Wiener Filter Modifications for Gravity Field Data Using Different Resolution Levels and Non-Stationary Noise

C. Kotsakis, F. A. Bayoud, M.G. Sideris
Department of Geomatics Engineering, University of Calgary
2500 University Drive N.W., Calgary, Alberta, Canada, T2N 1N4, Email: ckotsaki@ucalgary.ca

Abstract. One of the most basic tools in optimal spectral gravity field modelling is the method of Wiener filtering. Originally developed for applications in analog 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 drawbacks of Wiener filtering that make its use in certain geodetic estimation applications problematic stem from the stationarity assumption for both the signal and the noise involved in the approximation problem. In this paper, we introduce a modified Wiener-type linear estimation filter that can be used with noisy data obtained from an arbitrary deterministic field under the masking of generally non-stationary additive noise. In addition, the sampling resolution of the input data is explicitly taken into account within our estimation algorithm, resulting in a resolution-dependent optimal noise filter.

1 Introduction – Problem statement

Wiener filtering is a well-established and efficient optimal estimation method that can be used for geodetic data ‘de-noising’ in the frequency domain. Its theoretical framework is based on the famous Wiener-Kolmogorov linear prediction theory for stationary random fields, in the presence of stationary additive noise (Kailath 1974, Priestley 1981). The application of the Wiener filter in geodesy, both as an independent practical tool for data pre-processing and as an integral part of a more general linear estimation methodology (input-output systems theory, Sideris 1996), has been primarily focused on problems related to optimal spectral gravity field modelling. Many numerical 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 geological features (Pawlowski and Hansen 1990), 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 theoretical discussion on the use of the Wiener-Kolmogorov linear filtering theory in gravity field estimation, and its relationship with other linear approximation techniques traditionally used in geodesy (i.e. least-squares collocation), can be found in Sansò and Sideris (1997); see also Sideris (1996).

In order to employ the classical Wiener spectral filtering algorithm with noisy geodetic data, a stationarity assumption has to be made for both the true (unknown) signal and the random measurement errors. Such a modelling choice becomes quite problematic and unrealistic for many gravity field applications, since the underlying true signals (e.g. gravity anomaly) cannot admit a stochastic/probabilistic interpretation (and thus the stationarity property becomes meaningless in this case), and the additive data noise does not usually follow a spatially uniform statistical behaviour.

In this paper, we present a modified mean-square-error (MSE) spectral optimization procedure which, in conjunction with a certain translation-invariance condition, leads to a linear estimation filter that can be applied to an arbitrary finite-energy deterministic field with compact support (such as a gravity anomaly signal or the Earth’s topography signal) under the masking of generally non-stationary additive noise. Furthermore, the sampling resolution of the input data is explicitly taken into account within our estimation algorithm, resulting in a resolution-dependent
optimal noise filter. This provides a more ‘operational-friendly’ approach to the Wiener filtering spectral technique for geodetic applications, since the data resolution parameter has not been directly incorporated in previous formulations of frequency-domain estimation problems for gravity field signals. Due to space limitations and for the sake of notation economy, all the following developments will be restricted in a simple one-dimensional (1D) framework. Two-dimensional planar generalizations are quite straightforward and they will soon be published elsewhere.

2 Methodology – Problem formulation

The main problem that is studied in this paper is the frequency-domain estimation of an unknown deterministic field \( g(x) \in L^2(\mathbb{R}) \) using its noisy gridded samples \( d(nh) \in \mathbb{Z} \) according to the observation equation

\[
d(nh) = g(nh) + v(nh)
\]

where \( g(nh) \) are the true signal values, \( v(nh) \) is a random non-stationary noise sequence, and \( h \) corresponds to the data sampling resolution. The unknown signal will be assumed to have a compact spatial support over the real line, covering the interval \( 0 \leq x \leq X \). In this way, the integer sampling index \( n \) in Eq. (1) can be essentially restricted within the finite range \( 0 \leq n \leq (N-1) \), where \( X = (N-1)h \).

The associated stochastic model used to describe the behaviour of the measurement random errors, in terms of second-order moment information, is defined by the equations

\[
E\{v(nh)\} = 0 \quad (2a)
\]

\[
E\{v^2(nh)\} = \sigma^2_v(nh) = \sigma_v(nh,nh) \quad (2b)
\]

\[
E\{v(nh)\ v(mh)\} = \sigma_v(nh,mh) \quad (2c)
\]

where \( E \) is the probabilistic expectation operator. The symbol \( \sigma^2_v(\cdot) \) denotes the noise variance at a specific data point, whereas \( \sigma_v(\cdot,\cdot) \) corresponds to the noise covariance (CV) between two data points. The Fourier transform of the noiseless signal grid will be denoted by \( \mathcal{G}(\omega) \) and it is given by the formula (Oppenheim and Schafer 1989)

\[
\mathcal{G}(\omega) = \sum_{n=0}^{N-1} g(nh) e^{-in\omega h}
\]

where the overbar symbol will generally be used to indicate a periodic function. We will also use the notation \( \mathcal{V}(\omega) \) for the Fourier transform of the input data noise, which is defined as follows:

\[
\mathcal{V}(\omega) = \sum_{n=0}^{N-1} v(nh) e^{-in\omega h}
\]

Similarly, the Fourier transform of the gridded data values \( d(nh) \) is given by the equation

\[
\mathcal{D}(\omega) = \sum_{n=0}^{N-1} d(nh) e^{-in\omega h} = \mathcal{G}(\omega) + \mathcal{V}(\omega)
\]

Note that the noise signal is zero everywhere outside the input data grid \((N \text{ points})\), since the underlying unknown field \( g(x) \) has been assumed to have compact spatial support and thus no measurements are performed outside this region.

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 \( \hat{g}(x) \) that is independent of the reference system used to describe the physical position of the data points. Stated in a simplified way, if we change the origin of the coordinate system on the real line by some arbitrary translation \( t_x \) (without ‘moving’ the unknown field or the associated data grid), we want the new signal approximation to be just a translated version \( \hat{g}(x + t_x) \) of the initial estimate in the original reference system. The justification of such a modelling choice for spatial estimation problems relies basically on our logic and mathematical intuition, and it is not affected by the physical properties of the true signal and noise involved in the approximation procedure. If one chooses to follow a non translation-invariant methodology, he 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 the discrete input data. Note that the translation-invariance condition has often been applied in the theoretical formulation of optimal estimation methods using errorless discrete data (Sansó 1980, Kotsakis 2000a), although its justification is not altered by the presence of (stationary or non-stationary) noise in the input observations.
Based on the two assumptions of linearity and translation-invariance, the signal estimation formula will have the general convolution-type expression

$$\tilde{g}(x) = \sum_{n=0}^{N-1} d(nh) \xi_h(x - nh)$$

(6)

where $\xi_h(x)$ is an unknown 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 level. The last equation can be illustrated through the linear system shown in Fig. 1.

3 Estimation kernel optimization

The output signal error produced by the filtering formula in Eq. (6) can be decomposed into two distinct components, i.e.

$$e(x) = g(x) - \tilde{g}(x) = e_h(x) + e_v(x)$$

(7)

where $e_h(x)$ is the part of the total estimation error caused from the use of discrete data with finite sampling resolution (aliasing error), and $e_v(x)$ is the additional part due to the noise presence in the signal samples.

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

$$\tilde{g}(x) = \sum_{n=0}^{N-1} g(nh) \phi_h(x - nh)$$

(8a)

or, equivalently, in the frequency domain

$$\tilde{G}(\omega) = \Phi_h(\omega) \sum_{n=0}^{N-1} g(nh) e^{-in\omega}$$

$$= \Phi_h(\omega) \overline{G}(\omega)$$

(8b)

where $\phi_h(x)$ is some interpolating/modelling kernel that generally depends on the sampling interval $h$. The noise-dependent estimation error will be measured with respect to such a linear interpolating model for the unknown field, i.e.

$$e_v(x) = \tilde{g}(x) - \tilde{g}(x)$$

(9a)

whereas the (pure) aliasing error is

$$e_h(x) = g(x) - \tilde{g}(x)$$

(9b)

The actual form of the modelling kernel $\phi_h(x)$ in Eq. (8a) is irrelevant for the purpose of this paper. A very popular choice that covers many different linear interpolating (or quasi-interpolating) schemes, including band-limited (Shannon) interpolation, spline-based interpolation and also more general wavelet approximation models, is based on the use of certain scaling functions $\phi(x)$ which adapt to the data grid resolution through a dilation operation (see, e.g., Unser 2000), i.e.

$$\phi_h(x) = \phi\left(\frac{x}{h}\right)$$

(10)

The optimal determination, the intrinsic properties and the connection of such interpolating scaling kernels with the statistical collocation framework are discussed in detail in Kotsakis (2000a, b). For the purpose of this paper, it is sufficient to consider $\phi_h(x)$ in Eq. (8a) as an arbitrarily chosen interpolating kernel with a well-defined Fourier transform $\Phi_h(\omega)$, that is used to obtain a continuous signal approximation in the absence of any noise from the discrete input data.

![Fig. 1](image-url)  
Fig. 1 Linear and translation-invariant signal estimation using discrete noisy data.
In addition, it can be assumed that \( \phi_h(x) \) is such that: (i) the signal expansion in Eq. (8a) is always stable, and (ii) the aliasing error component \( e_h(x) \) vanishes as the data resolution increases; for more details, see Kotsakis (2000b) and Unser (2000).

The unknown filtering kernel \( \xi_h(x) \) in Eq. (6) will now be determined by minimizing the noise-dependent part \( e_v(x) \) of the total signal error. The optimization procedure will be carried out exclusively in the frequency domain using the familiar mean-square-error (MSE) criterion

\[
P_{e_v}(\omega) = E\left\{|E_v(\omega)|^2\right\} = \text{minimum} \quad (11)
\]

where \( E_v(\omega) \) is the Fourier transform of \( e_v(x) \), and \( P_{e_v}(\omega) \) is the (noise-dependent) mean error power spectrum of the estimated 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’ interpolating model \( \phi_h(x) \) where the MSE is defined in a spatio-statistical deterministic sense. In Kotsakis (2000a, b) the optimal determination of the modelling kernel \( \phi_h(x) \) was based on the spatio-statistical power spectrum of the deterministic error component \( e_h(x) \), whereas here the optimization of the noise filtering kernel \( \xi_h(x) \) employs the mean power spectrum of the stochastic error component \( e_v(x) \).

Using the previous equations, it is easy to show that the (noise-dependent) mean error power spectrum of the output signal is given by the formula

\[
P_{e_v}(\omega) = \Phi_h(\omega)\Phi^*_h(\omega)\left|\mathcal{F}(\omega)\right|^2
\]

\[
- \Phi_h(\omega)\Xi^*_h(\omega)\left|\mathcal{F}(\omega)\right|^2
\]

\[
+ \Xi_h(\omega)\Xi^*_h(\omega)\left|\mathcal{F}(\omega)\right|^2
\]

\[
\Xi_h(\omega) = E\left\{|\mathcal{F}(\omega)|^2\right\} - E\left\{|\mathcal{F}(\omega)|^2\right\}
\]

\[
= \text{minimum} \quad (12)
\]

The auxiliary term \( \Xi_h(\omega) \) in the last equation corresponds to the quantity

\[
\Xi_h(\omega) = E\left\{|\mathcal{F}(\omega)|^2\right\} = E\left\{|\mathcal{F}(\omega)|^2\right\}
\]

For the derivation of the result in Eq. (12) we have used the fact that \( E\{\mathcal{F}(\omega)\} = 0 \), in accordance with the zero-mean stochastic model introduced for the data noise in Eq. (2a); for more details, see Kotsakis (2000b). The optimal estimation filter \( \Xi_h(\omega) \) can now be determined using Eqs. (11) and (12). The underlying procedure is straightforward and it gives the final result

\[
\Xi_h(\omega) = \frac{\left|\mathcal{F}(\omega)\right|^2}{\left|\mathcal{F}(\omega)\right|^2 + \Phi_h(\omega)} \Phi_h(\omega) = \mathcal{F}(\omega)\Phi_h(\omega) \quad (14)
\]

In the next section, the separable Wiener-like form of the above filter is explained in detail.

### 4 The separable structure of the optimal estimation filter

The final result in Eq. (14) indicates that the optimal estimation procedure can be decomposed into two individual steps (filters) which are connected in a linear cascading manner. The first step, expressed by the periodic filter component \( \mathcal{F}(\omega) \), has the role of ‘de-noising’ the discrete input data using information about the average behaviour of the input noise and the unknown field at the given resolution level. The second filter component \( \Phi_h(\omega) \), on the other hand, is solely used to obtain a continuous representation for the output signal \( \hat{g}(x) \), based on an a-priori selected interpolating/modelling kernel \( \phi_h(x) \). These two basic steps of the optimal estimation procedure are illustrated in the linear system of Fig. 2. Note that, even though the optimal estimation principle was applied to the continuous (noise-dependent) signal error \( e_v(x) \), the noise filtering part of the linear estimation algorithm takes place at a discrete level and it is not affected by the choice of the ‘reference’ interpolating model.

As it can be seen from Fig. 2, it is not really necessary to modify the interpolating kernel \( \phi_h(x) \) of the reference signal model in Eq. (8a) when dealing with noisy input data. The optimization of the noise-dependent output error adds only an intermediate periodic filter that is applied to the original data grid \( d(nh) \) and it produces a new estimated signal sequence \( \hat{g}(nh) \) in which the effect of the random observational errors has been minimized in a certain translation-invariant linear fashion. We can then use this ‘synthetic’ grid as input to the interpolating model of Eq. (8a), in order to get a continuous (also linear and translation-invariant) approximation of the unknown field at the given resolution level.
The structure of the optimal filter in Eq. (14) is very similar to the classic Wiener estimation filter (i.e. they are both defined in terms of a certain signal-to-noise ratio (SNR) expression). However, there do exist conceptual differences between the two filtering schemes, since 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 in this case what is the exact meaning of the two frequency-domain terms that appear in the expression of our SNR-type optimal noise filter $\mathcal{W}(\omega)$. From Eq. (14), we have that

$$\mathcal{W}(\omega) = \frac{1}{N} \left[ \frac{1}{N} \left[ \mathcal{G}(\omega) \right]^2 + \frac{1}{N} \mathcal{R}(\omega) \right]$$

(15)

where $N$ is the total number of points in the input data grid. The auxiliary functions, $\mathcal{A}(\omega)$ and $\mathcal{B}(\omega)$, in the last equation correspond to the Fourier transforms of two associated sequences which have the following CV-like expressions (for a proof, see Kotsakis 2000b, ch. 5):

$$a(nh) = \frac{1}{N} \sum_{m=0}^{N-1} g(mh) g(mh + nh)$$

(16a)

and

$$b(nh) = \frac{1}{N} \sum_{m=0}^{N-1} \sigma_v(mh, mh + nh)$$

(16b)

The first sequence in Eq. (16a) can easily be identified as the discrete (spatio-statistical) CV function of the true deterministic signal at the given data resolution level, and thus the term $\mathcal{A}(\omega)$ in Eq. (15) is just the power spectrum of the true signal values $g(nh)$. The second sequence in Eq. (16b), on the other hand, does not exactly correspond to the discrete noise CV function and, as a result, the frequency-domain quantity $\mathcal{B}(\omega)$ in Eq. (15) should not generally be viewed as the power spectral density (PSD) function of the data noise. Such an interpretation is possible only in the special case where the input noise is (weakly) stationary. Indeed, in such situation the noise covariance $\sigma_v$ in Eq. (16b) between two arbitrary data points with coordinates $mh$ and $mh + nh$ becomes a function of their distance only, which is obviously equal to $nh$. Therefore, $\sigma_v$ can be taken outside of the summation operator, leaving the summation result in Eq. (16b) equal to $N$. In the more general case of non-stationary noise, the sequence $b(nh)$ 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 data point of the input grid, i.e.

$$b(0) = \frac{1}{N} \sum_{m=0}^{N-1} \sigma_v(mh, mh) = \frac{1}{N} \sum_{m=0}^{N-1} \sigma_v^2(mh)$$

(17)

whereas its values at the other points correspond to ‘averages’ of the noise covariance over pairs of data points with coordinate difference equal to $nh$. Note also that both sequences in Eqs. (16a) and (16b) are always symmetric, and they take zero values outside the range $-(N - 1) \leq n \leq (N - 1)$ due to the finite spatial support of the input signal $g(x)$. More details and comments on the properties and practical implementation of the optimal noise filter $\mathcal{W}(\omega)$ can be found in Kotsakis (2000b, ch. 5).
5 Numerical experiment

In order to examine the performance of the optimal noise filter $\mathcal{W}(\omega)$ from Eq. (15) at different data resolution levels, and under the presence of non-stationary observational noise, a simple numerical experiment was performed. First, a one-dimensional deterministic signal $g(x)$, assumed to represent some gravity anomaly profile, was synthesized using a truncated Fourier series expansion with a record length of 200 km (see Fig. 3). Four different sampling resolution levels ($h$) were used to create the various gridded data sets, namely 0.1, 0.5, 1.0 and 5.0 km. All the signal grids $g(nh)$, at every resolution level, were partitioned into three equal spatial blocks, labeled as left ($L$), central ($C$), and right ($R$). The simulated data noise, which is subsequently added to the true signal samples, will have a different stochastic behaviour in each of the three grid blocks.

![Image of the true (simulated) signal.](image)

**Fig. 3** The true (simulated) signal.

A zero-mean noise sequence $v(nh)$ was added to the samples $g(nh)$ of the true signal in order to generate the input data sets $d(nh)$ at each resolution level, according to the observation equation given in Eq. (1). The noise values originated from a non-stationary and uncorrelated Gaussian stochastic process, using the routines for random number generation of the MATLAB software package. The noise variance $\sigma^2(nh)$ was constant within each sub-block ($L$, $C$, $R$) of the data grids, with its values set to 20 mGals$^2$, 3 mGals$^2$ and 6 mGals$^2$, respectively. The sample statistics of the total noise sequence, at every resolution level, are given in Table 1.

The optimal noise filter $\mathcal{W}(\omega)$, according to Eqs. (15), (16a) and (16b), was computed via a fast Fourier transform (FFT) algorithm for each of the four sampling resolution levels ($h$), and it was then multiplied by the FFT of the noisy gridded data $d(nh)$. The result was finally transformed back to the space domain using an inverse FFT algorithm.

<table>
<thead>
<tr>
<th>Data resolution (in km)</th>
<th>0.1</th>
<th>0.5</th>
<th>1.0</th>
<th>5.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>14.57</td>
<td>15.34</td>
<td>8.91</td>
<td>8.27</td>
</tr>
<tr>
<td>Mean</td>
<td>0.09</td>
<td>0.24</td>
<td>0.27</td>
<td>0.11</td>
</tr>
<tr>
<td>Min</td>
<td>-15.09</td>
<td>-10.10</td>
<td>-11.98</td>
<td>-8.01</td>
</tr>
<tr>
<td>Std</td>
<td>3.15</td>
<td>3.08</td>
<td>3.09</td>
<td>3.01</td>
</tr>
<tr>
<td>RMS</td>
<td>3.15</td>
<td>3.09</td>
<td>3.10</td>
<td>3.01</td>
</tr>
</tbody>
</table>

**Table 1.** Statistics of the total additive noise at various data resolution levels (in mGals).

In Fig. 4, the original noisy data $d(nh)$ and the filtered signal estimates $\hat{g}(nh)$ are shown for some selective resolution levels $h$. The differences between the true signal samples and the corresponding estimated values are also shown in Fig. 5, whereas their statistics are given in Table 2.

![Image of noisy and filtered signal values at two different sampling resolution levels.](image)

**Fig. 4** Noisy and filtered signal values (in mGals) at two different sampling resolution levels. The vertical dashed lines mark the boundaries between the three blocks ($L$, $C$, $R$) of the input data grid with the different noise variances at each block.
Fig. 5 Differences between the filtered and the true signal values at two different sampling resolution levels.

Table 2. Statistics of the differences between the true and the filtered signal values at various data resolution levels (in mGals).

<table>
<thead>
<tr>
<th>Data resolution (in km)</th>
<th>0.1</th>
<th>0.5</th>
<th>1.0</th>
<th>5.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>1.47</td>
<td>1.28</td>
<td>3.50</td>
<td>6.48</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.09</td>
<td>-0.24</td>
<td>-0.26</td>
<td>-0.11</td>
</tr>
<tr>
<td>Min</td>
<td>-2.06</td>
<td>-3.38</td>
<td>-4.15</td>
<td>-6.47</td>
</tr>
<tr>
<td>Std</td>
<td>0.54</td>
<td>0.91</td>
<td>1.30</td>
<td>2.53</td>
</tr>
<tr>
<td>RMS</td>
<td>0.55</td>
<td>0.94</td>
<td>1.33</td>
<td>2.54</td>
</tr>
</tbody>
</table>

It is interesting to observe that the output estimation error for the filtered signal values is decreasing as the data sampling interval $h$ becomes smaller. This is evident from the comparison of the two graphs shown in Fig. 5, as well as from the error RMS values given in Table 2. 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. For more details, see Kotsakis (2000b).

6 Conclusions

We have presented a useful modification of the classic Wiener filtering algorithm which allows us to work with deterministic (instead of stochastic) unknown fields that are masked by additive non-stationary noise. The informal similarities of our frequency-domain estimation framework with the Wiener filtering formalism stem from the initial modelling choice in Eq. (6) that the optimal signal estimate should be linear and translation-invariant. This led to a convolution SNR-type computational scheme that can always be implemented very efficiently using FFT techniques. Of special importance in our derivations was the decomposition of the total signal estimation error into an aliasing-only component and a noise-dependent component (see Sect. 3). A detailed discussion on this subject, along with some comments on the problems encountered when we attempt to apply a ‘one-step’ optimization of the total signal estimation error, can be found in Kotsakis (2000b).

The spatial resolution of the noisy input data was also taken directly into account within our estimation procedure, revealing new interesting aspects related to the structure and the performance of the optimal noise filter as a function of the data grid density.

In terms of future work, our efforts should concentrate on extending the methodology presented herein for multi-dimensional problems, both in planar and spherical domains. Certain additional modifications are also needed in order to handle geodetic estimation applications that involve more than one type of noisy signal data (multiple-input/single-output systems), and not just the simple single-input/single-output case that was studied here. Nevertheless, the presented methodology can be proven a useful tool in many existing geodetic problems, such as the optimal spectral geoid determination from noisy gridded gravity data or the FFT-based computation of various height-dependent gravity field quantities from noisy digital terrain models.

References