$ (mGal^2) variance from EMPCOV	$\hat{d}$ (°)
1	10409	8.62	501.1	541.98	0.76
2	13559	7.69	397.7	458.32	0.59
3	16371	6.38	1590.5	1465.58	0.81
4	18726	-17.49	753.7	1848.09	1.11
5	92340	10.66	347.4	345.02	0.95
6	33636	7.45	292.5	208.75	0.53
7	28234	1.01	219.3	211.97	0.74
8	21759	-6.69	205.9	285.58	0.64
9	34166	-5.45	213.0	246.63	0.84
10	1764	-0.81	250.2	181.59	0.63
11	29508	-3.88	347.6	224.72	0.69
12	12846	-9.40	430.4	913.79	0.94
13	11191	-7.66	477.1	571.12	0.84
14	2343	-29.36	320.0	1497.23	1.00
15	14452	2.07	340.6	320.93	0.56
16	33486	0.02	1305	1226.73	1.41
17	26635	-23.89	1188.6	2200.58	1.20
18	26819	7.74	97.9	130.40	0.90
19	5908	1.70	144.1	67.76	0.53
21	127460	-1.18	238.1	169.16	0.58
22	57644	4.10	348.5	903.50	1.01
23	16194	6.12	242.8	310.62	0.79
24+25	8065	4.76	164.8	233.62	0.78
26	77915	10.83	307.1	520.43	0.68
27	80582	9.40	733.6	1461.37	1.04
28	27397	1.86	124.0	115.47	0.87
29	18778	-0.34	345.3	309.21	0.88
30	13407	-4.59	379.3	440.55	1.29
31	3070	21.12	353.3	1282.11	0.90
32	43637	3.79	920.2	434.98	0.75
33	49176	-0.14	191.8	90.27	0.62
34	23626	-5.26	188.6	217.84	0.76
35	27423	-0.40	130.8	89.04	0.74
37	4641	12.49	738.1	1293.03	0.52
38	16144	4.21	369.1	345.40	0.75
39	4155	-1.35	278.7	332.60	0.49
40	61084	33.12	547.2	1793.10	0.59

Table 2 Non-stationary versus stationary covariance functions

Name	stationary	non-stationary
Cauchy (e.g., Moritz 1980)	$C(\mathbf{r}) = \frac{C_0}{1+(\frac{\mathbf{r}}{d})^2}$	$C(\mathbf{s}_i, \mathbf{s}_j) = C_0 \lvert\Sigma_i\rvert^{\frac{1}{4}} \lvert\Sigma_j\rvert^{\frac{1}{4}} \left\lvert \frac{\lvert\Sigma_i\rvert+\lvert\Sigma_j\rvert}{2} \right\rvert^{-\frac{1}{2}} \left(\frac{1}{1+Q_{ij}}\right)$
Gaussian (e.g., Kearsley 1977)	$C(\mathbf{r}) = C_0 \exp\left(-\frac{\mathbf{r}^2}{d^2}\right)$	$C(\mathbf{s}_i, \mathbf{s}_j) = C_0 \lvert\Sigma_i\rvert^{\frac{1}{4}} \lvert\Sigma_j\rvert^{\frac{1}{4}} \left\lvert \frac{\lvert\Sigma_i\rvert+\lvert\Sigma_j\rvert}{2} \right\rvert^{-\frac{1}{2}} \exp(-Q_{ij})$

Table 3 Comparison of distance measures for covariance functions in terms of their stationarity

Covariance function	Distance measure	Parameter
stationary isotropic	$r^2 = (\mathbf{s}_i - \mathbf{s}_j)^T \Sigma^{-1} (\mathbf{s}_i - \mathbf{s}_j)$	$\Sigma = I$
stationary anisotropic	$r^2 = (\mathbf{s}_i - \mathbf{s}_j)^T \Sigma^{-1} (\mathbf{s}_i - \mathbf{s}_j)$	$\Sigma = \Theta(\alpha)\Delta(d_{min}, d_{max})\Theta(\alpha)^T$
non-stationary	$Q_{ij} = (\mathbf{s}_i - \mathbf{s}_j)^T [\frac{1}{2}(\Sigma_i + \Sigma_j)]^{-1} (\mathbf{s}_i - \mathbf{s}_j)$	Σ_i, Σ_j

Table 4 Cauchy model parameter estimates for the global and east-west segmented stationary covariances

	mean ($mGal$)	var (mGal^2)	$\hat{d}$ (°)
global	85.008	2287.889	.3
west	41.665	186.964	.09 (d_{max})
east	134.386	100.088	.04 (d_{min})

Table 5 Statistics from the external cross-validation of the differences between observed and predicted residual free-air gravity anomalies (units in mGal)

	No. points	Min	Max	Mean	STD	RMS
stationary covariance	529	-6.343	8.063	-0.165	1.180	1.192
non-stationary covariance	529	-6.191	5.827	-0.018	1.130	1.131

Table 6 Statistics of internal errors from LSC covariance propagation (units in mGal)

	No. points	Min	Max	Mean	STD	RMS
stationary covariance	529	0.369	1.347	0.412	0.115	0.428
non-stationary covariance	529	0.001	0.302	0.015	0.033	0.087