J Geodesy
DOI 10.1007/s00190-007-0133-5

ORIGINAL ARTICLE

Least-squares collocation with covariance-matching constraints

Christopher Kotsakis

Received: 18 August 2006 / Accepted: 28 December 2006

Abstract Most geostatistical methods for spatial random field (SRF) prediction using discrete data, including least-squares collocation (LSC) and the various forms of kriging, rely on the use of prior models describing the spatial correlation of the unknown field at hand over its domain. Based upon an optimal criterion of maximum local accuracy, LSC provides an unbiased field estimate that has the smallest mean squared prediction error, at every computation point, among any other linear prediction method that uses the same data. However, LSC field estimates do not reproduce the spatial variability which is implied by the adopted covariance (CV) functions of the corresponding unknown signals. This smoothing effect can be considered as a critical drawback in the sense that the spatio-statistical structure of the unknown SRF (e.g., the disturbing potential in the case of gravity field modeling) is not preserved during its optimal estimation process. If the objective for estimating a SRF from its observed functionals requires spatial variability to be represented in a pragmatic way then the results obtained through LSC may pose limitations for further inference and modeling in Earth-related physical processes, despite their local optimality in terms of minimum mean squared prediction error. The aim of this paper is to present an approach that enhances LSC-based field estimates by eliminating their inherent smoothing effect, while preserving most of their local prediction accuracy. Our methodology consists of correcting a posteriori the optimal result obtained from LSC in such a way that the new field estimate matches the spatial correlation structure implied by the signal CV function. Furthermore, an optimal criterion is imposed on the CV-matching field estimator that minimizes the loss in local prediction accuracy (in the mean squared sense) which occurs when we transform the LSC solution to fit the spatial correlation of the underlying SRF.

Keywords Least-squares collocation · Spatial random field · Prediction · Smoothing · Covariance matching

C. Kotsakis (✉)
Department of Geodesy and Surveying,
The Aristotle University of Thessaloniki, University Box 440,
Thessaloniki 54124, Greece
e-mail: kotsaki@topo.auth.gr

1 Introduction

The prediction of the functional values of a continuous spatial random field (SRF), using a set of observed values of the same and/or other SRFs, is a fundamental inverse problem in geosciences. The mathematical model describing such a problem is commonly formulated in terms of the system of observation equations

$$y_i = L_i(u) + v_i, \quad i = 1, 2, \ldots, n \tag{1}$$

where $u(P)$ denotes the primary random field of interest ($P \in D$, with D being a bounded or unbounded, not necessarily Euclidean, spatial domain) that needs to be determined, at one or more points, using n discrete measurements $\{y_i\}$ which are taken on the same and/or other locations. The symbols $L_i(\cdot)$ correspond to bounded linear or linearized functionals of the unknown field, depending on the physical model that relates the observable quantities with the underlying SRF itself. The additive terms $\{v_i\}$ contain the effect of measurement random noise, including possible errors due to model uncertainty in the specification of the field functionals $L_i(\cdot)$ that should not exceed the data noise level.

Typical examples in geodesy that fall within the realm of the aforementioned SRF prediction scheme include the determination of the disturbing gravity potential on or outside the Earth using various types of observed gravity field functionals, the prediction of stationary or non-stationary ocean circulation patterns from satellite altimetry data, the prediction and modeling of atmospheric fields (tropospheric, ionospheric) from the tomographic inversion of GPS data, the de-noising and optimal separation of geodetic and geophysical signals from a given data record, the generation of error correction "meta-surfaces" for network-based mobile positioning and GPS-based leveling in a local vertical datum, and the spatio-temporal prediction of crustal deformation fields from geodetic data.

Several methods exist in geosciences for tackling problems of spatial signal interpolation and prediction from discrete noisy measurements, and all of them depend to some extent on the way in which some prior information about the primary unknown SRF is assimilated into the data inversion algorithm. The predominant approach that is followed in geodesy for the optimal solution of such problems is least-squares collocation (LSC) which was introduced by Krarup (1969) in a deterministic context as a rigorous approximation method in separable Hilbert spaces with reproducing kernels, and formulated in parallel by Moritz (1970, 1973) in a probabilistic setting as a statistical prediction technique for spatially correlated random variables and stochastic processes; see also Sanso (1980, 1986), Dermanis (1976), Tscherning (1986). Most of the conceivable geodetic measurements, as well as unknown parameters and signals, can be simultaneously handled by the LSC method (Moritz 1980; Dermanis 1980) which has formed the basis of the integrated geodesy concept as introduced in Eeg and Krarup (1973). Similar methods have also been developed in the general area of geostatistics and spatial data analysis, where various kriging-type estimators are widely used for applications of SRF prediction in hydrology, geology, mining engineering, environmental monitoring and applied geophysics (Christakos 1992; Cressie 1993; Matheron 1971); for more details on the similarities and differences between kriging and LSC, see Dermanis (1984), and Reguzzoni et al. (2005).

Although originally developed for optimal estimation problems in spatially varying geodetic signals, the LSC method is closely associated with the pioneering work of Kolmogorov (1941) and Wiener (1949) on the interpolation, extrapolation and smoothing of stationary time series. The formalism of Wiener–Kolmogorov prediction theory in terms of Fourier transforms, spectral filters and input–output (I/O) linear systems has actually been amalgamated in the traditional LSC framework (Sanso and Sideris 1997; Nash and Jordan 1978) and it offers an alternative frequency-domain approach for dealing in a computationally efficient manner with optimal prediction problems in geodetic applications (see e.g., Eren 1982; Schwarz et al. 1990; Sideris 1995).

One of the critical aspects in LSC is the smoothing effect on the predicted signal values $\hat{u}(P)$, which typically exhibit less spatial variability than the actual true field $u(P)$. As a consequence, small field values are overestimated and large values are underestimated, thus introducing a likely conditional bias in the final results and possibly creating artifact structures in SRF maps generated through the LSC estimation process. Note that smoothing is a characteristic which is not solely associated with the LSC method and it is shared by most interpolation techniques aiming at the unique approximation of an unknown continuous function from a finite number of observed functionals. Its merit is that it guarantees that the recovered field does not produce artificial details not inherent or proven by the actual data, which is certainly a reasonable and desirable characteristic for an optimal signal interpolator. However, the use of smoothed SRF images or maps generated by techniques such as LSC or kriging provides a shortfall for applications sensitive to the presence of extreme signal values, patterns of field continuity and spatial correlation structure. While founded on *local* optimality criteria that minimize the mean squared error (MSE) independently for each prediction point, the LSC approach overlooks to some extent a feature of reality that is often important to capture, namely spatial variability. The latter can be considered a *global* attribute of an optimal field estimate, since it only has meaning in the context of the relationship of all predicted values to one another in space. As a result of the smoothing effect, ordinary LSC estimates do not reproduce either the histogram of the underlying true SRF, or the spatial correlation structure as implied by the adopted model of its covariance (CV) function.

If the objective for estimating a SRF from its observed functionals requires spatial variability to be represented in a pragmatic way (at least according to a theoretical or empirical model of the signal CV function) then the results obtained through LSC, or other similar techniques like kriging, may pose limitations for further inference and modeling in Earth-related physical processes, despite their local optimality in terms of minimum mean squared prediction error. It is because of this overly smooth representation of reality that Journel (1990, p. 31) cautions against the actual mapping of SRF prediction results obtained from a linear unbiased estimator with minimum MSE: "*In all rigor, estimates based on a local accuracy criterion such as kriging should only be tabulated; to map them is to entice the user to read on the map patterns of spatial continuity that may be artifacts of the data configuration and the kriging smoothing effect*".

Considering the previous arguments, the objective of this paper is to present a prudent approach that enhances LSC-based field estimates by eliminating their inherent smoothing effect, while preserving most of their local prediction accuracy. Our approach consists of correcting a posteriori the optimal result obtained from the LSC technique for the inversion of Eq. (1), in such a way that the new field estimate matches the spatial correlation structure implied by the signal CV function that was used to construct the initial LSC solution. In contrast to stochastic simulation techniques which provide *multiple* equiprobable signal realizations according to some CV model of spatial variability (Christakos 1992; Deutsch and Journel 1998), the methodology presented herein gives a *unique* field estimate that is statistically consistent with a prior model of its spatial CV function. The uniqueness condition is imposed though an optimal criterion that minimizes the loss in local prediction accuracy (in the MSE sense) which occurs when we transform the LSC solution to match the spatial correlation of the underlying unknown SRF.

The paper is organized as follows: in Sect. 2 a review of the traditional LSC technique for the optimal estimation of an unknown SRF is presented, and an analysis of its inherent smoothing effect is given. The proposed "de-smoothing" approach for LSC-based field estimates is introduced in Sect. 3, and some related numerical tests are presented in Sect. 4 using sets of simulated gravity data. Finally, Sect. 5 concludes the paper by presenting some ideas for future work, along with a few remaining open questions that need further study.

2 Ordinary least-squares collocation

2.1 General concept

Based on the general observation model of Eq. (1), let us briefly review the LSC technique as a means to obtain an optimal SRF estimate $\hat{u}(\cdot)$, at an arbitrary number of prediction points $\{P'_i\}$, from a set of discrete noisy measurements. The symbol $\mathbf{u}$ denotes the random vector that contains the values of the primary SRF $u(\cdot)$ at the selected prediction points, while the vector $\hat{\mathbf{u}}$ contains their corresponding estimates obtained from the LSC method.

Denoting by $s_i = L_i(u)$ the signal part in the available data, the system (1) can be written in vector form as

$$\mathbf{y} = \mathbf{s} + \mathbf{v} \tag{2}$$

where $\mathbf{y}$, $\mathbf{s}$ and $\mathbf{v}$ are n-dimensional random vectors containing the known measurements, and the unknown signal and noise values, respectively, at all observation points $\{P_i\}$. The LSC scheme requires some prior information in the form of auxiliary hypotheses placed on the first and second order moments of the signal and measurement noise. In particular, the additive signal and noise components in (2) are considered uncorrelated with each other (a crucial simplification that is regularly applied in practice), and of known statistical properties in terms of their given expectations and co-variances.

Assuming that the spatial variability of the primary SRF u (e.g., the disturbing potential in most physical geodesy applications) is described by a known CV function

$$C_u(P,Q) = E\{(u(P) - m_u(P))(u(Q) - m_u(Q))\} \tag{3}$$

with $m_u(P) = E\{u(P)\}$ being the spatial trend of u, then the elements of the CV matrix of the signal vector $\mathbf{s}$ are determined through a straightforward application of co-variance propagation (Moritz 1980)

$$\mathbf{C_s}(i,j) = L_i L_j C_u(P_i, P_j) \tag{4}$$

where L_i and L_j correspond to the functionals associated with the ith and jth observation, respectively. In the same way, the cross-CV matrix between the primary field values (at the selected prediction points) and the observed signal values (at the observation points) is obtained as

$$\mathbf{C_{us}}(i,j) = L_j C_u(P'_i, P_j) \tag{5}$$

The CV matrix of the data noise is also considered known, based on the availability of an appropriate stochastic model describing the statistical behavior of the zero-mean measurement errors.

$$\mathbf{C_v}(i,j) = E\{v_i v_j\} = \sigma_{v_i v_j} \tag{6}$$

Within the LSC framework it is not required, in principle, to impose any particular constraints on the functional form of the signal CV function $C_u(P,Q)$, apart from the fact that it corresponds to a positive-definite bivariate kernel. In practice, a stationarity modeling assumption is usually adopted in order to infer the signal and/or the noise covariance structure from available data records and to simplify the overall SRF prediction process.

An additional postulate on the spatial trend $m_u(P)$ of the primary SRF is often employed as an auxiliary hypothesis for the LSC inversion of Eqs. (1) or (2). In fact, various LSC prediction algorithms may arise in practice, depending on how we treat the signal de-trending problem. For the purpose of this paper and without any essential loss of generality, it will be assumed that we deal only with zero-mean SRFs and signals ($E\{u(P)\} = 0, E\{\mathbf{s}\} = \mathbf{0}$). More general cases can be handled either through a remove-restore approach by first subtracting a known deterministic trend from the available data and then restoring it back to the LSC field estimate, or through the simultaneous estimation of the field trend within the LSC algorithm using a suitable parameterization for the quantities $E\{u(P)\}$ and/or $E\{\mathbf{s}\}$; see Dermanis (1990) for more details.

Based on the previous hypotheses, the LSC estimator of the primary SRF at all selected prediction points $\{P'_i\}$ is given by the well known matrix formula

$$\hat{\mathbf{u}} = \mathbf{C_{us}}(\mathbf{C_s} + \mathbf{C_v})^{-1}\mathbf{y} \tag{7}$$

which corresponds to the linear unbiased solution with minimum mean squared prediction error (Moritz 1980; Sanso 1986).

2.2 LSC smoothing effect

The inherent smoothing effect in LSC prediction can be easily identified from the CV structure of its optimal result. Applying co-variance propagation to the SRF estimate $\hat{\mathbf{u}}$ in Eq. (7), we obtain the result

$$\mathbf{C_{\hat{u}}} = \mathbf{C_{us}}(\mathbf{C_s} + \mathbf{C_v})^{-1}\mathbf{C}_{\mathbf{us}}^{\mathrm{T}} \tag{8}$$

which generally differs from the CV matrix $\mathbf{C_u}$ of the original SRF at the particular set of prediction points, i.e.,

$$\mathbf{C_u}(i,j) = C_u(P'_i, P'_j) \neq \mathbf{C_{\hat{u}}}(i,j) \tag{9}$$

Moreover, if we consider the vector of the prediction errors $\mathbf{e} = \hat{\mathbf{u}} - \mathbf{u}$, it holds that

$$\mathbf{C_{\hat{u}}} = \mathbf{C_u} - \mathbf{C_e} \tag{10}$$

where the error CV matrix is given by the equation (Moritz 1980)

$$\mathbf{C_e} = \mathbf{C_u} - \mathbf{C_{us}}(\mathbf{C_s} + \mathbf{C_v})^{-1}\mathbf{C}_{\mathbf{us}}^{\mathrm{T}} \tag{11}$$

The fundamental relationship in Eq. (10) conveys the meaning of the smoothing effect in LSC, which essentially acts as an optimal low-pass filter to the input data. The spatial variability of the LSC prediction errors, in terms of their variances and co-variances, is exactly equal to the deficit in spatial variability of the LSC estimator $\hat{\mathbf{u}}$ with respect to the original SRF $\mathbf{u}$.

Note that the decrease in spatial variability between the estimated and the original true signal, according to Eq. (10), is valid also for the case of noiseless data ($\mathbf{C_v} = \mathbf{0}$). The only exception occurs when the group of prediction points is identical to the group of data points, and the observed signals are of the same type as the primary unknown SRF ($\mathbf{u} = \mathbf{s}, \mathbf{C_{us}} = \mathbf{C_s} = \mathbf{C_u}$). In this special case we have that $\mathbf{C_e} = \mathbf{0}$ and $\mathbf{C_{\hat{u}}} = \mathbf{C_u}$, which manifests the well known data reproduction property of LSC in the presence of noiseless measurements (Moritz 1980).

An example of the smoothing effect that can take place when using the LSC technique for SRF prediction is given in Fig. 1. This particular example refers to a standard interpolation problem, namely the construction of a gravity anomaly grid from irregularly distributed noiseless point data. The image shown in Fig. 1a is the realization of a free-air gravity anomaly field which has been simulated within an $50 \times 50\,\mathrm{km}^2$ area with a uniform sampling resolution of 2 km, according to the Hirvonen model of a planar isotropic CV function (Meier 1981)

$$C_u(P,Q) = \frac{C_o}{1 + \left(\frac{r_{PQ}}{a}\right)^2}$$

where $C_o = 750\,\mathrm{mgal}^2$, r_{PQ} is the planar distance between points P and Q, and the parameter a is selected such that the correlation length of the gravity anomaly field is equal to 8.5 km. The locations of the irregular sampling points for this specific experiment are plotted in Fig. 1b, whereas the resulting LSC-based grid with the corresponding prediction errors are shown in Fig. 1c, d, respectively. The LSC smoothing effect can be further identified in the histograms of the true (simulated) and predicted gravity anomaly grids (Fig. 2), as well as in their corresponding sample statistics that are listed in Table 1.

Remark 1 The smoothing effect according to Eq. (10) inflicts a kind of "model denial" in the optimal LSC solution. That is because the vector estimate $\hat{\mathbf{u}}$ obtained from Eq. (7) refutes its fundamental building component, namely the CV structure of the underlying true SRF. In this respect, the merit of LSC as a stochastic interpolation technique for the optimal recovery of a continuous field from a finite number of observed functionals, *using its known CV function* $C_u(P,Q)$, can be challenged based on the argument that the spatio-statistical structure of the SRF $u(\cdot)$ is not preserved through the optimal estimation process.

Remark 2 The LSC prediction algorithm is not affected by the spatial distribution of the prediction points. If we denote by P'_i and P'_j two different points where the optimal prediction of the primary SRF $u(\cdot)$ is sought, then the LSC method yields the following results, separately

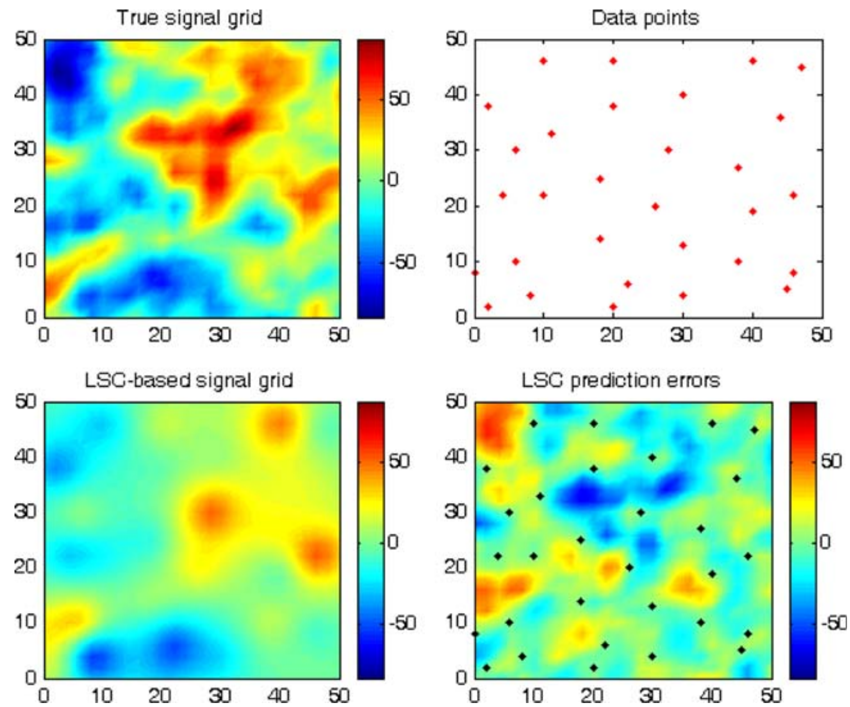


Fig. 1 Plots of the true free-air gravity anomaly signal (*upper left*), the locations of the noiseless irregular point data (*upper right*), the LSC signal solution obtained from the irregular point data (*lower left*), and the actual LSC prediction errors (*lower right*)

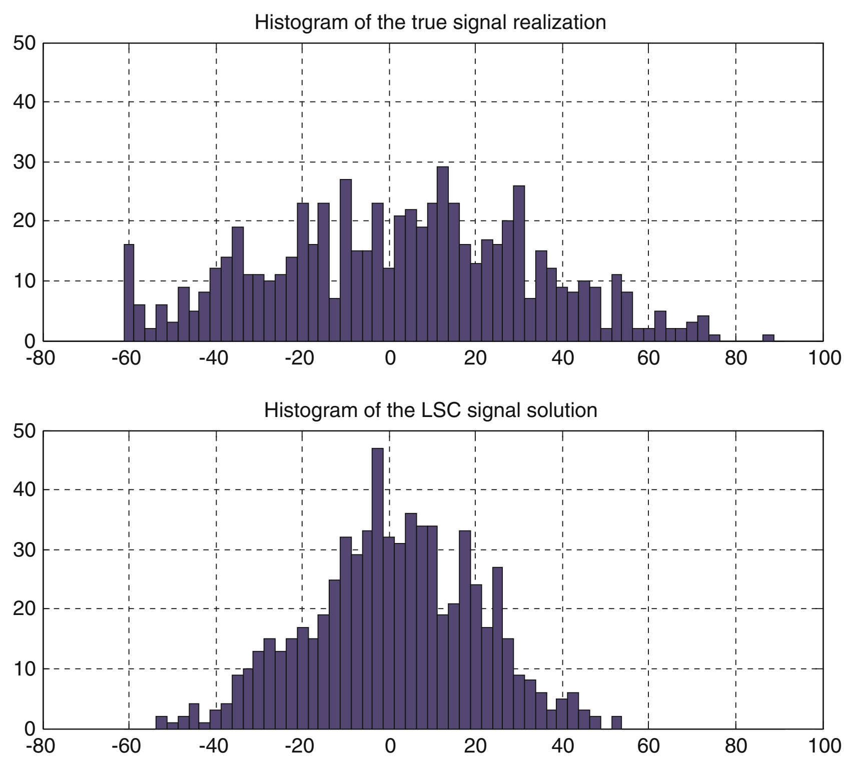


Fig. 2 Histograms of the true free-air gravity anomaly signal (*upper graph*) and the corresponding LSC solution using a sample of noiseless irregular point data (*lower graph*)

Table 1 Statistics of the true gridded values and the predicted gridded values using the LSC method with noiseless irregular point data (all values in mgals)

	Max	Min	Mean	σ
True grid	87.64	−83.70	2.04	31.04
LSC-predicted grid	52.86	−53.21	1.06	19.28

for each point

$$\hat{u}(P_i') = \mathbf{c}_i^{\mathrm{T}} (\mathbf{C_s} + \mathbf{C_v})^{-1} \mathbf{y} \tag{12}$$

and

$$\hat{u}(P_j') = \mathbf{c}_j^{\mathrm{T}} (\mathbf{C_s} + \mathbf{C_v})^{-1} \mathbf{y} \tag{13}$$

The cross-CV vectors $\mathbf{c}_i$ and $\mathbf{c}_j$ are determined through co-variance propagation from the basic CV function $C_u(P, Q)$, i.e.,

$$\mathbf{c}_i = \left[L_1 C_u(P'_i, P_1)\ L_2 C_u(P'_i, P_2)\ \cdots\ L_n C_u(P'_i, P_n)\right]^{\mathrm{T}} \quad (14)$$

$$\mathbf{c}_j = \left[L_1 C_u(P'_j, P_1)\ L_2 C_u(P'_j, P_2)\ \cdots\ L_n C_u(P'_j, P_n)\right]^{\mathrm{T}} \quad (15)$$

The values that are obtained from the prediction formulae (12) and (13) are not directly influenced by the relative spatial position of points P'_i and P'_j. More importantly, these values remain unaffected by the number and/or the geometry of other prediction points where additional field estimates may be computed. Note that the vector formula in Eq. (7) originates essentially by stacking each individual LSC estimate $\hat{u}(\cdot)$ for all desired prediction points.

The LSC results from Eqs. (12) and (13) are both affected by the common data-dependent factor $(\mathbf{C_s} + \mathbf{C_v})^{-1}\mathbf{y}$. Furthermore, the cross-CV vectors in Eqs. (14) and (15) are both generated from the same fundamental CV kernel, thus creating an implicit inter-dependence on the values $\hat{u}(P'_i)$ and $\hat{u}(P'_j)$. *However, what is lacking from the LSC framework is an explicit constraint that would jointly control the numerical values of the field estimates $\hat{u}(\cdot)$ at all prediction points, in a manner that their spatial variation would be statistically consistent with the CV function $C_u(P, Q)$ of the primary SRF.*

As it will be seen in the next section, the formulation of such a constrained SRF prediction approach can lead to an "adaptive" field estimator which, in contrast to the traditional LSC solution $\hat{\mathbf{u}}$, emulates the spatial variation patterns that are inherent into the CV function model $C_u(P, Q)$ for the particular distribution of prediction points $\{P'_i\}$. Depending on the specific problem at hand, the preservation of such patterns in the spatial variation of the primary SRF may be critical for its truthful approximation from discrete noisy data and the modeling or the computation of other quantities of interest that depend on it.

3 An optimal "de-smoothing" scheme for least-squares collocation

The main issue raised in the previous section is the inability of the classic LSC algorithm to reproduce the spatial variability of the primary SRF that needs to be predicted on the basis of a finite set of discrete observations. The term spatial variability is used here to describe the joint statistical behavior of the SRF values, as this is imposed and dictated by an a-priori CV function model. It should be noted that for increasing data sampling resolution, the result of LSC prediction converges to the true field (Tscherning 1978) and thus its inherent smoothing effect ceases to be a problem. However, whether the smoothing effect in LSC is a problem or not in a particular application depends on the relationship between the actual data resolution and the degree of roughness of the primary unknown field, as well as on the data noise level. A data sampling resolution level that may be considered sufficient for one problem, may not be sufficient for another problem with a less smooth field and/or more noisy data. Moreover, the justification for the effectiveness of any estimation or prediction method, including LSC and other least-squares techniques, should not be based solely on its behaviour under some ideal conditions that ensure nice properties of mathematical convergence, but it must also consider what happens in an arbitrary general situation where the resolution of the available data set may not "match" the spatial variability of the primary unknown field that is dictated by its CV function.

Our objective in this paper is to develop a post-processing correction algorithm that can be applied to any optimal field estimate obtained through LSC for the purpose of reducing its inherent smoothing effect, while sustaining most of its local prediction accuracy. In general terms, we seek a "de-smoothing" transformation to act upon the LSC estimator

$$\hat{\mathbf{u}}' = \Re(\hat{\mathbf{u}}) \quad (16)$$

so that the CV structure of the primary SRF $u(P)$ is recovered. This means that the transformation $\Re(\cdot)$ should guarantee that

$$\mathbf{C}_{\hat{\mathbf{u}}'} = \mathbf{C_u} \quad (17)$$

where $\mathbf{C_u}$ is the CV matrix that is formed from the CV function $C_u(P, Q)$ of the unknown SRF; see Eq. (9).

In addition, the prediction errors $\mathbf{e}' = \hat{\mathbf{u}}' - \mathbf{u}$ associated with the field estimator $\hat{\mathbf{u}}'$ should remain small in some sense, so that the new solution can provide not only a CV-adaptive representation for the SRF variation patterns, but also locally accurate predicted values on the basis of the given data. For this purpose, the formulation of the de-smoothing operator $\Re(\cdot)$ should additionally incorporate an optimality principle by minimizing, for example, the trace of the new error CV matrix $\mathbf{C}_{\mathbf{e}'}$.

3.1 Optimal linear de-smoothing

Let us introduce a straightforward linear approach to modify the LSC estimator $\hat{\mathbf{u}}$ by setting

$$\hat{\mathbf{u}}' = \mathbf{R}\hat{\mathbf{u}} \quad (18)$$

where $\mathbf{R}$ is a square filtering matrix that needs to be determined according to some optimal properties for the new estimator $\hat{\mathbf{u}}'$.

The field estimate obtained from Eq. (18) should reproduce the CV structure of the primary SRF, in the sense that $\mathbf{C}_{\hat{\mathbf{u}}'} = \mathbf{C}_{\mathbf{u}}$ for the given spatial distribution of all prediction points $\{P_i'\}$. As a result, the filtering matrix $\mathbf{R}$ should satisfy the CV-matching constraint

$$\mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{R}^{\mathrm{T}} = \mathbf{C}_{\mathbf{u}} \tag{19}$$

where $\mathbf{C}_{\mathbf{u}}$ and $\mathbf{C}_{\hat{\mathbf{u}}}$ correspond to the known CV matrices of the true and the LSC-predicted SRFs, respectively.

The assessment of the prediction accuracy of the new solution $\hat{\mathbf{u}}'$ can be made through its error CV matrix

$$\mathbf{C}_{\mathbf{e}'} = E\{(\hat{\mathbf{u}}' - \mathbf{u})(\hat{\mathbf{u}}' - \mathbf{u})^{\mathrm{T}}\} \tag{20}$$

which, taking into account Eq. (18), yields.

$$\mathbf{C}_{\mathbf{e}'} = \mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{R}^{\mathrm{T}} + \mathbf{C}_{\mathbf{u}} - \mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}\mathbf{u}} - \mathbf{C}_{\mathbf{u}\hat{\mathbf{u}}}\mathbf{R}^{\mathrm{T}} \tag{21}$$

Using the following relations that are always valid for the LSC estimator (assuming that there is zero correlation between the observed signals $\mathbf{s}$ and the measurement noise $\mathbf{v}$)

$$\mathbf{C}_{\mathbf{u}} = \mathbf{C}_{\hat{\mathbf{u}}} + \mathbf{C}_{\mathbf{e}} \tag{22}$$

$$\mathbf{C}_{\hat{\mathbf{u}}\mathbf{u}} = \mathbf{C}_{\hat{\mathbf{u}}} \tag{23}$$

the new error CV matrix $\mathbf{C}_{\mathbf{e}'}$ can be finally expressed as

$$\mathbf{C}_{\mathbf{e}'} = \mathbf{C}_{\mathbf{e}} + (\mathbf{I} - \mathbf{R})\mathbf{C}_{\hat{\mathbf{u}}}(\mathbf{I} - \mathbf{R})^{\mathrm{T}} \tag{24}$$

where $\mathbf{C}_{\mathbf{e}}$ is the error CV matrix of the usual LSC solution.

Evidently, the pointwise prediction accuracy of the modified solution $\hat{\mathbf{u}}'$ will always be worse than the prediction accuracy of the original LSC solution $\hat{\mathbf{u}}$, regardless of the form of the matrix $\mathbf{R}$. This is expected since LSC provides the best (in the MSE sense) linear predictor from the available measurements, which cannot be further improved by additional linear, or even nonlinear in the case of normally distributed SRFs, operations.

Our aim here is to determine an optimal filtering matrix $\mathbf{R}$ that will satisfy the CV-matching constraint in (19), while minimizing the loss of the MSE prediction accuracy in the sense that

$$\mathrm{trace}(\mathbf{C}_{\mathbf{e}'} - \mathbf{C}_{\mathbf{e}}) = \mathrm{trace}(\delta\mathbf{C}_{\mathbf{e}'}) = \text{minimum} \tag{25}$$

According to Eq. (24), the residual matrix

$$\delta\mathbf{C}_{\mathbf{e}'} = (\mathbf{I} - \mathbf{R})\mathbf{C}_{\hat{\mathbf{u}}}(\mathbf{I} - \mathbf{R})^{\mathrm{T}} \tag{26}$$

represents the part of the error CV matrix of the new estimator $\hat{\mathbf{u}}'$ which depends on the choice of the filtering matrix $\mathbf{R}$.

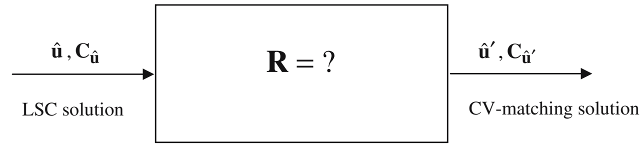


Fig. 3 De-smoothing of the LSC estimator $\hat{\mathbf{u}}$ as a filtering operation in a SISO linear system

Remark 3 To obtain an idea about the degradation in the MSE prediction accuracy caused by the modified estimator $\hat{\mathbf{u}}'$, *under the presence of the CV-matching constraint* (19), let us consider the case where the filtering matrix takes the general form $\mathbf{R} = \mathbf{I} + \delta\mathbf{R}$, with $\delta\mathbf{R}$ being a non negative-definite matrix (i.e., all eigenvalues of $\mathbf{R}$ are larger, or equal, than one). Then, it follows from Eq. (24) that

$$\begin{aligned}\mathrm{trace}(\mathbf{C}_{\mathbf{e}'}) &= \mathrm{trace}(\mathbf{C}_{\mathbf{e}} + \mathbf{C}_{\hat{\mathbf{u}}} + \mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{R}^{\mathrm{T}} - \mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}} - \mathbf{C}_{\hat{\mathbf{u}}}\mathbf{R}^{\mathrm{T}})\\ &= \mathrm{trace}(\mathbf{C}_{\mathbf{e}}) + \mathrm{trace}(\mathbf{C}_{\hat{\mathbf{u}}}) + \mathrm{trace}(\mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{R}^{\mathrm{T}})\\ &\quad -\mathrm{trace}(\mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}}) - \mathrm{trace}(\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{R}^{\mathrm{T}})\end{aligned} \tag{27a}$$

Taking into account the matrix decomposition $\mathbf{R} = \mathbf{I} + \delta\mathbf{R}$, the CV-matching constraint (19), and the well known properties $\mathrm{trace}(\mathbf{AB}) = \mathrm{trace}(\mathbf{BA}) = \mathrm{trace}((\mathbf{AB})^{\mathrm{T}})$, we have

$$\begin{aligned}\mathrm{trace}(\mathbf{C}_{\mathbf{e}'}) &= \mathrm{trace}(\mathbf{C}_{\mathbf{e}}) + \mathrm{trace}(\mathbf{C}_{\hat{\mathbf{u}}}) + \mathrm{trace}(\mathbf{C}_{\mathbf{u}})\\ &\quad -2\mathrm{trace}(\mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}})\\ &= \mathrm{trace}(\mathbf{C}_{\mathbf{e}}) + \mathrm{trace}(\mathbf{C}_{\hat{\mathbf{u}}}) + \mathrm{trace}(\mathbf{C}_{\mathbf{u}})\\ &\quad -2\mathrm{trace}(\mathbf{C}_{\hat{\mathbf{u}}}) - 2\mathrm{trace}(\delta\mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}})\\ &= \mathrm{trace}(\mathbf{C}_{\mathbf{e}}) + \mathrm{trace}(\mathbf{C}_{\mathbf{u}}) - \mathrm{trace}(\mathbf{C}_{\hat{\mathbf{u}}})\\ &\quad -2\mathrm{trace}(\delta\mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}})\\ &= \mathrm{trace}(\mathbf{C}_{\mathbf{e}}) + \mathrm{trace}(\mathbf{C}_{\hat{\mathbf{u}}}) + \mathrm{trace}(\mathbf{C}_{\mathbf{e}})\\ &\quad -\mathrm{trace}(\mathbf{C}_{\hat{\mathbf{u}}}) - 2\mathrm{trace}(\delta\mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}})\\ &= 2\mathrm{trace}(\mathbf{C}_{\mathbf{e}}) - 2\mathrm{trace}(\delta\mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}}).\end{aligned} \tag{27b}$$

Since both $\delta\mathbf{R}$ and $\mathbf{C}_{\hat{\mathbf{u}}}$ are non negative-definite matrices, then it holds that $\mathrm{trace}(\delta\mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}}) \geq 0$ (Theobald 1974, p. 104). Thus, we infer that

$$\mathrm{trace}(\mathbf{C}_{\mathbf{e}'}) \leq 2\mathrm{trace}(\mathbf{C}_{\mathbf{e}}) \tag{28}$$

Note that the above inequality is true for every filtering matrix that satisfies the CV-matching constraint (19) and has the form $\mathbf{R} = \mathbf{I} + \delta\mathbf{R}$, with $\delta\mathbf{R}$ being a non negative-definite matrix.

3.2 An alternative SISO-type formulation

The optimal de-smoothing process that was presented in the previous section can be formulated in a different, yet equivalent, way as follows. Using a single-input single-output (SISO) linear system approach, we take the LSC

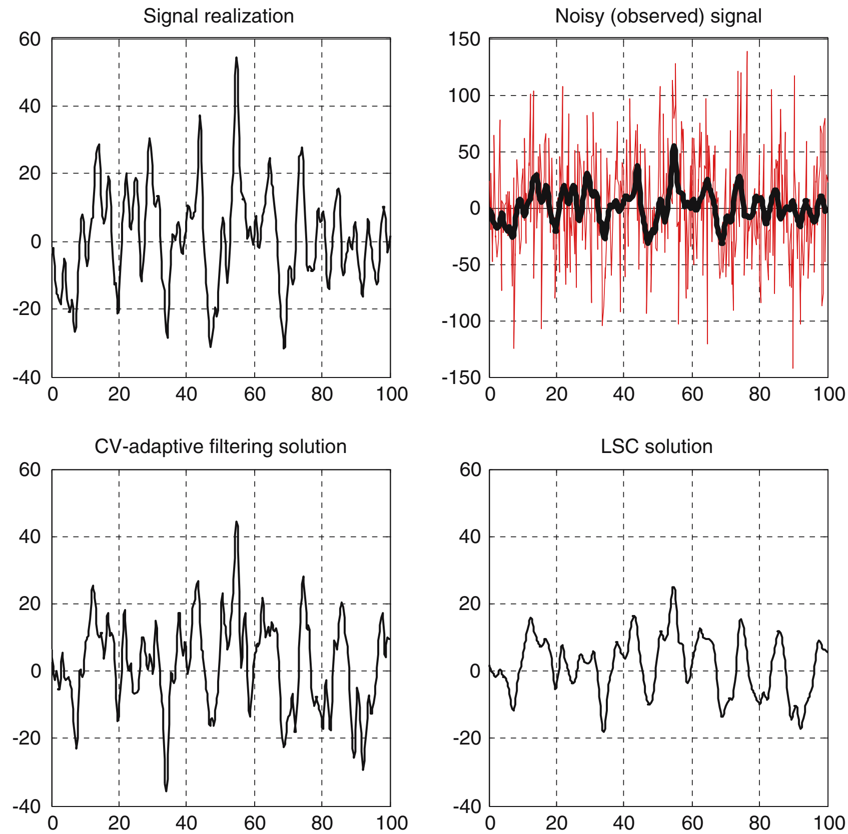


Fig. 4 Plots of the true gravity anomaly signal (*upper left*), the noisy observed signal (*upper right*), the CV-matching signal solution (*lower left*), and the LSC signal solution (*lower right*)

estimate $\hat{\mathbf{u}}$ as the known input to a SISO system which is implemented in terms of an arbitrary square filtering matrix $\mathbf{R}$ (see Fig. 3). The output of the SISO system corresponds to a new estimate $\hat{\mathbf{u}}'$ of the primary SRF, whose properties are controlled by the system matrix.

We want to impose a specific CV structure on the output field $\hat{\mathbf{u}}'$, which must reproduce the spatial variation patterns of the underlying primary SRF at all prediction points $\{P_i'\}$. Hence, the filtering matrix of the SISO system must satisfy the same CV-matching constraint that was already given in Eq. (19).

In contrast to the optimal principle in (25) that minimizes the loss of the MSE prediction accuracy for the new estimate $\hat{\mathbf{u}}'$ with respect to the original LSC solution, we shall now introduce a different criterion for the unique determination of the filtering matrix $\mathbf{R}$. In particular, we require that the differences between the original LSC solution and the new CV-adaptive field estimate are as small as possible. This requirement is formulated in terms of the following minimization principle which enforces an optimal fit, in the mean squared sense, between the input and the output signals in the SISO system of Fig. 3

$$\left\|\hat{\mathbf{u}}' - \hat{\mathbf{u}}\right\|^2 = E\{(\hat{\mathbf{u}}' - \hat{\mathbf{u}})^{\mathrm{T}}(\hat{\mathbf{u}}' - \hat{\mathbf{u}})\} = \text{minimum}. \tag{29}$$

The above principle is algebraically equivalent to the one given in Eq. (25). Indeed, if we use the following property which holds for every zero-mean random vector $\mathbf{x}$ and symmetric matrix $\mathbf{Q}$ (Koch 1999, p. 134)

$$E\{\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x}\} = \operatorname{trace}(\mathbf{Q}\mathbf{C}_{\mathbf{x}}).$$

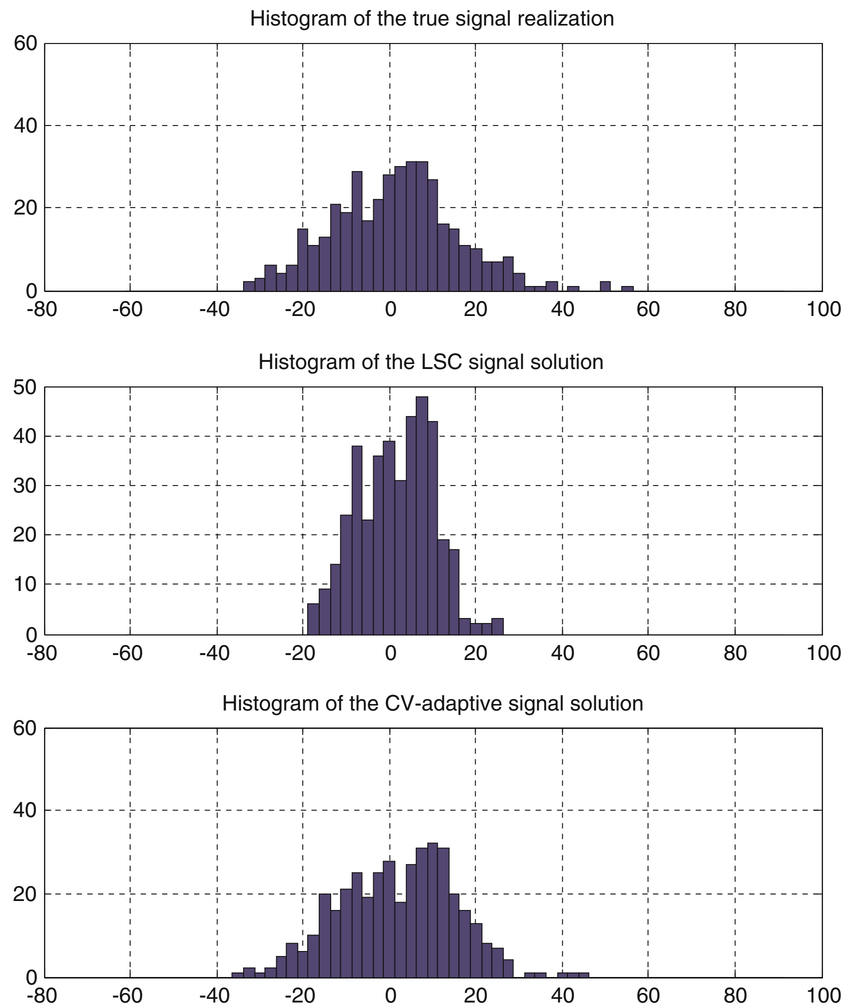
we have that

$$\begin{aligned}
E\{(\hat{\mathbf{u}}' - \hat{\mathbf{u}})^{\mathrm{T}}(\hat{\mathbf{u}}' - \hat{\mathbf{u}})\} &= E\{(\mathbf{R}\hat{\mathbf{u}} - \hat{\mathbf{u}})^{\mathrm{T}}(\mathbf{R}\hat{\mathbf{u}} - \hat{\mathbf{u}})\} \\
&= E\{(\hat{\mathbf{u}}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}} - \hat{\mathbf{u}}^{\mathrm{T}})(\mathbf{R}\hat{\mathbf{u}} - \hat{\mathbf{u}})\} \\
&= E\{\hat{\mathbf{u}}^{\mathrm{T}}(\mathbf{R}^{\mathrm{T}} - \mathbf{I})(\mathbf{R} - \mathbf{I})\hat{\mathbf{u}}\} \\
&= E\{\hat{\mathbf{u}}^{\mathrm{T}}(\mathbf{I} - \mathbf{R})^{\mathrm{T}}(\mathbf{I} - \mathbf{R})\hat{\mathbf{u}}\} \\
&= \operatorname{trace}((\mathbf{I} - \mathbf{R})^{\mathrm{T}}(\mathbf{I} - \mathbf{R})\mathbf{C}_{\hat{\mathbf{u}}}\} \\
&= \operatorname{trace}((\mathbf{I} - \mathbf{R})\mathbf{C}_{\hat{\mathbf{u}}}(\mathbf{I} - \mathbf{R})^{\mathrm{T}}\} \\
&= \operatorname{trace}(\mathbf{C}_{\mathbf{e}'} - \mathbf{C}_{\mathbf{e}}) \\
&= \operatorname{trace}(\delta\mathbf{C}_{\mathbf{e}'}).
\end{aligned} \tag{30}$$

Consequently, the desired filtering matrix $\mathbf{R}$ for implementing our CV-adaptive prediction scheme according to (18) and (19) yields two basic properties for the field solution $\hat{\mathbf{u}}'$, namely:

(i) minimum loss in the MSE prediction accuracy with respect to the original LSC solution, and
(ii) maximum statistical agreement, in the mean squared sense, between the CV-adaptive solution and the LSC solution.

Fig. 5 Histograms of the true gravity anomaly signal (*upper graph*), the LSC signal solution (*middle graph*), and the CV-matching signal solution (*lower graph*)



3.3 The optimal filtering matrix

The solution to the optimization problem that was formulated in the last two sections, namely the determination of the filtering matrix $\mathbf{R}$ that satisfies the CV-matching constraint (19) and also minimizes the loss in the prediction accuracy for the linear estimator $\hat{\mathbf{u}}'$, is analytically given in the appendix. Note that the following result was originally derived in Eldar (2001, 2003) in a completely different context than the one discussed in this paper, focusing on applications such as matched-filter detection, quantum signal processing and subspace signal whitening; see also Eldar and Oppenheim (2003).

The optimal filtering matrix will be given by the equation

$$\mathbf{R} = \mathbf{C}_{\mathbf{u}}^{1/2}(\mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2})^{-1/2}\mathbf{C}_{\mathbf{u}}^{1/2} \tag{31}$$

or equivalently,

$$\mathbf{R} = \mathbf{C}_{\hat{\mathbf{u}}}^{-1/2}(\mathbf{C}_{\hat{\mathbf{u}}}^{1/2}\mathbf{C}_{\mathbf{u}}\mathbf{C}_{\hat{\mathbf{u}}}^{1/2})^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}^{-1/2} \tag{32}$$

Table 2 Statistics of the true gravity anomaly signal, the LSC signal solution and the CV-matching signal solution (all values in mgals)

	Max	Min	Mean	σ
True grid	54.18	−31.59	1.26	14.62
LSC filtered signal	25.14	−18.00	1.60	8.69
CV-matching filtered signal	44.20	−35.49	1.98	13.46

Using the matrix identity $\mathbf{S}\mathbf{T}^{1/2}\mathbf{S}^{-1} = (\mathbf{S}\mathbf{T}\mathbf{S}^{-1})^{1/2}$, we can also obtain the following equivalent matrix expressions from the previous equations:

$$\mathbf{R} = (\mathbf{C}_{\mathbf{u}}\mathbf{C}_{\hat{\mathbf{u}}})^{-1/2}\mathbf{C}_{\mathbf{u}} \tag{33}$$

$$\mathbf{R} = (\mathbf{C}_{\mathbf{u}}\mathbf{C}_{\hat{\mathbf{u}}})^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}^{-1} \tag{34}$$

$$\mathbf{R} = \mathbf{C}_{\hat{\mathbf{u}}}^{-1}(\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}})^{1/2} \tag{35}$$

$$\mathbf{R} = \mathbf{C}_{\mathbf{u}}(\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}})^{-1/2} \tag{36}$$

The numerical implementation of any of the last four formulae requires the use of an appropriate algorithm

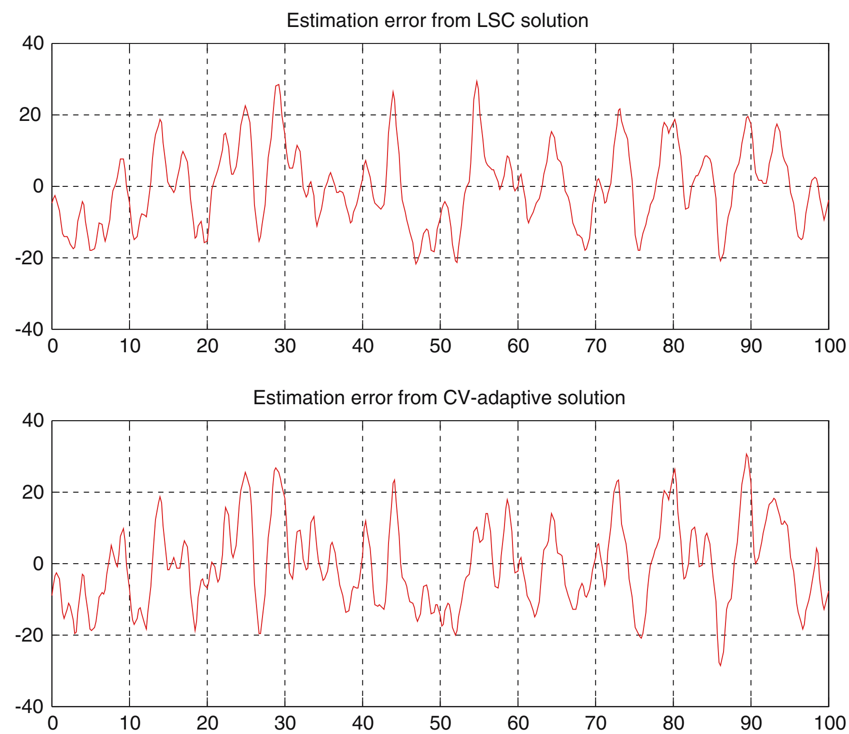


Fig. 6 Plots of the actual prediction errors in the LSC signal solution (*upper graph*), and the CV-matching signal solution (*lower graph*)

Table 3 Statistics of the actual predictions errors in the LSC signal solution and the CV-matching signal solution (all values in mgals)

	Max	Min	Mean	σ
LSC prediction errors	29.03	−21.89	−0.33	11.03
CV-matching prediction errors	30.24	−28.91	−0.71	11.83

for the computation of the square root of a generally non-symmetric matrix; see, e.g., Higham (1987, 1997, 2006), Iannazzo (2006), and Smith (2003).

Following the general expression given in (24), the error CV matrix of the CV-matching field estimator $\hat{\mathbf{u}}'$ that corresponds to the previous equivalent optimal choices for the filtering matrix $\mathbf{R}$ will be given by the formula

$$\mathbf{C_{e'}} = \mathbf{C_e} + \mathbf{C_{\hat{u}}} + \mathbf{C_u} - (\mathbf{C_{\hat{u}}C_u})^{1/2} - (\mathbf{C_u C_{\hat{u}}})^{1/2} \tag{37}$$

which can be used for the assessment of the prediction error originating from the de-smoothing process $\hat{\mathbf{u}}' = \mathbf{R}\hat{\mathbf{u}}$ that is applied to the standard LSC solution.

4 Numerical tests

A number of simulation experiments have been worked out to demonstrate the performance of the CV-adaptive field estimator $\hat{\mathbf{u}}'$ in contrast to the classic LSC estimator $\hat{\mathbf{u}}$. The examples presented in this section refer to a standard SRF prediction problem that often appears in many geodetic applications, namely the optimal denoising of observed gravity field functionals.

The first simulation regards a signal $u(x)$, with $x \in [0, 100]$, thought of as a free-air gravity anomaly profile obtained from airborne gravity measurements. In particular, we have simulated a 4 Hz ensemble of a zero-mean SRF $u(x)$ based on the CV function model

$$C_u(P,Q) = \frac{C_o}{1 + \left(\frac{r_{PQ}}{a}\right)^2} \tag{38}$$

where $C_o = 225\,\text{mgal}^2, r_{PQ}$ is the distance between points P and Q, and the parameter a is selected such that the correlation length of the gravity anomaly profile is 10 km. The "observed" data record is obtained by adding white noise to the simulated signal values, with the noise variance set to $1{,}025\,\text{mgal}^2$.

Two filtering solutions for $u(x)$ have been determined based on the aforementioned simulation setting. The first field estimate comes from the usual LSC algorithm that is applied directly to the available data (a standard *Wiener filter* in this particular case), while the second field estimate corresponds to the CV-matching optimal solution that was described in the previous sections. The two signal solutions, along with the original SRF and the observed noisy data, are illustrated in Fig. 4. As it can be seen from these plots, the CV-matching solution emulates much closer than the LSC solution the

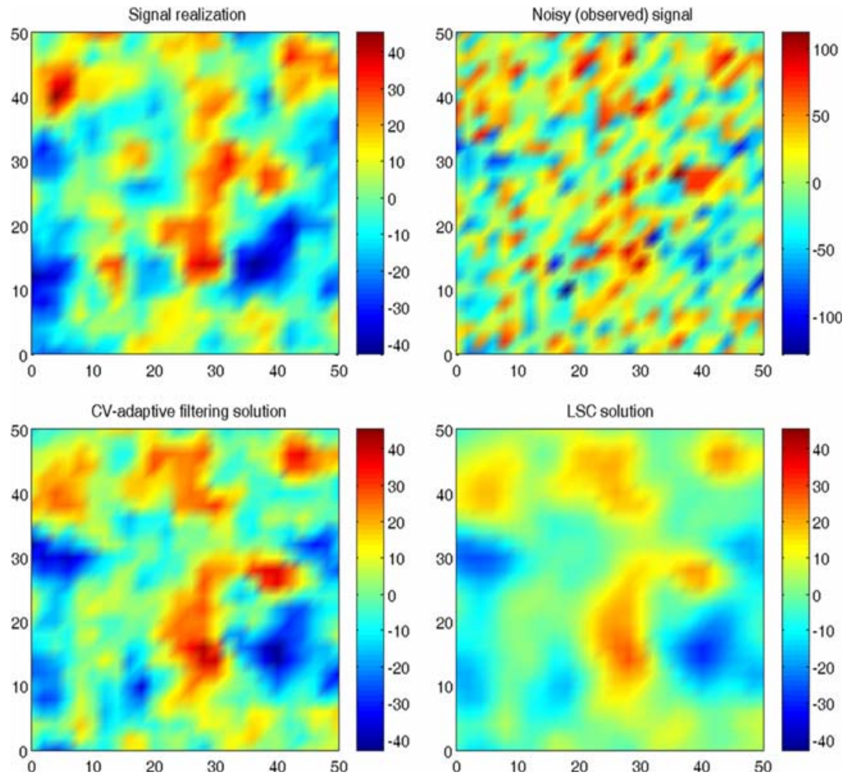


Fig. 7 Plots of the true gravity anomaly signal (*upper left*), the noisy observed signal (*upper right*), the CV-matching signal solution (*lower left*), and the LSC signal solution (*lower right*)

spatial variability and the range of values of the original SRF $u(x)$, despite LSC's optimal pointwise prediction accuracy.

The improved replication of the spatial variability of the unknown SRF by the CV-matching solution $\hat{\mathbf{u}}'$, in contrast to the smoothed representation obtained by the LSC result $\hat{\mathbf{u}}$, can also be seen in the histogram plots in Fig. 5, as well as in the signal statistics in Table 2.

The actual prediction errors in the LSC solution and the CV-matching solution (i.e., $\mathbf{e} = \hat{\mathbf{u}} - \mathbf{u}$ and $\mathbf{e}' = \hat{\mathbf{u}}' - \mathbf{u}$, respectively), are plotted in Fig. 6, while their corresponding statistics are given in Table 3. From the spatial behavior and the magnitude of their prediction errors, we conclude that the performance of both solutions is similar in terms of average prediction accuracy, with the LSC result being marginally better than the CV-matching field estimate, as should be expected.

The second example is essentially a repetition of the previous noise filtering experiment, with the difference being that now a two-dimensional gravity anomaly SRF is considered. In particular, using the same signal and noise simulation parameters and the same CV function model of Eq. (38), a two-dimensional free-air gravity anomaly grid of "true" and "observed" values is simulated over a $50 \times 50\,\mathrm{km}^2$ with a uniform resolution of 1 km (see Fig. 7). The LSC and the CV-matching filtering solutions obtained from this test are also shown in Fig. 7, while their corresponding histograms are presented in Fig. 8, and their actual prediction errors in Fig. 9.

Similarly to the previous example, we see again the improvement achieved by the CV-matching solution in the representation of the spatial variability patterns and the range of signal values, at the cost of a small increase in the signal prediction errors; see also Tables 4 and 5.

As a final note, let us mention that the improvement obtained from the CV-matching prediction algorithm over the LSC solution, with respect to the representation of spatial signal variability and the reproduction of the global statistics of the underlying SRF, becomes more evident in cases with particularly low signal-to-noise ratio (SNR), while for high SNR situations both field estimates give practically identical results. A characteristic example is shown in Figs. 10 and 11 for two noise filtering cases with high and low SNR values, respectively. Specifically, in Fig. 10 we see the results obtained from the CV-matching and the LSC prediction methods when the ratio between the signal and the noise standard deviation in the simulation of a free-air gravity anomaly field is set to SNR = 1.5, while in Fig. 11 the corresponding results are shown when the simulation experiment takes place with SNR = 0.1. From the values of the statistics given in the corresponding Tables 6, 7, 8 and 9, we confirm the superiority of the CV-matching solution when the SNR of the problem is low, and the equivalency with the LSC solution when the SNR is high.

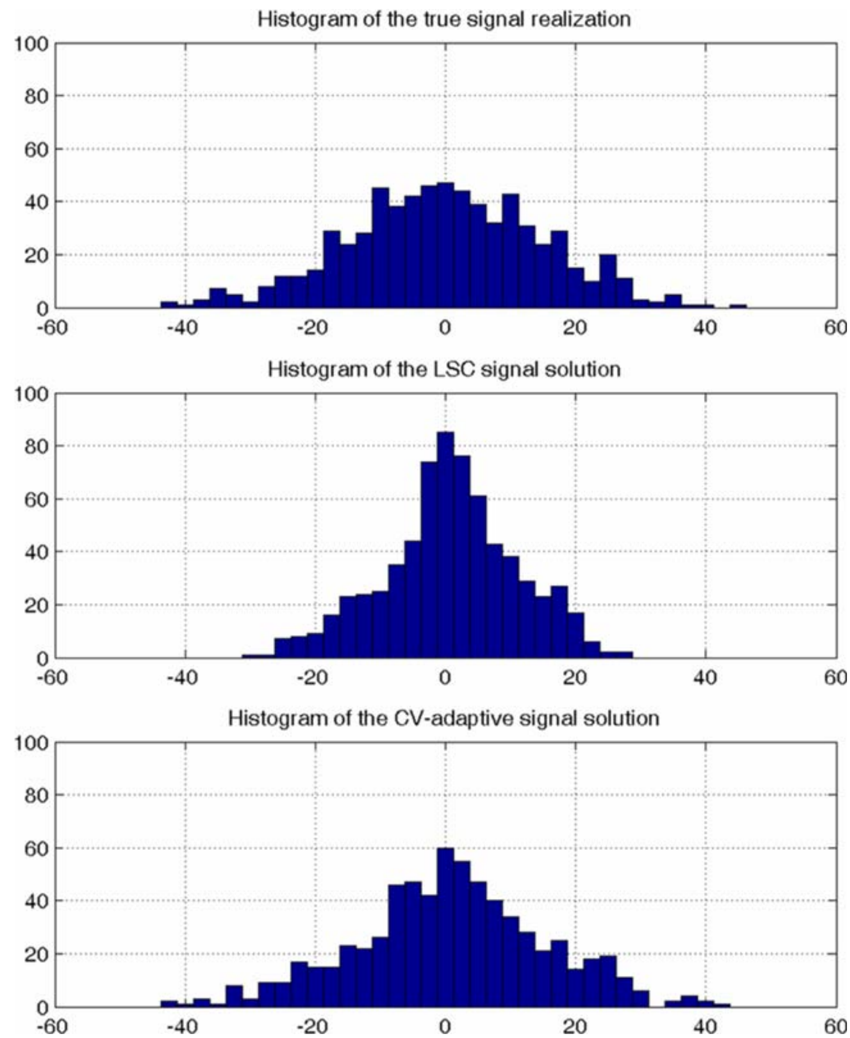


Fig. 8 Histograms of the true gravity anomaly signal (*upper graph*), the LSC signal solution (*middle graph*), and the CV-matching signal solution (*lower graph*). The plots refer to the 2D noise filtering experiment

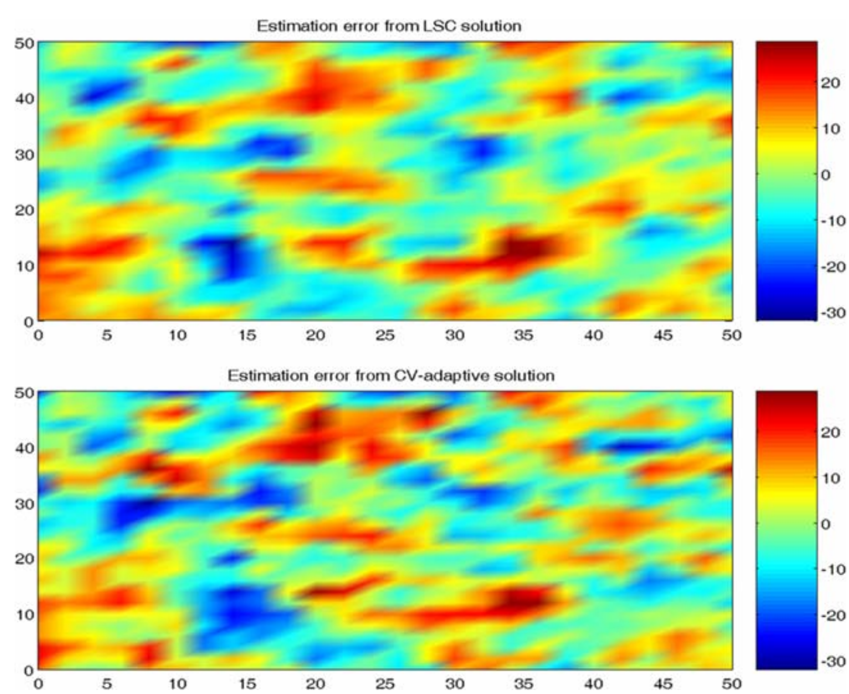


Fig. 9 Plots of the actual prediction errors in the LSC signal solution (*upper graph*), and the CV-matching signal solution (*lower graph*). The plots refer to the 2D noise filtering experiment

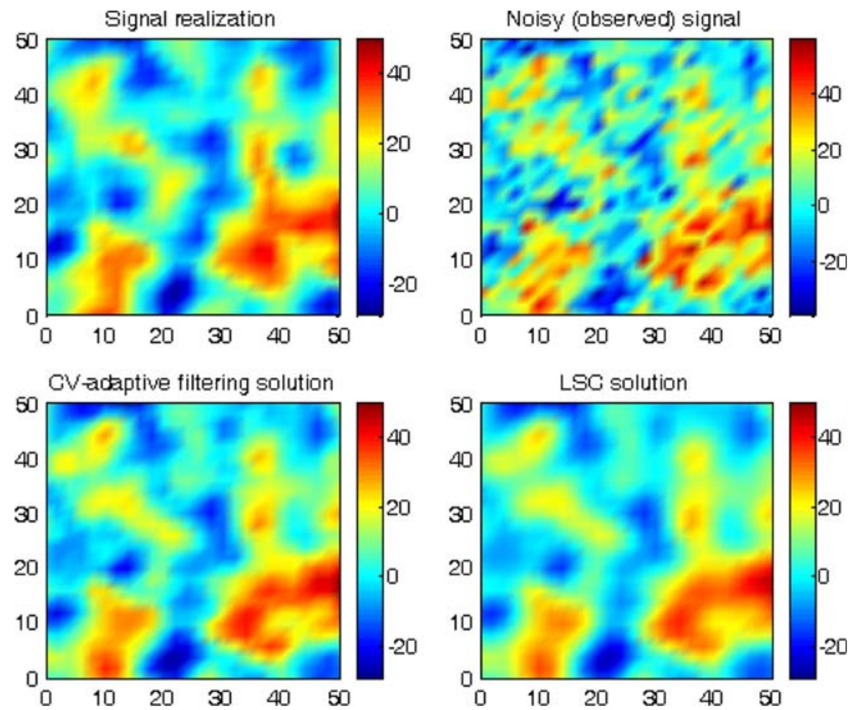


Fig. 10 Plots of the true gravity anomaly signal (*upper left*), the noisy observed signal (*upper right*), the CV-matching signal solution (*lower left*), and the LSC signal solution (*lower right*), for the case where SNR = 1.5

Table 4 Statistics of the true gravity anomaly signal, the LSC signal solution and the CV-matching signal solution for the two-dimensional noise filtering experiment (all values in mgals)

	Max	Min	Mean	σ
True grid	45.33	−42.88	−0.04	14.96
LSC filtered signal	27.42	−29.10	0.74	10.29
CV-matching filtered signal	41.47	−42.48	0.73	14.64

Table 5 Statistics of the actual predictions errors in the LSC signal solution and the CV-matching signal solution for the two-dimensional noise filtering experiment (all values in mgals)

	Max	Min	Mean	σ
LSC prediction errors	28.70	−31.94	0.78	9.88
CV-matching prediction errors	31.98	−32.05	0.77	11.01

5 Conclusions

Most geostatistical methods for SRF prediction problems using discrete data, including LSC and the various forms of kriging, rely on the use of prior models that describe the spatial correlation of the unknown field at hand over its domain. Built upon an optimal criterion of maximum local accuracy, LSC provides an unbiased field estimate that has the smallest mean squared prediction error, at every computation point, among any other linear prediction method using the same data. However, LSC field estimates do not reproduce the spatial variability of the corresponding unknown SRFs which is implied by their adopted CV functions. From a theoretical viewpoint, this can be considered as a critical drawback in the sense that the spatial correlation structure of the SRF (e.g., the disturbing potential in the case of gravity field modeling) is not preserved during its optimal estimation process. On the practical side, the LSC algorithm acts as a low-pass filter which smoothens the spatial variability patterns originating from the adopted, either theoretical or empirical, CV function model of the unknown field.

Although smoothing is not necessarily a bad side-effect in signal interpolation and prediction problems, there is an apparent paradox in the idea that the optimality criteria used in the LSC method give field estimates that do not achieve the global statistics of the underlying SRF, thus suggesting that "local" prediction accuracy (in the sense of unbiased signal estimates with minimum pointwise MSE) and "global" field mapping (in the sense of preserving the spatial correlation structure of the unknown signal according to a CV function model) cannot be both optimally accomplished. Generally speaking, this dichotomy may be a LSC user's analogue to the well known Heisenberg's uncertainty principle. Whether or not this paradox represents something more fundamental, the current needs for SRF prediction in geodesy and geosciences make it clear that the objective(s) in every application must be clearly defined in advance (e.g., maximum local accuracy, preservation of global statistics, preservation of non-stationary patterns in spatial variability, etc.) since no single method

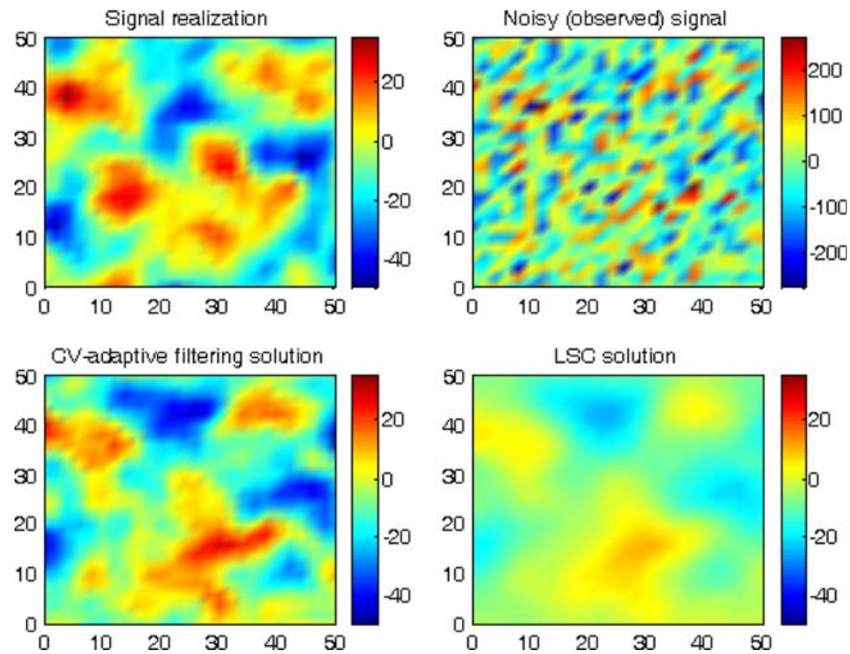


Fig. 11 Plots of the true gravity anomaly signal (*upper left*), the noisy observed signal (*upper right*), the CV-matching signal solution (*lower left*), and the LSC signal solution (*lower right*), for the case where SNR = 0.1

Table 6 Statistics of the true gravity anomaly signal, the LSC signal solution and the CV-matching signal solution (all values in mgals), for the case where SNR = 1.5

	Max	Min	Mean	σ
True grid	49.68	−28.82	7.16	14.29
LSC filtered signal	46.91	−25.78	7.14	13.32
CV-matching filtered signal	50.92	−27.15	7.16	14.11

Table 7 Statistics of the true gravity anomaly signal, the LSC signal solution and the CV-matching signal solution (all values in mgals), for the case where SNR = 0.1

	Max	Min	Mean	σ
True grid	34.62	−49.37	−6.32	14.55
LSC filtered signal	9.82	−24.72	−6.22	7.01
CV-matching filtered signal	28.03	−47.45	−7.76	13.94

Table 8 Statistics of the actual predictions errors in the LSC signal solution and the CV-matching signal solution (all values in mgals) for the case where SNR = 1.5

	Max	Min	Mean	σ
LSC prediction errors	13.47	−16.35	−0.02	4.70
CV-matching prediction errors	13.34	−17.65	−0.00	5.04

exists that allows the simultaneous optimal reaching of all possible objectives in a given prediction problem. Choosing the best solution based solely on a minimum MSE criterion may be considered a prejudiced approach that does not necessarily lead to the most realistic picture of physical reality.

Table 9 Statistics of the actual predictions errors in the LSC signal solution and the CV-matching signal solution (all values in mgals), fir the case where SNR = 0.1

	Max	Min	Mean	σ
LSC prediction errors	29.54	−32.11	0.10	12.04
CV-matching prediction errors	31.90	−49.92	−1.44	13.79

The SRF prediction method presented in this paper is not as locally accurate as LSC, but is more globally representative for an unknown signal with a known CV function. Our methodology employs a post-processing filter that is applied to the usual LSC solution and yields a unique field estimate that reproduces the global statistics of the primary SRF based on a CV-matching constraint. Besides having the same spatial CV structure with the primary SRF, the new field estimate retains most of the local prediction accuracy associated with the original LSC solution by minimizing the increase of the error variance at each prediction point. In that sense, the proposed technique offers an interesting alternative to LSC for SRF filtering problems in cases where the final result should follow a given CV function model. It can provide a useful post-processing tool, in particular, for physical geodesy estimation problems with low SNR and/or high degree of smoothness in their regularized least-squares solution (e.g., Wiener filtering of airborne gravity data, downward continuation, etc.), for calibration or validation of signal and/or error CV models using discrete data, for computing stochastic simulations (according to some a-priori CV model) of various gravity field functionals that need to be additionally conditioned upon actual data values, and finally for adjusting heterogeneous gravity field data sets according to a common CV model of spatial variability.

Acknowledgements The author would like to acknowledge the valuable suggestions and comments by C.C. Tscherning and D. Roman, which led to an improved revised version of this paper.

Appendix: Determination of the optimal filtering matrix

We want to determine the optimal filtering matrix $\mathbf{R}$ that is used for obtaining a new SRF estimate $\hat{\mathbf{u}}'$ from the original LSC solution $\hat{\mathbf{u}}$ according to $\hat{\mathbf{u}}' = \mathbf{R}\hat{\mathbf{u}}$, which satisfies the CV-matching constraint

$$\mathbf{C}_{\hat{\mathbf{u}}'} = \mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{R}^{\mathrm{T}} = \mathbf{C}_{\mathbf{u}} \tag{A1}$$

and also minimizes the loss in MSE prediction accuracy as stated by the optimal principle (see Sect. 3.1)

$$\varphi = \mathrm{trace}[(\mathbf{I}-\mathbf{R})\mathbf{C}_{\hat{\mathbf{u}}}(\mathbf{I}-\mathbf{R})^{\mathrm{T}}] = \mathrm{minimum} \tag{A2}$$

where $\mathbf{C}_{\mathbf{u}}$ and $\mathbf{C}_{\hat{\mathbf{u}}}$ are known $n \times n$ positive-definite matrices, with $\mathbf{C}_{\hat{\mathbf{u}}} \neq \mathbf{C}_{\mathbf{u}}$ in general. Note that the MSE principle in the last equation is equivalent to the following expression

$$\varphi = E\{(\hat{\mathbf{u}}' - \hat{\mathbf{u}})^{\mathrm{T}}(\hat{\mathbf{u}}' - \hat{\mathbf{u}})\} = \mathrm{minimum} \tag{A3}$$

as explained in Sect. 3.2.

Let us apply an auxiliary invertible transformation to the LSC field estimate $\hat{\mathbf{u}}$ and the corresponding CV-matching field estimate $\hat{\mathbf{u}}' = \mathbf{R}\hat{\mathbf{u}}$ (note that there is a different transformation matrix applied to each of the vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{u}}'$):

$$\hat{\mathbf{w}} = \mathbf{C}_{\mathbf{u}}^{1/2}\hat{\mathbf{u}} \tag{A4}$$

$$\hat{\mathbf{w}}' = \mathbf{C}_{\mathbf{u}}^{-1/2}\hat{\mathbf{u}}' \tag{A5}$$

In this way, the transformed SRF estimates are related through the filtering equation

$$\hat{\mathbf{w}}' = \mathbf{R}_w\hat{\mathbf{w}} \tag{A6}$$

where the matrix $\mathbf{R}_w$ is

$$\mathbf{R}_w = \mathbf{C}_{\mathbf{u}}^{-1/2}\mathbf{R}\mathbf{C}_{\mathbf{u}}^{-1/2} \tag{A7}$$

Based on Eqs. (A4) and (A5), we have that

$$\mathbf{C}_{\hat{\mathbf{w}}} = \mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2} \tag{A8}$$

$$\mathbf{C}_{\hat{\mathbf{w}}'} = \mathbf{C}_{\mathbf{u}}^{-1/2}\mathbf{C}_{\hat{\mathbf{u}}'}\mathbf{C}_{\mathbf{u}}^{-1/2} \tag{A9}$$

Taking into account the CV-matching constraint (A1), the last equation yields the following whitening constraint for the CV matrix of the transformed vector $\hat{\mathbf{w}}'$

$$\mathbf{C}_{\hat{\mathbf{w}}'} = \mathbf{C}_{\mathbf{u}}^{-1/2}\mathbf{C}_{\mathbf{u}}\mathbf{C}_{\mathbf{u}}^{-1/2} = \mathbf{I} \tag{A10}$$

Also, using the transformation formulae in (A4) and (A5), the MSE principle from (A3) can be now expressed as

$$\begin{aligned}
\varphi &= E\{(\mathbf{C}_{\mathbf{u}}^{1/2}\hat{\mathbf{w}}' - \mathbf{C}_{\mathbf{u}}^{-1/2}\hat{\mathbf{w}})^{\mathrm{T}}(\mathbf{C}_{\mathbf{u}}^{1/2}\hat{\mathbf{w}}' - \mathbf{C}_{\mathbf{u}}^{-1/2}\hat{\mathbf{w}})\}\\
&= E\{\hat{\mathbf{w}}'^{\mathrm{T}}\}\mathbf{C}_{\mathbf{u}}\hat{\mathbf{w}}' + \hat{\mathbf{w}}^{\mathrm{T}}\mathbf{C}_{\mathbf{u}}^{-1}\hat{\mathbf{w}} - \hat{\mathbf{w}}'^{\mathrm{T}}\hat{\mathbf{w}} - \hat{\mathbf{w}}^{\mathrm{T}}\hat{\mathbf{w}}'\}\\
&= E\{(\hat{\mathbf{w}}'^{\mathrm{T}}\hat{\mathbf{w}}' + \hat{\mathbf{w}}^{\mathrm{T}}\hat{\mathbf{w}} - \hat{\mathbf{w}}'^{\mathrm{T}}\hat{\mathbf{w}} - \hat{\mathbf{w}}^{\mathrm{T}}\hat{\mathbf{w}}')\\
&\quad +(\hat{\mathbf{w}}'^{\mathrm{T}}\mathbf{C}_{\mathbf{u}}\hat{\mathbf{w}}' + \hat{\mathbf{w}}^{\mathrm{T}}\mathbf{C}_{\mathbf{u}}^{-1}\hat{\mathbf{w}} - \hat{\mathbf{w}}'^{\mathrm{T}}\hat{\mathbf{w}}' - \hat{\mathbf{w}}^{\mathrm{T}}\hat{\mathbf{w}})\}\\
&= E\{(\hat{\mathbf{w}}' - \hat{\mathbf{w}})^{\mathrm{T}}(\hat{\mathbf{w}}' - \hat{\mathbf{w}})\} + E\{\hat{\mathbf{w}}'^{\mathrm{T}}\mathbf{C}_{\mathbf{u}}\hat{\mathbf{w}}'\\
&\quad +\hat{\mathbf{w}}^{\mathrm{T}}\mathbf{C}_{\mathbf{u}}^{-1}\hat{\mathbf{w}} - \hat{\mathbf{w}}'^{\mathrm{T}}\hat{\mathbf{w}}' - \hat{\mathbf{w}}^{\mathrm{T}}\hat{\mathbf{w}})\}\\
&= \mathrm{minimum}
\end{aligned} \tag{A11}$$

Note that both vectors $\hat{\mathbf{w}}$ and $\hat{\mathbf{w}}'$ have zero means, since the LSC solution $\hat{\mathbf{u}}$ (and thus also $\hat{\mathbf{u}}' = \mathbf{R}\hat{\mathbf{u}}$) is also a zero-mean random vector based on the initial assumptions stated in Sect. 2.1. Hence, the previous MSE principle can finally take the form

$$\begin{aligned}
\varphi &= E\{(\hat{\mathbf{w}}' - \hat{\mathbf{w}})^{\mathrm{T}}(\hat{\mathbf{w}}' - \hat{\mathbf{w}})\} + \mathrm{trace}(\mathbf{C}_{\mathbf{u}}\mathbf{C}_{\hat{\mathbf{w}}'})\\
&\quad + \mathrm{trace}(\mathbf{C}_{\mathbf{u}}^{-1}\mathbf{C}_{\hat{\mathbf{w}}}) - \mathrm{trace}\mathbf{C}_{\hat{\mathbf{w}}'} - \mathrm{trace}\mathbf{C}_{\hat{\mathbf{w}}}\\
&= \mathrm{minimum}
\end{aligned} \tag{A12}$$

or, by substituting from (A8) and (A10),

$$\begin{aligned}
\varphi &= \tilde{\varphi} + \mathrm{trace}(\mathbf{C}_{\mathbf{u}} + \mathbf{C}_{\mathbf{u}}^{-1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2} - \mathbf{I} - \mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2})\\
&= \mathrm{minimum}
\end{aligned} \tag{A13}$$

where

$$\tilde{\varphi} = E\{(\hat{\mathbf{w}}' - \hat{\mathbf{w}})^{\mathrm{T}}(\hat{\mathbf{w}}' - \hat{\mathbf{w}})\} \tag{A14}$$

Since $\mathbf{C}_{\mathbf{u}}$ and $\mathbf{C}_{\hat{\mathbf{u}}}$ are both known and fixed matrices not depending on $\mathbf{R}$, we conclude that finding the optimal filtering matrix $\mathbf{R}$ that minimizes the quantity φ in (A3), subject to the CV-matching constraint (A1), is equivalent to finding the filtering matrix $\mathbf{R}_w = \mathbf{C}_{\mathbf{u}}^{-1/2}\mathbf{R}\mathbf{C}_{\mathbf{u}}^{-1/2}$ that minimizes the quantity $\tilde{\varphi}$ in (A14), subject to the CV-matching whitening constraint (A10).

Let us consider the orthogonal decomposition of the CV matrix $\mathbf{C}_{\hat{\mathbf{w}}}$ given in (A8)

$$\mathbf{C}_{\hat{\mathbf{w}}} = \mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2} = \mathbf{V}\mathbf{D}\mathbf{V}^{\mathrm{T}} \tag{A15}$$

where $\mathbf{V}$ is a unitary matrix whose columns correspond to the orthonormal eigenvectors of $\mathbf{C}_{\hat{\mathbf{w}}}$, and $\mathbf{D}$ is a diagonal matrix that contains the corresponding eigenvalues.

If we apply the following invertible transformation to the random vectors $\hat{\mathbf{w}}$ and $\hat{\mathbf{w}}'$

$$\hat{\mathbf{z}} = \mathbf{V}^{\mathrm{T}}\hat{\mathbf{w}} \tag{A16}$$

$$\hat{\mathbf{z}}' = \mathbf{V}^{\mathrm{T}}\hat{\mathbf{w}}' \tag{A17}$$

then we have immediately that

$$\mathbf{C}_{\hat{\mathbf{z}}} = \mathbf{V}^{\mathrm{T}}\mathbf{C}_{\hat{\mathbf{w}}}\mathbf{V} \overset{\text{Eq. (A15)}}{=} \mathbf{D} \tag{A18}$$

$$\mathbf{C}_{\hat{\mathbf{z}}'} = \mathbf{V}^{\mathrm{T}}\mathbf{C}_{\hat{\mathbf{w}}'}\mathbf{V} \overset{\text{Eq. (A10)}}{=} \mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I} \tag{A19}$$

so that both vectors $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ are uncorrelated.

Furthermore, it holds that

$$\hat{\mathbf{z}}' = \mathbf{R}_z\hat{\mathbf{z}} \tag{A20}$$

where the matrix $\mathbf{R}_z$ is related with the previous filtering matrix $\mathbf{R}_w$ according to the equation

$$\mathbf{R}_z = \mathbf{V}^{\mathrm{T}}\mathbf{R}_w\mathbf{V} \tag{A21}$$

Also, due to the orthogonality of the vector transformation applied in (A16) and (A17), the quantity $\tilde{\varphi}$ can be equivalently expressed as

$$\tilde{\varphi} = E\{(\hat{\mathbf{z}}' - \hat{\mathbf{z}})^{\mathrm{T}}(\hat{\mathbf{z}}' - \hat{\mathbf{z}})\} \tag{A22}$$

Taking into account that both $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ are zero-mean random vectors (since they are obtained through successive linear transformations applied to the zero-mean LSC solution $\hat{\mathbf{u}}$ and the CV-matching field solution $\hat{\mathbf{u}}' = \mathbf{R}\hat{\mathbf{u}}$, respectively), the above quantity takes the form

$$\begin{aligned}\tilde{\varphi} &= E\{\hat{\mathbf{z}'}^{\mathrm{T}}\hat{\mathbf{z}}'\} + E\{\hat{\mathbf{z}}^{\mathrm{T}}\hat{\mathbf{z}}\} - 2E\{(\hat{\mathbf{z}}^{\mathrm{T}}\hat{\mathbf{z}}'\} \\ &= \mathrm{trace}(\mathbf{C}_{\hat{\mathbf{z}}'}) + \mathrm{trace}(\mathbf{C}_{\hat{\mathbf{z}}}) - 2E\{(\hat{\mathbf{z}}^{\mathrm{T}}\hat{\mathbf{z}}'\} \\ &= \mathrm{trace}(\mathbf{I}) + \mathrm{trace}(\mathbf{D}) - 2E\{(\hat{\mathbf{z}}^{\mathrm{T}}\hat{\mathbf{z}}'\}\end{aligned} \tag{A23}$$

The first two terms in the above equation are fixed and they do not depend on the desired filtering matrix $\mathbf{R}$ (or any of its transformed versions $\mathbf{R}_w = \mathbf{C}_{\mathbf{u}}^{-1/2}\mathbf{R}\mathbf{C}_{\mathbf{u}}^{-1/2}$ and $\mathbf{R}_z = \mathbf{V}^{\mathrm{T}}\mathbf{R}_w\mathbf{V}$). Note that the diagonal matrix $\mathbf{D}$ is uniquely specified through the eigenvalue decomposition in (A15), since $\mathbf{C}_{\mathbf{u}}$ and $\mathbf{C}_{\hat{\mathbf{u}}}$ are both known and given symmetric matrices; see (A15). In this way, the minimization of $\tilde{\varphi}$ with respect to the choice of the filtering matrix $\mathbf{R}_z$ is equivalent to the minimization of the term

$$\rho = -2E\{(\hat{\mathbf{z}}^{\mathrm{T}}\hat{\mathbf{z}}'\} \tag{A24}$$

Taking into account the linearity of the expectation operator $E\{\cdot\}$, we can express the quantity ρ in the alternative form

$$\rho = -2\sum_{k=1}^{n} E\{\hat{z}[k]\hat{z}'[k]\} \tag{A25}$$

where $\hat{z}[k]$ and $\hat{z}'[k]$ denote the k-th elements of the vectors $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$, respectively. In order to find the minimum of ρ we will use the well known *cosine inequality* (Papoulis 1991, p. 154)

$$E\{x_1x_2\} \leq \sqrt{E\{x_1^2\}E\{x_2^2\}} \tag{A26}$$

which is valid for any pair (x_1, x_2) of real-valued random variables. It needs to be emphasized that the equality sign in the previous inequality holds if and only if the random variables are linearly related, $x_2 = ax_1$, where a is an arbitrary real scalar.

Putting the elements $\hat{z}[k]$ and $\hat{z}'[k]$ in place of the arbitrary random variables x_1 and x_2, the cosine inequality of (A26) yields

$$E\{\hat{z}[k]\hat{z}'[k]\} \leq \sqrt{E\{(\hat{z}[k])^2\}E\{(\hat{z}'[k])^2\}} \tag{A27}$$

or equivalently

$$-2E\{\hat{z}[k]\hat{z}'[k]\} \geq -2\sqrt{E\{(\hat{z}[k])^2\}E\{(\hat{z}'[k])^2\}} \tag{A28}$$

and by summing over all possible values of the index k, we finally obtain

$$-2\sum_{k=1}^{n} E\{\hat{z}[k]\hat{z}'[k]\} \geq -2\sum_{k=1}^{n}\sqrt{E\{(\hat{z}[k])^2\}E\{(\hat{z}'[k])^2\}} \tag{A29}$$

or equivalently

$$\rho \geq -2\sum_{k=1}^{n}\sqrt{E\{(\hat{z}[k])^2\}E\{(\hat{z}'[k])^2\}} \tag{A30}$$

The equality sign in (A30)s holds *if and only if* the cosine inequality in (A27) becomes an equality for every value of the index k, i.e.,

$$E\{\hat{z}[k]\hat{z}'[k]\} = \sqrt{E\{(\hat{z}[k])^2\}E\{(\hat{z}'[k])^2\}}, \quad k=1,2,\ldots,n \tag{A31}$$

which obviously holds if and only if the random variables $\hat{z}[k]$ and $\hat{z}'[k]$ are linearly related

$$\hat{z}'[k] = \lambda_k\hat{z}[k], \quad k = 1, 2, \ldots, n \tag{A32}$$

where λ_k denotes an arbitrary scalar factor. Hence, we can conclude that the minimization of the term ρ requires that the random variables $\hat{z}[k]$ and $\hat{z}'[k]$ should be related according to above linear equation.

In view of (A32), and considering the fact that $\mathbf{C}_{\hat{\mathbf{z}}} = \mathbf{D}$ and $\mathbf{C}_{\hat{\mathbf{z}}'} = \mathbf{I}$ [see Eqs. (A18) and (A18)], we easily deduce that the vectors $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ should be related through a *diagonal* matrix $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_k, \ldots, \lambda_n)$

$$\hat{\mathbf{z}}' = \Lambda\hat{\mathbf{z}} \tag{A33}$$

such that $\Lambda = \mathbf{D}^{-1/2}$. In this way, the matrix $\mathbf{R}_z$ according to the filtering formula in (A20) becomes

$$\mathbf{R}_z = \mathbf{D}^{-1/2} \tag{A34}$$

From Eq. (A21), and taking also into account Eq. (A15), we obtain the filtering matrix $\mathbf{R}_w$

$$\begin{aligned}\mathbf{R}_w &= \mathbf{V}\mathbf{R}_z\mathbf{V}^{\mathrm{T}} = \mathbf{V}\mathbf{D}^{-1/2}\mathbf{V}^{\mathrm{T}} \\ &= (\mathbf{C}_{\hat{\mathbf{w}}})^{-1/2} \\ &= \left[\mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2}\right]^{-1/2}\end{aligned} \tag{A35}$$

Using the last result, the required filtering matrix $\mathbf{R}$ is finally determined from (A7) as

$$\begin{aligned}\mathbf{R} &= \mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{R}_w\mathbf{C}_{\mathbf{u}}^{1/2} \\ &= \mathbf{C}_{\mathbf{u}}^{1/2}\left[\mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2}\right]^{-1/2}\mathbf{C}_{\mathbf{u}}^{1/2}\end{aligned} \tag{A36}$$

In order to conclude the proof, it is required to show that the above result satisfies the CV-matching constraint in (A1). We have

$$\begin{aligned}\mathbf{R}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{R}^{\mathrm{T}} &= \mathbf{C}_{\mathbf{u}}^{1/2}\left[\mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2}\right]^{-1/2}\mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2} \\ &\quad \left[\mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2}\right]^{-1/2}\mathbf{C}_{\mathbf{u}}^{1/2} \\ &= \mathbf{C}_{\mathbf{u}}^{1/2}\left[\mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2}\right]^{-1/2} \\ &\quad \left[\mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2}\right]\left[\mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\hat{\mathbf{u}}}\mathbf{C}_{\mathbf{u}}^{1/2}\right]^{-1/2}\mathbf{C}_{\mathbf{u}}^{1/2} \\ &= \mathbf{C}_{\mathbf{u}}^{1/2}\mathbf{C}_{\mathbf{u}}^{1/2} \\ &= \mathbf{C}_{\mathbf{u}}\end{aligned} \tag{A37}$$

which concludes our proof.

References

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

Christakos G (1992) Random field models for earth sciences. Academic, New York

Dermanis A (1976) Probabilistic and deterministic aspects of linear estimation in geodesy. Report No. 244, Department of Geodetic Science, Ohio State University, Columbus

Dermanis A (1980) Adjustment of geodetic observations in the presence of signals. In: Proceedings of the International School of advanced geodesy. Boll Geod Sci Affi, vol. 38, pp. 419–445

Dermanis A (1984) Kriging and collocation—a comparison. Manusc. Geod 9:159–167

Dermanis A (1990) A unified approach to linear estimation and prediction. Presented at the 20th IUGG General Assembly, Vienna, August 28–19, 1990. (available at http://der.topo.auth.gr/DERMANIS/ENGLISH/Index_ENGLISH.htm)

Deutsch CV, Journel AG (1998) GSLIB: Geostatistical software library and user's guide, 2nd edn. Oxford University Press, New York

Eeg J, Krarup T (1973) Integrated geodesy. Danish Geodetic Institute, Internal report no. 7, Copenhagen

Eldar YC (2001) Quantum signal processing. PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge

Eldar YC (2003) Minimum mean-squared error covariance shaping. In: Proceedings of international conference on acoust, speech, signal processing (ICASSP-2003), Hong-Kong, vol. 6, pp 713–716

Eldar YC, Oppenheim AV (2003) MMSE whitening and subspace whitening. IEEE Trans Info Theory 49(7):1846–1851

Eren K (1982) Toeplitz matrices and frequency domain collocation. Manusc. Geod 7:85–116

Higham NJ (1987) Computing real square roots of a real matrix. Linear Algebra Appl 88/89:405–430

Higham NJ (1997) Stable iterations for the matrix square root. Numer Algorithm, 15:227–242

Higham NJ (2006) The matrix computation toolbox. http://www.ma.man.ac.uk/~higham/mctoolbox

Iannazzo B (2006) On the Newton method for the matrix pth root. SIAM J Matrix Anal 28:503–523

Journel AG (1990) Fundamentals of geostatistics in five lessons. American Geophysical Union, Washington DC

Koch K-R (1999) Parameter estimation and hypothesis testing in linear models, 2nd edn. Springer, Berlin

Kolmogorov AN (1941) Interpolation and extrapolation of stationary random sequences. Izv Akad Nauk SSSR, Seriia Mat 5:3–14

Kotsakis C (2000) The multiresolution character of collocation. J Geod 74(3–4):275–290

Krarup T (1969) A contribution to the mathematical foundation of physical geodesy. Danish Geodetic Institute, Report no. 44, Copenhagen

Matheron G (1971) The theory of regionalized variables and its applications. Les Cahiers du Centre de Morphologie Mathematique de Fontainebleu No 5, Ecole Nationale Superieure de Mines de Paris, Paris

Meier S (1981) Planar geodetic covariance functions. Rev Space Phys 19(4):673–686

Moritz H (1970) Least-squares estimation in physical geodesy. Deutsche Geodatische Kommission, Reihe A, Heft 69, Munchen

Moritz H (1973) Least-squares collocation. Deutsche Geodatische Kommission, Reihe A, Heft 75, Munchen

Moritz H (1980) Advanced physical geodesy. Herbert Wichmann Verlag, Karlsruhe

Nash RA, Jordan SK (1978) Statistical geodesy: an engineering perspective. Proc IEEE 66(5):532–550

Papoulis A (1991) Probability, random variables and stochastic processes, 3rd edn. McGraw-Hill, New York

Reguzzoni M, Sanso F, Venuti G (2005) The theory of general kriging with applications to the determination of a local geoid. Geoph J Int, doi: 10.1111/j.1365-246X.2005.02662.x

Sanso F (1980) The minimum mean squared estimation error principle in physical geodesy (stochastic and non-stochastic interpretation). Boll Geod Sci Affi 39(2):112–129

Sanso F (1986) Statistical methods in physical geodesy. In: Sunkel H (ed) Mathematical and numerical techniques in physical geodesy. Lecture notes in Earth sciences, vol 7, pp 49–156, Springer, Berlin

Sanso F, Sideris MG (1997) On the similarities and differences between systems theory and least-squares collocation in physical geodesy. Boll Geod Sci Affi 54(2):173–206

Schwarz K-P, Sideris MG, Forsberg R (1990) The use of FFT techniques in physical geodesy. Geophys J Int 100:485–514

Sideris MG (1995) On the use of heterogeneous noisy data in spectral gravity field modeling methods. J Geod 70:470–479

Smith MI (2003) A Schur algorithm for computing matrix pth roots. SIAM J Matrix Anal Appl 24:971–989

Theobald CM (1974) Generalizations of the mean square error applied to ridge regression. J R Stat Soc Ser B 36(1):103–106

Tscherning CC (1978) On the convergence of least squares collocation. Boll Geod Sci Affi 37(2,3):507–517

Tscherning CC (1986) Functional methods for gravity field approximation. In: Sunkel H (ed) Mathematical and numerical techniques in physical geodesy. Lecture notes in Earth sciences, vol 7, pp 3–48, Springer, Berlin

Wiener N (1949) Extrapolation, interpolation and smoothing of stationary time series. MIT Press, Cambridge