Journal of Geodesy (2001) 75: 647–660

Journal of Geodesy

A modified Wiener-type filter for geodetic estimation problems with non-stationary noise

C. Kotsakis[1], M. G. Sideris[2]

[1] Department of Geodesy and Surveying, Faculty of Engineering, Aristotle University of Thessaloniki, Univ. Box 440, Thessaloniki 540 06, Greece; e-mail: kotsakis@alumni.ucalgary.ca; Fax: +30-31-995948
[2] Department of Geomatics Engineering, University of Calgary, 2500 University Drive NW, Calgary T2N 1N4, Alberta, Canada; e-mail: sideris@ucalgary.ca; Fax: +1-403-2841980

Received: 1 November 2000 / Accepted: 19 June 2001

Abstract. One of the most basic and important tools in optimal spectral gravity field modelling is the method of Wiener filtering. Originally developed for applications in analogue signal analysis and communication engineering, Wiener filtering has become a standard linear estimation technique of modern operational geodesy, either as an independent practical tool for data de-noising in the frequency domain or as an integral component of a more general signal estimation methodology (input–output systems theory). Its theoretical framework is based on the Wiener–Kolmogorov linear prediction theory for stationary random fields in the presence of additive external noise, and thus it is closely related to the (more familiar to geodesists) method of least-squares collocation with random observation errors. The main drawback of Wiener filtering that makes its use in many geodetic applications problematic stems from the stationarity assumption for both the signal and the noise involved in the approximation problem. A modified Wiener-type linear estimation filter is introduced that can be used with noisy data obtained from an arbitrary deterministic field under the masking of non-stationary random observation errors. In addition, the sampling resolution of the input data is explicitly taken into account within the estimation algorithm, resulting in a resolution-dependent optimal noise filter. This provides a more insightful approach to spectral filtering techniques for noise reduction, since the data resolution parameter has not been directly incorporated in previous formulations of frequency-domain estimation problems for gravity field signals with discrete noisy data.

Key words: Wiener filter – Non-stationary noise – Data resolution – Translation-invariant estimation – Fast Fourier transform

Correspondence to: C. Kotsakis

1 Introduction: problem statement

Wiener filtering is a well known and efficient optimal estimation method that can be used for geodetic data 'de-noising' in the frequency domain. Its theoretical background is based on the famous Wiener–Kolmogorov (W-K) linear prediction theory for stationary random fields in the presence of stationary additive noise (Kolmogorov 1941; Wiener 1949). For a comprehensive mathematical exposition of various issues related to Wiener filtering theory and its extensions, see the review paper by Kailath (1974) which contains circa 350 additional relevant references on this topic. The application of the Wiener filter in geodesy, either as an independent practical tool for data pre-processing or as an integral component of a more general linear estimation methodology (input–output systems theory, Sideris 1996), has been mainly focused on problems related to optimal spectral gravity field modelling. Many studies have been performed using, implicitly or explicitly, the Wiener filtering procedure for various physical geodesy estimation problems, including: de-noising of gravity anomaly data prior to gravimetric geoid computations (Li and Sideris 1994); optimal separation of the gravity anomaly signal from external noise (and other residual) effects for the identification of certain crustal geological features (Pawlowski and Hansen 1990); optimal spectral combination of shipborne gravity and altimetric data for marine gravity field modelling (Li 1996; Tziavos et al. 1996, 1998); simultaneous optimal noise filtering of airborne gravity vector data (Wu and Sideris 1995); and optimal frequency-domain estimation of the anomalous potential from airborne gradiometry data (Vassiliou 1986). A detailed discussion on the use of the W–K filtering theory in gravity field estimation, and its relationship with other linear approximation techniques traditionally used in geodesy [e.g. least-squares (LS) collocation], can be found in Sansò and Sideris (1997); see also Sideris (1996) and Schwarz et al. (1990).

There are basically two main drawbacks associated with the usual formulation of the Wiener spectral filtering algorithm that make its use in geodetic applications problematic. First, the necessary assumption for noise stationarity can be quite unrealistic in many practical situations, since the observational errors in the input data often have significant spatial (or temporal) variations in their statistical behaviour. A typical example of such case arises in gravimetric geoid determination from a single grid of gravity anomalies that have been obtained from heterogeneous sources, such as terrestrial, shipborne and airborne gravity surveys. Combined gravity anomaly grids for large-scale geoid computations may also be constructed from terrestrial gravimetric databases obtained at various epochs with different instrumentation and reduction procedures, and thus different accuracy levels. Similar problems are encountered in airborne gravimetry/gradiometry, where the data noise level can change considerably with respect to time, depending on the dynamical behaviour of the gravity sensors (e.g. acceleration noise is usually amplified during aircraft turns). In all of the above cases, the non-stationary nature of the data noise makes the application of the classic Wiener filter impossible, at least from a theoretical point of view. Of course, optimal data de-noising and signal estimation with non-stationary observation errors can always be performed using the standard space-domain collocation formulae, which are nevertheless not very efficient (in terms of both computational speed and numerical stability), especially for large data sets with high sampling resolution.

The second problem of geodetic importance in the Wiener filtering theory originates from the modelling requirements for the true (unknown) signal, which is also assumed to represent a stationary stochastic process, usually uncorrelated with the external additive noise. In this way, the power spectral density (PSD) functions of the true signal and noise, both defined according to the fundamental *Wiener–Khinchine* probabilistic relationship (Bendat and Piersol 1993, pp. 55–56), are the two essential components needed for the computation of the optimal estimation/de-noising filter. In physical geodesy applications, however, a stochastic interpretation of the underlying true field values as random variables is rather questionable, to say the least, since repetitive noiseless signal measurements should always give the same result (Moritz and Sansò 1980). Although the stochastic/non-stochastic modelling dilemma in optimal gravity field approximation has been successfully resolved by Sansò (1980) within a *noiseless* linear estimation framework, a similar treatment does not currently exist for the case of spectral filtering methods with noisy input data. Furthermore, the behaviour of most geodetic signals is usually far from being uniform across their domain, thus making the stationarity assumption once again a poor modelling choice for many practical applications.

Taking into account the previous discussion and the important role that fast spectral techniques play in modern operational geodesy, we feel that a new revised formulation for the noise filtering problem of geodetic data is needed, in such a way that the two aforementioned limitations of the classic Wiener filter can be overcome. Hence, the aim of this paper is to present a modified Wiener-type optimal noise filter which can be used in linear estimation problems with arbitrary deterministic signals that are masked by additive non-stationary observation errors. An important point in our approach is that the sampling resolution of the input data will be explicitly taken into account within the optimization procedure, resulting in a *resolution-dependent* noise filter. This will help us identify an interesting interplay between measurement noise and data resolution in linear signal estimation, and clarify the nature of their individual contributions to the total output approximation error.

The structure of the paper is organized as follows. In Sect. 2, the main mathematical background and the modelling assumptions associated with our signal estimation methodology are presented in detail. In Sect. 3, a spectral optimization procedure for the linear noise filter is developed. The basic properties of our optimal estimation filter and its informal similarities with the classic Wiener filter are discussed in Sect. 4, whereas some additional remarks of both theoretical and practical interest are given in Sect. 5. A numerical simulation example has been included in Sect. 6 to demonstrate the performance of our optimal noise filter under non-stationary additive noise at different data resolution levels. Finally, the last section of the paper contains some conclusions and a brief discussion on remaining open problems that require further research.

2 Background

In this section, the main assumptions and the basic mathematical notation that will be used throughout the rest of this paper are presented. We have chosen to follow a relatively simple approximation framework in a two-dimensional (2-D) planar setting, which nevertheless can fit nicely to many spatial estimation problems of geodetic interest (local gravity field approximation, digital terrain modelling, etc.), as well as to other geomatics-related applications with noisy input data (e.g. digital image processing).

2.1 General formulation

The unknown object of our estimation procedure will be modelled as a 2-D deterministic signal $g(x,y)$ with compact spatial support over the real plane $\Re^2$. Its finite extent along the two orthogonal axes x and y is denoted by X and Y, respectively. No specific restrictions on the behaviour of the unknown field will be imposed, apart from the assumption that it posses a well-defined Fourier transform

$$G(\omega_x,\omega_y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\mathrm{e}^{-i(\omega_x x+\omega_y y)}\,\mathrm{d}x\,\mathrm{d}y = \int_{Y}\int_{X} g(x,y)\mathrm{e}^{-i(\omega_x x+\omega_y y)}\,\mathrm{d}x\,\mathrm{d}y \quad (1)$$

which essentially means that $g(x,y)$ is a member of the $L_1(\Re^2)$ space. In this way, the unknown signal is allowed to exhibit irregular local variations at any scale level. The two parameters ω_x and ω_y are the spatial circular frequencies along the x and y axes, respectively. A regular grid of the signal values will be denoted by $g(nh_x,mh_y)_{n,m\in Z}$ or simply $g(nh_x,mh_y)$, where the symbols h_x and h_y correspond to the sampling intervals along the x and y directions. Without loss of generality, we can assume that the compact support of the unknown field is enclosed by the parallelogram region $0\le x\le X$ and $0\le y\le Y$, and thus the integer sampling indices can practically be restricted within the finite range $0\le n\le N-1$ and $0\le m\le M-1$, where $X=(N-1)h_x$ and $Y=(M-1)h_y$. The 2-D Fourier transform of such a signal grid will be denoted by $\bar{G}(\omega_x,\omega_y)$ and it is given by the summation formula (Dudgeon and Mersereau 1984)

$$\bar{G}(\omega_x,\omega_y)=\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} g(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)} = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} g(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)} = \frac{1}{h_x h_y}\sum_{k=-\infty}^{+\infty}\sum_{l=-\infty}^{+\infty} G\left(\omega_x+\frac{2\pi k}{h_x},\omega_y+\frac{2\pi l}{h_y}\right) \quad (2)$$

where the last part in the above equation expresses the aliasing effect on the Fourier transform of the original continuous signal. The overbar symbol will be used to indicate a periodic function, and the lower case letters n, m, k and l are always reserved to denote integer numbers. Some technical mathematical details regarding the convergence of the third part in the above equation (Poisson summation formula) can be found in Gasquet and Witomski (1999, pp. 344–351).

The input data obtained from the unknown field $g(x,y)$ is given in a discrete gridded form according to the linear observation equation

$$d(nh_x,mh_y)=g(nh_x,mh_y)+v(nh_x,mh_y) \quad (3)$$

where $v(nh_x,mh_y)$ is a 2-D random noise sequence that is generally assumed to be non-stationary. The associated stochastic model used to describe the behaviour of the noise, in terms of second-order moment information, is defined by the equations

$$E\{v(nh_x,mh_y)\}=0 \quad (4a)$$

$$E\{v^2(nh_x,mh_y)\}=\sigma_v^2(nh_x,mh_y) = \sigma_v[(nh_x,mh_y)(nh_x,mh_y)] \quad (4b)$$

$$E\{v(nh_x,mh_y)v(kh_x,lh_y)\}=\sigma_v[(nh_x,mh_y)(kh_x,lh_y)] \quad (4c)$$

where E is the probabilistic expectation operator. The symbol $\sigma_v^2(\cdot)$ denotes the noise variance at a specific data point on the real plane, whereas $\sigma_v[(\cdot)(\cdot)]$ corresponds to the noise covariance (CV) between two data points. We will also use the symbol $\bar{V}(\omega_x,\omega_y)$ for the 2-D Fourier transform of the random observation errors, which is defined as

$$\bar{V}(\omega_x,\omega_y)=\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} v(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)} = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} v(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)} \quad (5)$$

Similarly, the 2-D Fourier transform of the gridded data is given by the equation

$$\bar{D}(\omega_x,\omega_y)=\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} d(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)} = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} d(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)} = \bar{G}(\omega_x,\omega_y)+\bar{V}(\omega_x,\omega_y) \quad (6)$$

Note that the noise is zero everywhere outside the input data grid (NM points) because the unknown field has been assumed to have finite spatial support and thus no measurements are performed outside this region. In this way, the existence/convergence of both Fourier transforms in Eqs. (5) and (6) is always guaranteed regardless of the behaviour of the noise (i.e. stationary or non-stationary), since the sequences $v(nh_x,mh_y)$ and $d(nh_x,mh_y)$ will have only a finite number of non-zero terms.

2.2 Continuous versus discrete noise

In the previous section we adopted a discrete. Gauss–Markov stochastic model for the data noise. This is in contrast to other existing formulations of optimal estimation problems in spectral gravity field approximation, where a spatially continuous (and stationary) stochastic model for the data errors is usually employed with the help of stationary noise CV functions (Sideris 1996; Tziavos et al. 1996, 1998; Li and Sideris 1997). The use of continuous noise models is a rather questionable practice within an estimation framework utilizing only discrete spatial data. The measurement noise does not generally exist in a physical sense as a continuous spatial signal (i.e. we do not 'sample the noise'), but originates only because we performed an observation with an imperfect instrument under certain external influences at a specific point.

On the other hand, there exist cases of signal approximation problems with discrete spatial data where it does make sense to consider continuous noise models.

For example, a data acquisition device may change its noise characteristics, as it moves from one spatial point to another, according to a continuous (time-dependent) error model. However, since we always collect and process observations at a discrete network of data points with finite resolution, the input noise in the estimation algorithm will still be a discrete signal (in a spatial sense) with an associated discrete stochastic model. The latter is actually determined in such cases, at the points of interest, from a continuous time-dependent error model and the spatio-temporal 'path' of the measuring device. We may also consider cases of estimation problems with discrete spatial data where the observational noise, or some part of it, does indeed exist in a spatially continuous sense (e.g. atmospheric effects in various types of geodetic measurements).

In any case, the main aspect to be emphasized here is that we do not strictly need continuous noise models in estimation schemes with finite-resolution discrete data (although their use often facilitates the accuracy description of the input data via simple analytical formulae). As it will be demonstrated in the next sections, all that is *algorithmically* necessary for obtaining the signal estimate is the second-order discrete model given in Eqs. (4a), (4b) and (4c), even if the underlying data noise is generated by continuous time/spatial phenomena. For the treatment of noise in continuous fields of measurements, see Sansò and Sona (1995).

2.3 A priori properties of the estimation algorithm

The main problem that is studied in this paper is the frequency-domain estimation of the unknown deterministic field $g(x,y)$ from its noisy gridded samples $d(nh_x, mh_y)$ according to Eq. (3). Two basic properties will be imposed a priori in the estimation procedure, namely *linearity* and *translation-invariance*. The reason for introducing the second property is to obtain a signal estimate $\hat{g}(x,y)$ that is independent of the reference system used to describe the position of the data points. Stated in a simplified way, if we change the origin of the reference coordinate system xy on the real plane by arbitrary shifts t_x and t_y (without 'moving' the unknown field or the associated data grid), we want the new solution to be just a translated version $\hat{g}(x+t_x, y+t_y)$ of the initial estimate in the original reference system (see Fig. 1). For a general discussion on the translation-invariance property in linear interpolation/estimation problems and its importance, see Thévenaz et al. (2000) and Unser (2000).

At this point, it can be argued that non-shift-invariant estimation algorithms may have better performance, under certain optimality criteria, than simpler translation-invariant methods. Indeed, this is the case in the classic W–K linear prediction theory where the optimal signal estimate becomes translation-invariant only if the unknown field and the additive noise are both stationary. A similar situation exists in semi-stochastic signal approximation schemes with discrete noisy data, such as the minimum-norm solution obtained from Tikhonov's quadratic regularization principle (Moritz 1980). In this case, the optimal estimate is shift-invariant only if the signal modelling takes place in a Hilbert space with a homogeneous reproducing kernel, and the random observation errors follow a stationary probabilistic model.

Here, on the other hand, we choose to apply a priori the translation-invariance restriction, in addition to linearity, in order to ensure that the result of the estimation algorithm is not affected by arbitrary shifts of the spatial reference system. (Note: for signal approximation problems on the sphere, instead of translations, we deal with arbitrary rotations of the spherical coordinate reference systems.) The justification of such a choice relies basically on simple logic and mathematical intuition, and it is *not* affected by the spatio-statistical properties of the true signal and noise involved in the approximation problem. If we choose to follow a non translation-invariant methodology, we should at least be able to explain physically the dependence of the output signal estimate on the origin of the coordinate system used to reference the unknown field and its discrete input data. Note that the translation-invariance

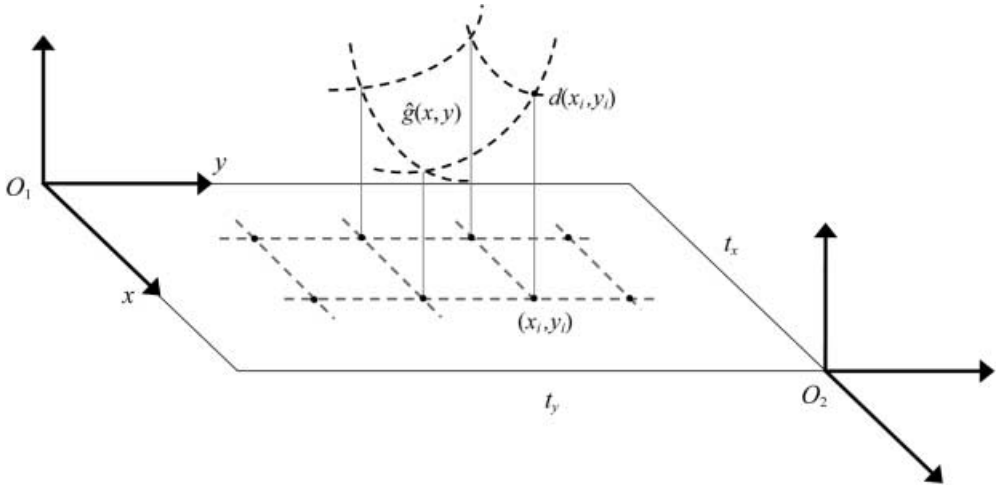

Fig. 1. A graphical explanation of the concept of translation-invariant estimation. Regardless of which reference coordinate system we use (O_1 or O_2), the signal estimate $\hat{g}(x,y)$ obtained from the data values $d(x_i,y_i)$ should always have the same shape/behaviour

condition has often been applied in the theoretical formulation of optimal estimation methods using errorless data (Sansò 1980; Kotsakis 2000a), although its justification is not altered by the presence of noise in the observations.

Based on the two assumptions of linearity and translation-invariance, the signal estimation formula will have the general convolution-type expression

$$\hat{g}(x,y) = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} d(nh_x, mh_y)\xi_h(x - nh_x, y - mh_y) \qquad (7)$$

where $\xi_h(x,y)$ is a 2-D filtering kernel that needs to be determined in some optimal sense. The subscript h is used to indicate that the estimation kernel will generally depend on the specific data resolution levels, h_x and h_y. The above equation can be illustrated through the linear input–output (I/O) system shown in Fig. 2.

Remark. It would be proper at this point to ask ourselves if the signal approximation model in Eq. (7) provides the most general choice in the class of linear and translation-invariant estimators, or whether it is restricted in some sense in order to simplify the following derivations. This important question was raised by Sansò (personal communication 2001), who pointed out that the use of a *single* kernel $\xi_h(x,y)$ in Eq. (7) confines the family of the estimators in which we will seek an optimal signal solution. In particular, consider the linear interpolation formula of least-squares collocation (LSC) which also gives a translation-invariant estimate (when used with a stationary/homogeneous covariance function), but it cannot be exactly reduced to the simpler expression of Eq. (7). Nevertheless, the usual matrix formula of LSC ignores the fact that the signal values are zero outside the input data grid. This additional information comes from the initial assumptions that were set in Sect. 2.1, i.e. the unknown object is a finite-support signal over the entire real plane. Actually, for the case of gridded data with uniform sampling resolution, it can be shown that the LSC algorithm can easily be transformed into the linear expression of Eq. (7), if we expand/augment the input data vector with an infinite number of zeros that are taken at the same resolution as the original data values. We certainly recognize that the above points deserve further and more detailed discussion which, however, is not possible to provide in this paper without deviating from its original aim. Hence, in order to close this remark and continue with the main scope of our study, we should just say that (for the modelling assumptions set here) Eq. (7) is adopted as the most general choice of a linear and translation-invariant signal estimator from discrete gridded data with non-stationary observation errors. Some interesting details can also be found in Brovelli et al. (2000).

3 Optimization of the estimation kernel

In order to determine an optimal form for the estimation kernel $\xi_h(x,y)$, a corresponding optimality criterion must first be introduced. The signal error produced by the filtering formula in Eq. (7) can generally be decomposed into two components

$$e(x,y) = g(x,y) - \hat{g}(x,y) = e_h(x,y) + e_v(x,y) \qquad (8)$$

where $e_h(x,y)$ is the part of the total estimation error caused from the use of discrete data with finite resolution (aliasing error), and $e_v(x,y)$ is the additional part due to the noise presence in the signal samples. Note that the noise-dependent error component $e_v(x,y)$ is not completely unaffected by the data resolution but actually depends on it strongly, as will be demonstrated in later sections of this paper.

In the absence of any noise from the input data, the best we can do is to obtain just an *interpolated model* $\tilde{g}(x,y)$ for the unknown field that will depend on the true signal values at the given sampling resolution. We will assume that such a signal model is given in terms of a linear and translation-invariant formula, as follows:

$$\tilde{g}(x,y) = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} g(nh_x, mh_y)\varphi_h(x - nh_x, y - mh_y) \qquad (9a)$$

or, equivalently, in the frequency domain

$$\begin{aligned}\tilde{G}(\omega_x,\omega_y) &= \Phi_h(\omega_x,\omega_y)\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} g(nh_x, mh_y)\mathrm{e}^{-i(nh_x\omega_x + mh_y\omega_y)} \\ &= \Phi_h(\omega_x,\omega_y)\bar{G}(\omega_x,\omega_y) \end{aligned} \qquad (9b)$$

where $\varphi_h(x,y)$ is some interpolating kernel that generally depends on the sampling intervals h_x and h_y. The noise-dependent estimation error in $\hat{g}(x,y)$ will be measured with respect to such a noiseless interpolated model for the unknown field, i.e.

$$e_v(x,y) = \tilde{g}(x,y) - \hat{g}(x,y) \qquad (10a)$$

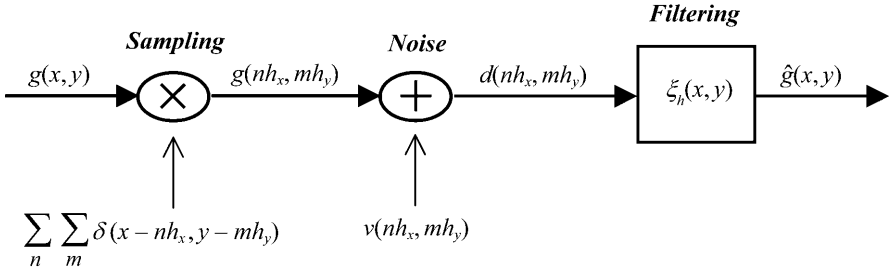

Fig. 2. Linear and translation-invariant signal estimation from discrete noisy samples

whereas the (pure) aliasing error is

$$e_h(x,y) = g(x,y) - \tilde{g}(x,y) \tag{10b}$$

The specific form of the kernel $\varphi_h(x,y)$ in Eq. (9a) is irrelevant for the purpose of this paper. A very popular choice that covers many different interpolating (or quasi-interpolating) schemes, including band-limited (Shannon) interpolation, B-spline interpolation and also more general wavelet-based approximation models, is based on the use of certain scaling functions $\varphi(x,y) \in L_2(\Re^2)$ which adapt to the data resolution through a dilation operation (Unser and Daubechies 1997; Unser 2000), i.e.

$$\varphi_h(x,y) = \varphi\left(\frac{x}{h_x}, \frac{y}{h_y}\right) \tag{11}$$

For the purpose of this paper, it is sufficient to consider $\varphi_h(x,y)$ as an arbitrarily chosen reference interpolating kernel with a well-defined Fourier transform $\Phi_h(\omega_x,\omega_y)$, which is used to obtain a continuous signal approximation in the absence of any noise from the discrete input data. In addition, it can be assumed that $\varphi_h(x,y)$ is such that: (i) the linear expansion in Eq. (9a) is always stable, and (ii) the aliasing error component $e_h(x,y)$ vanishes as the data resolution increases; for more details, see Blu and Unser (1999), Kotsakis (2000b) and Unser (2000).

The unknown filtering kernel in Eq. (7) will be determined by minimizing the *noise-dependent part* of the total signal error. In this way, the estimation problem is essentially reduced to a problem of finding an optimal modification for the reference signal interpolator $\varphi_h(x,y)$ that minimizes the propagated data noise in the final output field $\hat{g}(x,y)$. The optimization procedure will be carried out in the frequency domain using the familiar mean-square-error (MSE) criterion

$$P_{e_v}(\omega_x,\omega_y) = E\left\{\left|E_v(\omega_x,\omega_y)\right|^2\right\} = \text{minimum} \tag{12}$$

where $E_v(\omega_x,\omega_y)$ is the Fourier transform of $e_v(x,y)$, and $P_{e_v}(\omega_x,\omega_y)$ is the (noise-dependent) mean error power spectrum of the output signal. Note that the term 'mean' corresponds to its usual probabilistic interpretation, in contrast to the optimization scheme that is usually followed for the reference interpolator where the MSE is defined in a spatio-statistical deterministic sense. In Kotsakis (2000a, b) the optimal determination of the interpolation kernel $\varphi_h(x,y)$ was based on the minimization of the spatio-statistical power spectrum of the aliasing error component $e_h(x,y)$, whereas here the optimization of the noise filtering kernel $\xi_h(x,y)$ employs the mean power spectrum of the random error component $e_v(x,y)$.

From Eqs. (7), (9a) and (10a), we have that

$$\begin{aligned} E_v(\omega_x,\omega_y) = {} & \Phi_h(\omega_x,\omega_y)\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} g(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)} \\ & - \Xi_h(\omega_x,\omega_y)\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} d(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)} \end{aligned} \tag{13a}$$

where $\Xi_h(\omega_x,\omega_y)$ is the Fourier transform of the unknown filtering kernel $\xi_h(x,y)$. Using the shorthand notation according to Eqs. (2) and (6), the last equation takes the form

$$\begin{aligned} E_v(\omega_x,\omega_y) = {} & \Phi_h(\omega_x,\omega_y)\bar{G}(\omega_x,\omega_y) - \Xi_h(\omega_x,\omega_y)\bar{G}(\omega_x,\omega_y) \\ & - \Xi_h(\omega_x,\omega_y)\bar{V}(\omega_x,\omega_y) \end{aligned} \tag{13b}$$

By multiplying the above expression with its complex conjugate and taking the expected value, we can finally obtain

$$\begin{aligned} P_{e_v}(\omega_x,\omega_y) = {} & \Phi_h(\omega_x,\omega_y)\Phi_h^*(\omega_x,\omega_y)\left|\bar{G}(\omega_x,\omega_y)\right|^2 \\ & - \Phi_h(\omega_x,\omega_y)\Xi_h^*(\omega_x,\omega_y)\left|\bar{G}(\omega_x,\omega_y)\right|^2 \\ & - \Phi_h^*(\omega_x,\omega_y)\Xi_h(\omega_x,\omega_y)\left|\bar{G}(\omega_x,\omega_y)\right|^2 \\ & + \Xi_h(\omega_x,\omega_y)\Xi_h^*(\omega_x,\omega_y)\left|\bar{G}(\omega_x,\omega_y)\right|^2 \\ & + \Xi_h(\omega_x,\omega_y)\Xi_h^*(\omega_x,\omega_y)\bar{P}_v(\omega_x,\omega_y) \end{aligned} \tag{14}$$

where the asterisk $*$ denotes complex conjugation, and the auxiliary term $\bar{P}_v(\omega_x,\omega_y)$ corresponds to the noise PSD-like function

$$\bar{P}_v(\omega_x,\omega_y) = E\left\{\bar{V}(\omega_x,\omega_y)\bar{V}^*(\omega_x,\omega_y)\right\} = E\left\{\left|\bar{V}(\omega_x,\omega_y)\right|^2\right\} \tag{15}$$

For the derivation of the result in Eq. (14) we have used the fact that $E\{\bar{V}(\omega_x,\omega_y)\} = 0$, in accordance with the zero-mean stochastic model introduced for the data noise in Eq. (4a). The optimal estimation filter can now be determined using Eqs. (12) and (14). The underlying procedure is straightforward and it gives the final result

$$\begin{aligned} \Xi_h(\omega_x,\omega_y) &= \frac{\left|\bar{G}(\omega_x,\omega_y)\right|^2}{\left|\bar{G}(\omega_x,\omega_y)\right|^2 + \bar{P}_v(\omega_x,\omega_y)}\Phi_h(\omega_x,\omega_y) \\ &= \bar{W}(\omega_x,\omega_y)\Phi_h(\omega_x,\omega_y) \end{aligned} \tag{16}$$

The separable Wiener-like form of the above optimal filter is discussed in the next section.

4 The cascading structure of the optimal estimation filter

The final result in Eq. (16) indicates that the estimation algorithm can be decomposed into two individual steps which are connected in a linear cascading manner. The first step, expressed by the periodic filter component $\bar{W}(\omega_x,\omega_y)$, has the role of 'de-noising' the discrete input data using information on the average behaviour of the noise and the unknown field at the given resolution level. The second filter component $\Phi_h(\omega_x,\omega_y)$, on the other hand, is solely used to obtain a *continuous* representation for the estimated signal based on an a priori selected interpolating kernel $\varphi_h(x,y)$. These two basic procedures are illustrated in the linear input–output (I/O) system of Fig. 3. Note that, even though the optimization principle was applied to the continuous error term $e_v(x,y)$, the noise filtering part of the estimation algorithm always takes place at a discrete level (i.e. discrete input – discrete

output) and it is not affected by the choice of the reference interpolating model.

As it can be seen from Fig. 3, it is not necessary to 'modify' the interpolating kernel $\varphi_h(x, y)$ of the reference signal model in Eq. (9a) when dealing with noisy input data. The optimization of the noise-dependent output error adds only an intermediate filter that is applied to the original data grid $d(nh_x, mh_y)$, and it produces a new signal sequence $\hat{g}(nh_x, mh_y)$ from which the effect of the random observational errors has been minimized in a MSE sense. We can then use this 'denoised' grid as input to the interpolating model of Eq. (9a) in order to obtain a continuous approximation of the unknown field at the given resolution level. It should be noted that the interpolation filter $\Phi_h(\omega_x, \omega_y)$, shown in Fig. 3, can be additionally optimized by following a separate MSE approach that takes into account only the noise-free (aliasing) error component $e_h(x, y)$, as is described in Kotsakis (2000a, b).

The structure of the noise filter $\bar{W}(\omega_x, \omega_y)$ in Eq. (16) is very similar to the classic Wiener filter, since they are both defined in terms of a signal-to-noise ratio (SNR) expression. However, there do exist conceptual differences between the two filtering schemes because, in our formulation: (i) the unknown field has been modelled as a deterministic (instead of stochastic) signal; and (ii) the additive data noise has not been restricted to being stationary. Therefore, it is important to clarify the exact meaning of the two frequency-domain terms that appear in the SNR algebraic expression of our noise filter. From Eq. (16), we have that

$$\begin{aligned}\bar{W}(\omega_x, \omega_y) &= \frac{\left|\bar{G}(\omega_x, \omega_y)\right|^2}{\left|\bar{G}(\omega_x, \omega_y)\right|^2 + \bar{P}_v(\omega_x, \omega_y)} \\ &= \frac{\frac{1}{NM}\left|\bar{G}(\omega_x, \omega_y)\right|^2}{\frac{1}{NM}\left|\bar{G}(\omega_x, \omega_y)\right|^2 + \frac{1}{NM}\bar{P}_v(\omega_x, \omega_y)} \\ &= \frac{\bar{A}(\omega_x, \omega_y)}{\bar{A}(\omega_x, \omega_y) + \bar{B}(\omega_x, \omega_y)}\end{aligned} \tag{17}$$

where NM is the total number of points in the input data grid. The two auxiliary functions, $\bar{A}(\omega_x, \omega_y)$ and $\bar{B}(\omega_x, \omega_y)$, in the last equation correspond to the Fourier transforms of two associated sequences, $a(nh_x, mh_y)$ and $b(nh_x, mh_y)$, that have the CV-like expressions (for a proof, see the Appendix)

$$a(nh_x, mh_y) = \frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} g(kh_x, lh_y)\, g\big((k+n)h_x, (l+m)h_y\big) \tag{18a}$$

and

$$b(nh_x, mh_y) = \frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} \sigma_v\big[(kh_x, lh_y)\big((k+n)h_x, (l+m)h_y\big)\big] \tag{18b}$$

respectively. The first sequence in Eq. (18a) can easily be identified as the discrete spatial CV function of the true deterministic signal at the given data resolution, and thus the term $\bar{A}(\omega_x, \omega_y)$ in Eq. (17) corresponds to the power spectrum of the true signal values $g(nh_x, mh_y)$. Note that the sequence $a(nh_x, mh_y)$ contains less spatio-statistical information than the continuous signal CV function, since it takes into account only the discrete values of the unknown field at a certain resolution level. The knowledge of the continuous spatial CV function of the true signal is only needed in the optimization of the reference interpolator $\varphi_h(x, y)$, as explained in Kotsakis (2000a, b).

The second sequence in Eq. (18b), on the other hand, does not exactly correspond to the discrete noise CV function and, as a result, the frequency-domain quantity $\bar{B}(\omega_x, \omega_y)$ in Eq. (17) should not generally be viewed as the data noise PSD. Such an interpretation is possible only in the special case where the input noise is stationary. Indeed, in such a situation the noise covariance σ_v between two arbitrary data points with coordinates (kh_x, lh_y) and $((k+n)h_x, (l+m)h_y)$ becomes a function of their coordinate differences only, which are obviously equal to (nh_x, mh_y). Therefore, σ_v can be taken outside of the summation operator in Eq. (18b), leaving the summation result equal to NM. In the more general case of non-stationary noise, the sequence $b(nh_x, mh_y)$ can be interpreted as a 'mean' CV function of the random observation errors. Its value at the origin gives an average indication of the noise level at every point of the data grid, i.e.

$$\begin{aligned}b(0,0) &= \frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} \sigma_v\big[(kh_x, lh_y)(kh_x, lh_y)\big] \\ &= \frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} \sigma_v^2(kh_x, lh_y)\end{aligned} \tag{19}$$

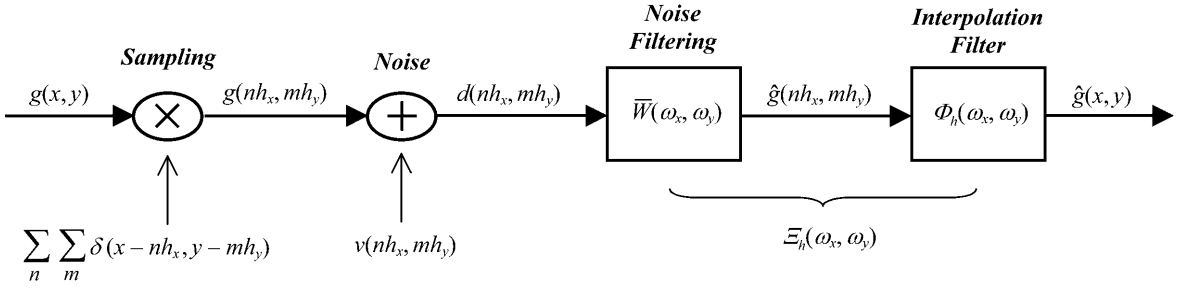

Fig. 3. The cascading structure of the optimal linear estimation filter

whereas its values $b(nh_x, mh_y)$ at other points correspond to 'averages' of the noise covariance over pairs of data points with coordinate differences (nh_x, mh_y).

Note that both sequences in Eqs. (18a) and (18b) are always symmetric, and they take zero values outside the range $-(N-1) \le n \le (N-1)$ and $-(M-1) \le m \le (M-1)$ due to the finite support of the input signal $g(x,y)$. Also, the numerical evaluation of the optimal noise filter $\bar{W}(\omega_x, \omega_y)$ in practice can take place only at discrete frequency values using the discrete Fourier transforms (DFTs) of the two CV-type sequences, $a(nh_x, mh_y)$ and $b(nh_x, mh_y)$.

5 Additional remarks

5.1 Why not 'total' MSE optimization?

A key point in the development of our estimation procedure was the decomposition of the output signal error into an aliasing (deterministic) part and a noise-dependent (stochastic) part; see Sect. 3. The advantage of such a partition is that it allows us to study and optimize individually the effects of the finite data resolution and the input noise on the signal estimate $\hat{g}(x,y)$, using appropriate error measures and criteria for each case. It should be kept in mind that the aliasing error component $e_h(x,y)$ is a purely deterministic signal whose average behaviour (at a given data resolution level) can be modelled in a spatio-statistical sense through the concept of different 'sampling phases' (Kotsakis 2000b; Kotsakis and Sideris 2000), whereas the noise-dependent error term $e_v(x,y)$ is a random signal whose average behaviour is described probabilistically with the notion of different 'experiment repetitions' (expectation operator E).

At this point we could naturally ask: why can't we just employ a single optimal estimation criterion for the total signal error produced by the linear approximation formula in Eq. (7)? We could have used, for example, the following 'global' MSE estimation principle:

$$P_e(\omega_x, \omega_y) = E\left\{ \left| E(\omega_x, \omega_y) \right|^2 \right\} = \text{minimum} \tag{20}$$

where $E(\omega_x, \omega_y)$ is the Fourier transform of the total error $e(x,y)$ given in Eq. (8), and $P_e(\omega_x, \omega_y)$ corresponds to the mean error power spectrum of the signal estimate $\hat{g}(x,y)$. In such a case, it is easy to show that the corresponding optimal estimation kernel will have the frequency-domain form

$$\Xi_h(\omega_x, \omega_y) = \frac{G(\omega_x, \omega_y)\bar{G}^*(\omega_x, \omega_y)}{\left|\bar{G}(\omega_x, \omega_y)\right|^2 + \bar{P}_v(\omega_x, \omega_y)} \tag{21}$$

which is slightly different from the previous optimal filter solution given in Eq. (16). In fact, if we replace in the above equation the Fourier transform $G(\omega_x, \omega_y)$ of the true unknown signal with the Fourier transform $\tilde{G}(\omega_x, \omega_y)$ of a linear and shift-invariant interpolating model according to Eq. (9b), then we automatically obtain the separable filter solution of Eq. (16).

The result shown in Eq. (21) is practically useless since it requires not just the knowledge of the signal spatial CV function, but the complete pointwise knowledge of the unknown field $g(x,y)$ beforehand. It is actually quite instructive to evaluate the performance of the optimal estimation filter in Eq. (21) when the input data is noiseless. By applying the Fourier transform to the general estimation formula in Eq. (7), we obtain

$$\begin{aligned} \hat{G}(\omega_x, \omega_y) &= \Xi_h(\omega_x, \omega_y)\bar{D}(\omega_x, \omega_y) \\ &= \Xi_h(\omega_x, \omega_y)(\bar{G}(\omega_x, \omega_y) + \bar{V}(\omega_x, \omega_y)) \end{aligned} \tag{22}$$

In the absence of any noise from the data [i.e. $\bar{V}(\omega_x, \omega_y) = 0$, $\bar{P}_v(\omega_x, \omega_y) = 0$] the use of the estimation filter from Eq. (21) yields

$$\begin{aligned} \hat{G}(\omega_x, \omega_y) &= \frac{G(\omega_x, \omega_y)\bar{G}^*(\omega_x, \omega_y)}{\left|\bar{G}(\omega_x, \omega_y)\right|^2 + 0}\left(\bar{G}(\omega_x, \omega_y) + 0\right) \\ &= G(\omega_x, \omega_y) \end{aligned} \tag{23a}$$

which means that, *regardless of the sampling resolution levels* h_x and h_x, the filter given in Eq. (21) can always recover the full field from its noiseless discrete values! On the other hand, the use of the separable filter from Eq. (16) gives a more reasonable signal estimate in the case of noise-free data, as follows:

$$\begin{aligned} \hat{G}(\omega_x, \omega_y) &= \bar{W}(\omega_x, \omega_y)\Phi_h(\omega_x, \omega_y)(\bar{G}(\omega_x, \omega_y) + 0) \\ &= \frac{\left|\bar{G}(\omega_x, \omega_y)\right|^2}{\left|\bar{G}(\omega_x, \omega_y)\right|^2 + 0}\Phi_h(\omega_x, \omega_y)\bar{G}(\omega_x, \omega_y) \\ &= \Phi_h(\omega_x, \omega_y)\bar{G}(\omega_x, \omega_y) = \tilde{G}(\omega_x, \omega_y) \end{aligned} \tag{23b}$$

where $\tilde{G}(\omega_x, \omega_y)$ is the Fourier transform of the interpolated signal model $\tilde{g}(x,y)$ according to the formula given in Eq. (9a).

Thus, it is seen that the partition of the output estimation error $e(x,y)$ into a noise-free and a noise-dependent component is more than an arbitrary modelling choice. It is actually a reasonable step in order to obtain meaningful signal estimators that take into account both the limited data resolution and the external noise effects. It should be remembered that the reference interpolation filter $\Phi_h(\omega_x, \omega_y)$ can be optimized by following a separate approach that takes into account only the noise-free error component $e_h(x,y)$, as described in Kotsakis (2000a, b).

5.2 Practical realization of the noise filter $\bar{W}(\omega_x, \omega_y)$

An important aspect for geodetic applications is the numerical computation of the optimal noise filter $\bar{W}(\omega_x, \omega_y)$ given in Eq. (17). Although the noise 'PSD' term $\bar{P}_v(\omega_x, \omega_y)$ can always be determined from the known noise variances and covariances using fast Fourier transform (FFT) techniques (see Sect. 4), the signal 'PSD' term $|\bar{G}(\omega_x, \omega_y)|^2$ is generally unknown in practice. In order to overcome this difficulty, we can use the power spectrum of the available noisy data

$d(nh_x, nh_y)$ to infer the behaviour of the signal 'PSD' function. From Eq. (6), we can express the power spectrum of the data values as follows:

$$\begin{aligned}\left|\bar{D}(\omega_x,\omega_y)\right|^2 &= (\bar{G}(\omega_x,\omega_y)+\bar{V}(\omega_x,\omega_y))(\bar{G}(\omega_x,\omega_y)\\ &\quad +\bar{V}(\omega_x,\omega_y))^*\\ &= \left|\bar{G}(\omega_x,\omega_y)\right|^2+\bar{G}(\omega_x,\omega_y)\bar{V}^*(\omega_x,\omega_y)\\ &\quad +\bar{G}^*(\omega_x,\omega_y)\bar{V}(\omega_x,\omega_y)+\left|\bar{V}(\omega_x,\omega_y)\right|^2 \quad (24a)\end{aligned}$$

and by applying the expectation operator to the above formula, we finally obtain

$$\begin{aligned}E\left\{\left|\bar{D}(\omega_x,\omega_y)\right|^2\right\} &= \left|\bar{G}(\omega_x,\omega_y)\right|^2+E\left\{\left|\bar{V}(\omega_x,\omega_y)\right|^2\right\}\\ &= \left|\bar{G}(\omega_x,\omega_y)\right|^2+\bar{P}_v(\omega_x,\omega_y) \quad (24b)\end{aligned}$$

The unknown signal term $|\bar{G}(\omega_x,\omega_y)|^2$ can now be determined empirically through the last equation by taking the available realization $|\bar{D}(\omega_x,\omega_y)|^2$ of the data power spectrum as an estimate of its expected value.

5.3 Higher data resolution leads to noise reduction

The use of the optimal estimation filter $\Xi_h(\omega_x,\omega_y)$ according to Eq. (16) leads to the following expression for the mean power spectrum of the noise-dependent signal error $e_v(x,y)$:

$$\begin{aligned}P_{e_v}(\omega_x,\omega_y) &= \frac{\left|\bar{G}(\omega_x,\omega_y)\right|^2}{\left|\bar{G}(\omega_x,\omega_y)\right|^2+\bar{P}_v(\omega_x,\omega_y)}\\ &\quad \times \bar{P}_v(\omega_x,\omega_y)\left|\Phi_h(\omega_x,\omega_y)\right|^2\\ &= \bar{W}(\omega_x,\omega_y)\bar{P}_v(\omega_x,\omega_y)\left|\Phi_h(\omega_x,\omega_y)\right|^2 \quad (25)\end{aligned}$$

The above formula can easily be derived by substituting the optimal result of Eq. (16) into Eq. (14). As was mentioned at the beginning of Sect. 3, the noise-dependent estimation error is additionally affected by the actual data resolution level. In order to see more clearly this important implicit relationship, let us first adopt a rather general model for the reference interpolating kernel $\varphi_h(x,y)$. In particular, we will consider the case

$$\varphi_h(x,y) = \varphi\left(\frac{x}{h_x},\frac{y}{h_y}\right) \quad (26)$$

where $\varphi(x,y)$ is some basic scaling interpolator (e.g. *sinc function*). Taking into account the fundamental scaling property of the 2-D Fourier transform (Dudgeon and Mersereau 1984), the frequency-domain expression in Eq. (25) can now be written as

$$\begin{aligned}P_{e_v}(\omega_x,\omega_y) &= (h_x h_y)^2\bar{W}(\omega_x,\omega_y)\\ &\quad \times \bar{P}_v(\omega_x,\omega_y)\left|\Phi(h_x\omega_x,h_y\omega_y)\right|^2 \quad (27)\end{aligned}$$

where $\Phi(\omega_x,\omega_y)$ is the Fourier transform of the scaling function $\varphi(x,y)$. It is seen from the last equation that the noise-dependent estimation error will decrease as the resolution of the input data increases (i.e. smaller sampling intervals h_x and h_y). Such a result is not surprising and it just confirms the (already well known from signal analysis theory) fact that oversampling leads to noise reduction in the final signal estimate; see e.g. Benedetto (1998) and Cvetković and Vetterli (1998). An additional justification of this interesting property is given in the next section using simulated numerical data.

6 Numerical example

The purpose of this section is to test numerically the noise filtering component of the optimal estimation filter $\Xi_h(\omega_x,\omega_y)$ that was derived in Sect. 3; see Eq. (16). Hence, we will not implement the whole linear I/O system shown in Fig. 3, but will restrict our attention only to its first de-noising part that transforms the original input data into an improved filtered signal sequence. The second interpolatory step through the use of a reference modelling kernel $\varphi_h(x,y)$ will be ignored.

Although our theoretical developments in this paper have been based on a 2-D planar framework, the numerical results presented in this section refer to the simpler 1-D-estimation case. This kind of simplification was chosen only for the sake of easier visualization of the noise filter performance under non-stationary observation errors, and it does not restrict the validity of the conclusions for higher-dimensional problems. For some 2-D numerical examples, see Kotsakis and Sideris (2001) and Kotsakis (2001). The various steps and the results of our simulated numerical experiments are summarized in the following list.

- A 1-D deterministic signal $g(x)$, assumed to represent some gravity anomaly profile on the geoid, was initially synthesized using a truncated Fourier series expansion with a record length of 200 km (see Fig. 4).
- The above signal was sampled at various resolution levels (h_x) to obtain noiseless gridded values $g(nh_x)$. Four different sampling resolution levels were selected, namely 0.1, 0.5, 1 and 5 km.
- All signal grids $g(nh_x)$, at each resolution level, were partitioned into three equal spatial blocks, labelled as left (L), central (C) and right (R). The simulated data noise, which is subsequently added to the true signal values at the next step, will have different statistical behaviour in each of the three grid blocks.
- A zero-mean noise sequence $v(nh_x)$ was added to the samples of the true signal in order to create the input data set, according to the form $d(nh_x) = g(nh_x) + v(nh_x)$. The noise values originated from a non-stationary and uncorrelated Gaussian stochastic process, using the routines for random number generation from the MATLAB™ software package. Note that the noise values $v(nh_x)$ were generated separately at each resolution, instead of simply decimating the noise sequence with the smaller sampling interval.
- The noise variance $\sigma_v^2(nh_x)$ was constant within each block (L, C, R) of the data grid, with its values set to 20, 3 and 6 mGal^2, respectively. The sample statistics of the total noise sequence $v(nh_x)$ at every resolution

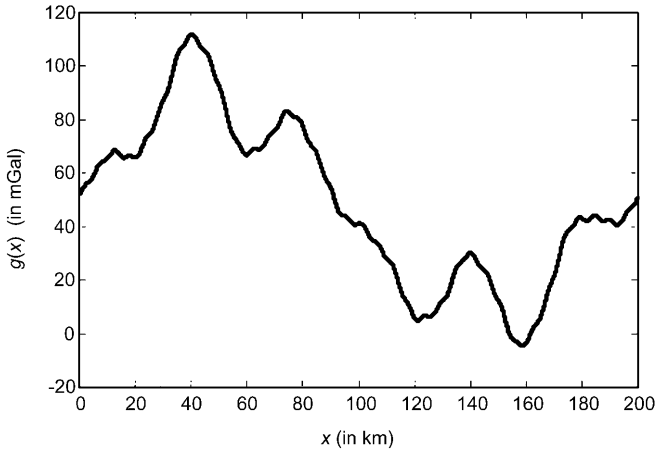

Fig. 4. The original (simulated) signal

level are given in Table 1, whereas the individual sample statistics of the noise values in the three different parts of the data grids are given in Table 2.

- The optimal noise filter $\bar{W}(\omega_x)$ was computed, for each resolution h_x, via an FFT algorithm according to the 1-D counterpart of the SNR expression in Eq. (17). It was then multiplied by the FFT of the noisy data $d(nh_x)$, and the result was finally transformed back to the space domain as an estimated ('de-noised') signal sequence $\hat{g}(nh_x)$. The original noisy data and the filtered signal values are plotted in Fig. 5 for some selective sampling resolution levels. The differences between the true signal samples and the estimated signal values are also shown in Fig. 6, whereas their statistics are given in Table 3.

It is interesting to observe that the estimation error of the filtered signal values $\hat{g}(nh_x)$ decreases as the data resolution increases. This is evident from the comparison of the three graphs shown in Fig. 6, as well as from the root-mean-square (RMS) error values given in Table 3, and it confirms our earlier theoretical remark at the end of Sect. 5. Note also that the part of the signal with the highest noise level (left grid block, $\sigma_v^2 = 20$ mGal^2) shows consistently larger errors after the data filtering than the other two grid blocks (see Fig. 6). This is a reasonable result since we should not expect to have the same quality level in the final signal estimate when using data with spatially varying accuracy.

The inhomogeneous character of the output (noise-dependent) estimation error can also be verified in a more rigorous analytical way. If the data noise is non-stationary then the input data values constitute a non-stationary random sequence as well, as it is easily seen from Eq. (3). As a result, the signal estimate obtained through the filtering formula in Eq. (7) becomes a non-stationary random process, and thus the noise-dependent estimation error $e_v(x, y)$ defined by Eq. (10a) will also correspond to a non-stationary random process. Note that the values plotted in Fig. 6 are basically the samples of this continuous (noise-dependent) signal estimation error taken at the data points.

Table 1. Sample statistics of the total additive noise at various data resolution levels (mGal)

Data resolution (km)	0.1	0.5	1.0	5.0
Max	14.57	15.34	8.91	8.27
Mean	0.09	0.24	0.27	0.11
Min	−15.09	−10.10	−11.98	−8.01
Std	3.15	3.08	3.09	3.01
RMS	3.15	3.09	3.10	3.01

7 Conclusions

A modification of the classic Wiener filtering method has been presented, which allows us to work with arbitrary deterministic signals that are masked by additive non-stationary noise at different sampling resolution levels. This provides a useful tool for many geodetic problems related to optimal spectral gravity field modelling, as discussed in the introductory section of this paper. It has been shown that non-stationary noise filtering using fast spectral Fourier techniques is possible, if we are willing to incorporate a simple shift-invariance condition into our signal approximation algorithm. In contrast to the W–K linear prediction theory for random fields, where the optimal estimators are translation-invariant only for stationary signals and noise, the methodology presented herein always uses convolution-type estimators, regardless of the statistical properties of the unknown quantities. The spatial resolution of the input data was also taken directly into account within the estimation algorithm, revealing some interesting aspects about the structure and the performance of the optimal noise filter as a function of the data grid density.

The informal algorithmic similarities of our signal approximation procedure with the Wiener filtering formalism stem from the initial assumptions in Eq. (7) of linearity and translation-invariance. These led to a convolution-based computational scheme in terms of an optimal SNR filter that is applied to the discrete input

Table 2. Sample statistics of the additive noise at various data resolution levels in the three different blocks [left (L), central (C) and right (R)] of the data grids (mGal)

Data resolution (km)	0.1			0.5			1.0			5.0		
	L	C	R	L	C	R	L	C	R	L	C	R
Max	14.57	5.06	8.96	15.34	4.12	5.67	8.91	5.58	6.70	8.27	2.60	4.01
Mean	0.22	0.02	0.04	0.79	−0.07	0.01	0.85	0.06	−0.11	−0.12	−0.16	0.66
Min	−15.09	−6.44	−7.58	−10.10	−3.94	−5.65	−11.98	−3.67	−5.75	−8.01	−2.16	−2.93
Std	4.53	1.70	2.51	4.38	1.79	2.34	4.32	1.74	2.53	4.31	1.46	2.40
RMS	4.54	1.70	2.51	4.45	1.79	2.34	4.41	1.74	2.53	4.31	1.47	2.49

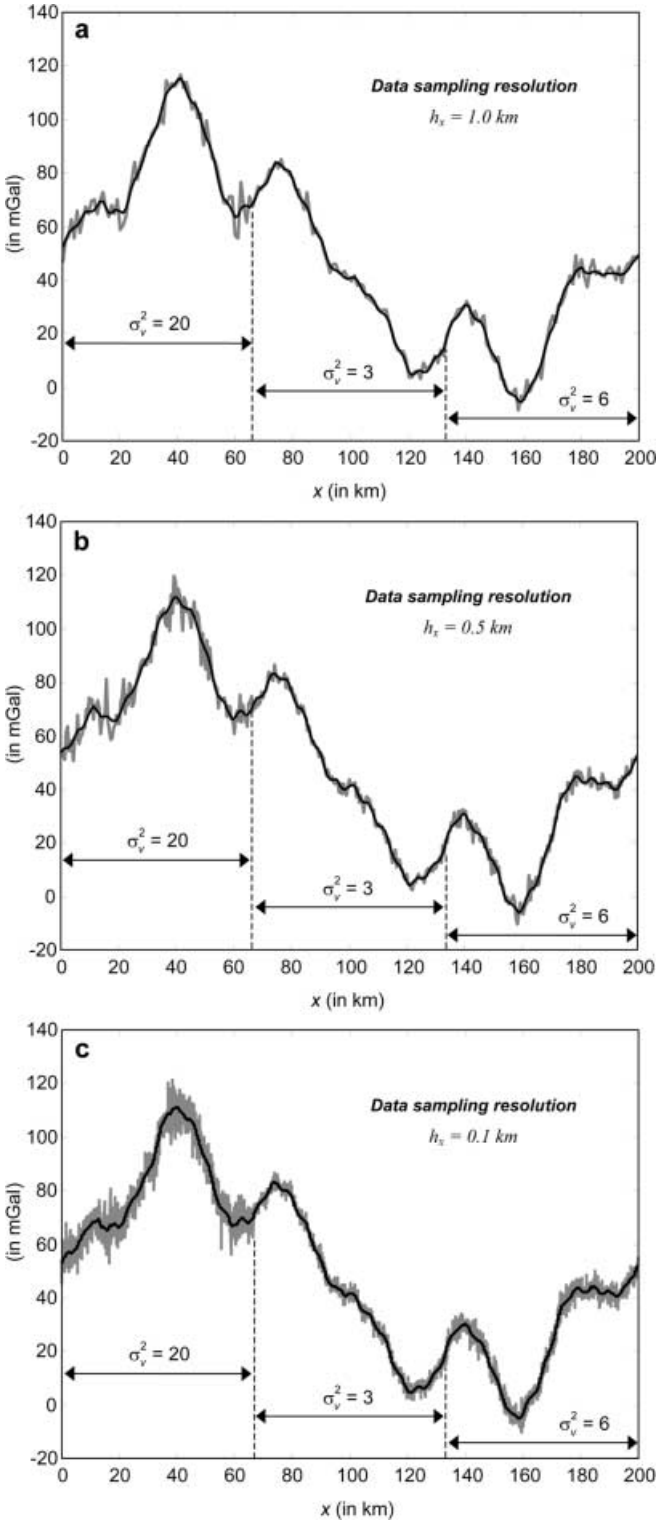

Fig. 5a–c. Noisy and filtered signal values at various sampling resolution levels. The vertical dashed lines indicate the boundaries between the three blocks (L, C, R) of the input data grid with the different noise variances

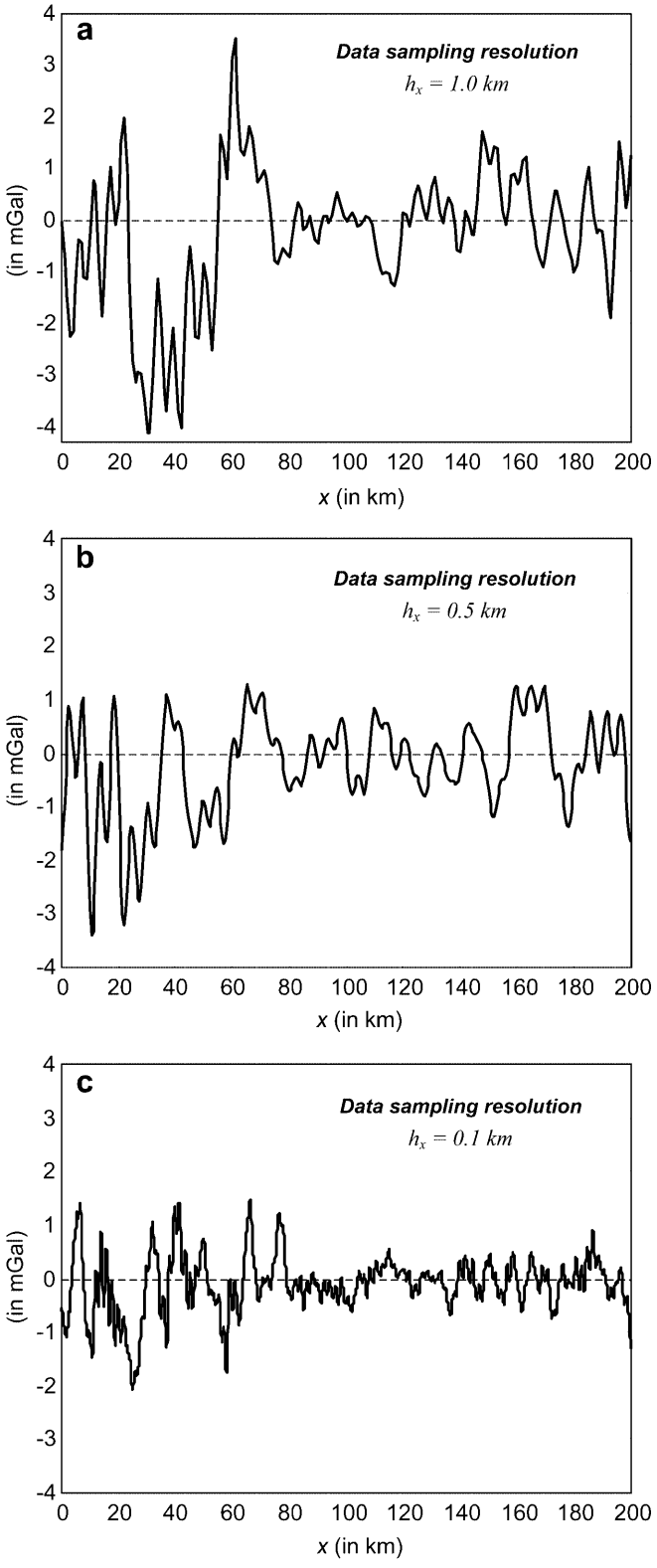

Fig. 6a–c. Differences between the filtered and the true signal values at various data resolution levels

data. On the other hand, the differences with the classic Wiener filtering theory are due to the modelling assumptions for the input signal and noise, which are not generally treated as stationary stochastic processes in our case. The unknown field, in particular, is not associated with any stochastic behaviour but is treated as a finite-energy finite-support deterministic signal. It is important to keep in mind that when the data noise is

Table 3. Sample statistics of the differences between the true and the filtered signal values at various data resolution levels (mGal)

Data resolution (km)	0.1	0.5	1.0	5.0
Max	1.47	1.28	3.50	6.48
Mean	−0.09	−0.24	−0.26	−0.11
Min	−2.06	−3.38	−4.15	−6.47
Std	0.54	0.91	1.30	2.53
RMS	0.55	0.94	1.33	2.54

not stationary, the W–K linear prediction theory can no longer lead to simple filtering/convolution operations. In such cases, the estimation algorithm is reduced to a Fredholm equation of the first kind (*Wiener–Hopf equation*), whose solution determines the best linear (but *not* translation-invariant) signal estimate in a probabilistic MSE sense; for more details, see Sansò and Sideris (1997). In our approach, the a priori imposed condition of translation-invariance allows us to treat both stationary and non-stationary noise cases within a unified linear filtering setting, which can be implemented in practice very efficiently via FFT techniques.

In terms of future work, our efforts should concentrate on extending the 2-D planar algorithms given in this paper for signal approximation problems on the sphere and/or on the ellipsoid. Additional modifications are also needed in order to handle applications that involve more than one type of input data (multiple-input/single-output systems), and not just the single-input/single-output case that was studied here. Nevertheless, the presented methodology can be proven to be a useful tool for many existing geodetic estimation problems of regional or local scale. Some examples of such problems include the optimal spectral geoid determination from noisy gridded gravity data, and the FFT computation of various terrain-dependent gravity field quantities (e.g. terrain correction, indirect effect, isostatic potential, etc.) from noisy digital elevation models.

Acknowledgements. This work was carried out during the first author's stay at the Department of Geomatics Engineering, University of Calgary, as a post-doctoral fellow. The financial support obtained through this post-doctoral fellowship is gratefully acknowledged. Research funding for this work has been provided through research grants awarded to the second author by the Natural Sciences and Engineering Research Council (NSERC) of Canada and by the Geomatics for Informed Decisions (GEOIDE) network of Centres of Excellence in Canada. Profs. F. Sansò, W. Keller and an anonymous reviewer are acknowledged for their valuable comments and criticism on the initial draft of the paper.

Appendix

We will prove that the 2-D Fourier transforms of the following sequences:

$$a(nh_x, mh_y) = \frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} g\big(kh_x, lh_y\big) \times g\big((k+n)h_x, (l+m)h_y\big) \tag{A1}$$

and

$$b(nh_x, mh_y) = \frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} \sigma_v[\big(kh_x, lh_y\big)\big((k+n)h_x, (l+m)h_y\big)] \tag{A2}$$

are given by the terms $(NM)^{-1}|\bar{G}(\omega_x, \omega_y)|^2$ and $(NM)^{-1}\bar{P}_v(\omega_x, \omega_y)$, respectively, that appear in the SNR expression of the optimal noise filter in Eq. (17). Let us recall that the unknown signal $g(x, y)$ is assumed to have compact support over the real plane, and thus both the signal values $g(kh_x, lh_y)$ and the measurement noise values $v(kh_x, lh_y)$ are always zero outside the input data grid (NM points); see Sect. 2.1.

The 2-D Fourier transform $\bar{A}(\omega_x, \omega_y)$ of the first sequence in Eq. (A1) can be obtained as follows:

$$\begin{aligned}
\bar{A}(\omega_x, \omega_y) &= \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} a(nh_x, mh_y)\mathrm{e}^{-i(nh_x\omega_x + mh_y\omega_y)} \\
&= \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}\left[\frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} g\big(kh_x, lh_y\big) \times g\big((k+n)h_x, (l+m)h_y\big)\right]\mathrm{e}^{-i(nh_x\omega_x + mh_y\omega_y)} \\
&= \frac{1}{NM}\sum_{p=-\infty}^{+\infty}\sum_{r=-\infty}^{+\infty}\left[\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} g(kh_x, lh_y)g(ph_x, rh_y)\right] \times \mathrm{e}^{-i\big((p-k)h_x\omega_x + (r-l)h_y\omega_y\big)} \\
&= \frac{1}{NM}\left[\sum_{p=-\infty}^{+\infty}\sum_{r=-\infty}^{+\infty} g(ph_x, rh_y)\mathrm{e}^{-i(ph_x\omega_x + rh_y\omega_y)}\right] \times \left[\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} g(kh_x, lh_y)\mathrm{e}^{i(kh_x\omega_x + lh_y\omega_y)}\right] \\
&= \frac{1}{NM}\left[\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} g(nh_x, mh_y)\mathrm{e}^{-i(nh_x\omega_x + mh_y\omega_y)}\right] \times \left[\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} g(kh_x, lh_y)\mathrm{e}^{i(kh_x\omega_x + lh_y\omega_y)}\right] \\
&= \frac{1}{NM}\left[\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} g(nh_x, mh_y)\mathrm{e}^{-i(nh_x\omega_x + mh_y\omega_y)}\right] \times \left[\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} g(kh_x, lh_y)\mathrm{e}^{i(kh_x\omega_x + lh_y\omega_y)}\right] \\
&= \frac{1}{NM}\bar{G}(\omega_x, \omega_y)\bar{G}^*(\omega_x, \omega_y) = \frac{1}{NM}\big|\bar{G}(\omega_x, \omega_y)\big|^2
\end{aligned} \tag{A3}$$

In the previous derivations we have used the notation according to Eq. (2) in Sect. 2.1 of this paper, and the assumption that the unknown signal is always real. Similarly, the 2-D Fourier transform $\bar{B}(\omega_x, \omega_y)$ of the second sequence given in Eq. (A2) can be expressed as follows:

$$
\begin{aligned}
\bar{B}(\omega_x,\omega_y) &= \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} b(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)} \\
&= \sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}\left[\frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1}\sigma_v\big[(kh_x,lh_y)\right. \\
&\quad \left.\big((k+n)h_x,(l+m)h_y\big)\big]\right]\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)} \\
&= \frac{1}{NM}\sum_{p=-\infty}^{+\infty}\sum_{r=-\infty}^{+\infty}\left[\sum_{k=0}^{N-1}\sum_{l=0}^{M-1}\sigma_v\big[(kh_x,lh_y)\right. \\
&\quad \left.(ph_x,rh_y)\big]\right]\mathrm{e}^{-i\big((p-k)h_x\omega_x+(r-l)h_y\omega_y\big)} \\
&= \frac{1}{NM}\sum_{p=-\infty}^{+\infty}\sum_{r=-\infty}^{+\infty}\left[\sum_{k=0}^{N-1}\sum_{l=0}^{M-1}E\big\{v(kh_x,lh_y)\right. \\
&\quad \left.\times v(ph_x,rh_y)\big\}\right]\mathrm{e}^{-i\big((p-k)h_x\omega_x+(r-l)h_y\omega_y\big)} \\
&= \frac{1}{NM}E\left\{\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1}v(kh_x,lh_y)\right. \\
&\quad \left.\times v(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)}\mathrm{e}^{i(kh_x\omega_x+lh_y\omega_y)}\right\} \\
&= \frac{1}{NM}E\left\{\left[\sum_{n=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty}v(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)}\right]\right. \\
&\quad \left.\times\left[\sum_{k=0}^{N-1}\sum_{l=0}^{M-1}v(kh_x,lh_y)\mathrm{e}^{i(kh_x\omega_x+lh_y\omega_y)}\right]\right\} \\
&= \frac{1}{NM}E\left\{\left[\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}v(nh_x,mh_y)\mathrm{e}^{-i(nh_x\omega_x+mh_y\omega_y)}\right]\right. \\
&\quad \left.\times\left[\sum_{k=0}^{N-1}\sum_{l=0}^{M-1}v(kh_x,lh_y)\mathrm{e}^{i(kh_x\omega_x+lh_y\omega_y)}\right]\right\} \\
&= \frac{1}{NM}E\big\{\bar{V}(\omega_x,\omega_y)\bar{V}^*(\omega_x,\omega_y)\big\} \\
&= \frac{1}{NM}E\Big\{\big|\bar{V}(\omega_x,\omega_y)\big|^2\Big\} = \frac{1}{NM}\bar{P}_v(\omega_x,\omega_y) \qquad \text{(A4)}
\end{aligned}
$$

where the frequency-domain quantities $\bar{V}(\omega_x,\omega_y)$ and $\bar{P}_v(\omega_x,\omega_y)$ have been defined previously in Eqs. (5) and (15), respectively.

References

Bendat JS, Piersol AG (1993) Engineering applications of correlation and spectral analysis. John Wiley, New York

Benedetto JJ (1998) Noise reduction in terms of the theory of frames. In: Zeevi Y, Coifman R (eds) Signal and image representation in combined spaces. Academic Press, New York, pp 259–284

Blu T, Unser M (1999) Quantitative Fourier analysis of approximation techniques: part I (Interpolators and projectors) and part II (Wavelets). IEEE Trans Sig Proc 47(10): 2783–2806

Brovelli M, Sansò F, Venuti G (2001) A discussion on the Wiener–Kolmogorov prediction principle with easy to compute and robust variants. Preprint

Cvetković Z, Vetterli M (1998) Overcomplete expansions and robustness. In: Zeevi Y, Coifman R (eds) Signal and image representation in combined spaces. Academic Press, New York, pp 301–338

Dudgeon DE, Mersereau RM (1984) Multidimensional digital signal processing. Prentice-Hall, Englewood Cliffs

Gasquet C, Witomski P (1999) Fourier analysis and applications. Springer, Berlin Heidelberg New York

Kailath T (1974) A view of three decades of linear filtering theory. IEEE Trans Info Theor IT-20(2): 146–181

Kolmogorov AN (1941) Interpolation and extrapolation of stationary random sequences. Bull Acad Sci USSR, Ser Math, vol 5 (Translated by Doyle W, Selin I, RAND Corporation, Santa Monica, CA, Memo RM-3090-PR, April 1962)

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

Kotsakis C (2000b) Multiresolution aspects of linear approximation methods in Hilbert spaces using gridded data. UCGE rep 20138, Department of Geomatics Engineering, University of Calgary

Kotsakis C (2001) Can we filter non-stationary noise from geodetic data with fast spectral (FFT) techniques? IGeS Bull 11: 93–114

Kotsakis C, Sideris MG (2000) Aliasing error modelling in single-input single-output linear estimation systems. Paper presented at the IAG Int Symp Gravity, Geoid and Geodynamics, Banff, 31 July–4 August

Kotsakis C, Sideris MG (2001) Non-stationary noise filtering of gravity data using fast spectral techniques. Paper presented at the XXVI EGS General Assembly, Nice, 26–30 March

Li J (1996) Detailed marine gravity field determination by combination of heterogeneous data. UCGE rep 20102, Department of Geomatics Engineering, University of Calgary

Li J, Sideris MG (1997) Marine gravity and geoid determination by optimal combination of satellite altimetry and shipborne gravimetry data. J Geod 71(4): 209–216

Li YC, Sideris MG (1994) Minimization and estimation of geoid undulation errors. Bull Geod 68: 201–219

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

Moritz H, Sansò F (1980) A dialogue on collocation. Boll Geod Sci Affi 39(1): 49–51

Pawlowski RS, Hansen RO (1990) Gravity anomaly separation by Wiener filtering. Geoph 55(5): 539–548

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

Sansò 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

Sansò F, Sona G (1995) The theory of optimal linear estimation for continuous fields of measurements. Manuscr Geod 20: 204–230

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

Sideris MG (1996) On the use of heterogeneous noisy data in spectral gravity field modelling methods. J Geod 70(8): 470–479

Thévenaz P, Blu T, Unser M (2000) Interpolation revisited. IEEE Trans Med Imag 19(7): 739–758

Tziavos IN, Sideris MG, Li J (1996) Optimal spectral combination of satellite altimetry and marine gravity data. Presented at the XXI EGS General Assembly, The Hague, 6–10 May. Published in Finnish Geod Inst Rep 96(2): 41–56

Tziavos IN, Forsberg R, Sideris MG (1998) Marine gravity field modelling using shipborne gravity and geodetic mission altimetry data. Geom Res Austr 69: 1–18

Unser M (2000) Sampling – 50 years after Shannon. IEEE Proc 88(4): 569–587

Unser M, Daubechies I (1997) On the approximation power of convolution-based least squares versus interpolation. IEEE Trans Sig Proc 47(7): 1697–1711

Vassiliou AA (1986) Numerical techniques for processing airborne gradiometer data. UCGE rep 20017, Department of Geomatics Engineering, University of Calgary

Wiener N (1949) Extrapolation, interpolation and smoothing of stationary time series with engineering applications. John Wiley, New York (originally issued in February 1942 as a classified US National Defense Research Council Report)

Wu L, Sideris MG (1995) Using multiple-input single-output system relationships in post processing of airborne gravity vector data. In: Schwarz KP (ed) Airborne gravimetry, Proc IAG Symp Airborne Gravity Field Determination, IUGG XXI General Assembly, Boulder, CO, 2–14 July, pp 87–94