Presented in Session G7: Regional and local gravity field approximation
XXVI General Assembly of the European Geophysical Society, Nice, March 25-30, 2001

Non-stationary noise filtering of gravity data using fast spectral techniques

C. Kotsakis and M.G. Sideris
Department of Geomatics Engineering, University of Calgary
2500 University Drive N.W.
Calgary, Alberta, Canada, T2N 1N4

Abstract

An important problem in the practical use of optimal spectral methods in gravity field modelling is the stationarity assumption for the data noise and the underlying unknown signals. Such a restriction is required, according to the standard Wiener-Kolmogorov estimation theory, in order to obtain signal approximation algorithms of simple convolution structure that can be evaluated very efficiently through fast Fourier transform (FFT) techniques. Often, the observation errors in the input data have significant spatial variations in their statistical behaviour, thus making the noise stationarity assumption unrealistic for many practical situations. Also, a stochastic interpretation of the true values of the various gravity field signals as random variables with similar statistical parameters is rather questionable, since they describe physical phenomena that are not random (probabilistic) and certainly do not have a uniform (stationary) behaviour over their domain. The aim of this paper is to present a spectral Wiener-type optimal filter which can be used in geodetic estimation problems with arbitrary deterministic signals that are masked by non-stationary observation errors.

1. Introduction

The method of Wiener filtering is a well known and efficient spectral tool that can be used for geodetic data 'de-noising' in the frequency domain. Its framework is based on the Wiener-Kolmogorov linear prediction theory for stationary random fields in the presence of stationary additive noise (KAILATH, 1974), and thus it is closely related to the method of least-squares collocation with random observation errors (MORITZ, 1980). 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 (i.e. input-output systems theory), has primarily focused on problems related to optimal spectral gravity field modelling. For some practical applications, see SIDERIS (1996), PAWLOWSKI AND HANSEN (1990), TZIAVOS ET AL. (1996). A detailed discussion on the use of the Wiener-Kolmogorov filtering theory in gravity field estimation, and its relationship with other linear approximation techniques traditionally used in geodesy, can be found in SANSO AND SIDERIS (1997).

In order to employ the classic Wiener filtering algorithm with noisy geodetic data, a stationarity assumption has to be made for both the true/unknown signal and the external random errors. Such restrictions become quite problematic for many gravity field applications, since (i) the underlying true signals cannot easily admit a stochastic interpretation (thus making the stationarity assumption meaningless) and, most importantly, (ii) the additive data noise does not usually follow a spatially uniform statistical behaviour; see KOTSAKIS AND SIDERIS (2001) for more discussion. In this paper we will outline a convolution-based algorithmic procedure that can be used in geodetic estimation problems regardless of the spatio-statistical properties of the underlying signals and the data noise. Our analysis will cover the simple case where an unknown deterministic field is observed under the masking of non-stationary random errors, and the desired output corresponds to an improved ('de-noised') linear interpolating model of the noisy input data. An important point in our approach is that the sampling resolution of the data will be taken into account within the optimization procedure, resulting in a resolution-dependent noise filter. A numerical example, using a synthetic two-dimensional gravity anomaly grid, has also been included to demonstrate the performance of our optimal noise filter under non-stationary additive noise, at different sampling resolution levels.

2. Methodology – Problem formulation

The main problem that is studied in this paper is the frequency-domain estimation of an unknown deterministic field $g(x,y)$ using its noisy gridded samples $d(nh_x, mh_y)_{n,m\in Z}$ according to the linear observation equation

$$d(nh_x, mh_y) = g(nh_x, mh_y) + v(nh_x, mh_y) \tag{1}$$

where $g(nh_x, mh_y)$ are the true signal values, $v(nh_x, mh_y)$ is a non-stationary noise sequence, and $h_x \times h_y$ corresponds to the sampling resolution level. The unknown signal will be assumed to have compact support over the real plane, and thus the integer sampling indices *n* and *m* in Eq. (1) can be practically restricted within a finite range $0 \le n \le (N-1)$ and $0 \le m \le (M-1)$. The Gauss-Markov stochastic model for the observation errors is

$$E\left\{ v(nh_x, mh_y) \right\} = 0 \tag{2a}$$

$$E\left\{ v^2(nh_x, mh_y) \right\} = \sigma_v^2(nh_x, mh_y) = \sigma_v\left[(nh_x, mh_y)\,(nh_x, mh_y)\right] \tag{2b}$$

$$E\left\{ v(nh_x, mh_y)\, v(kh_x, lh_y) \right\} = \sigma_v\left[(nh_x, mh_y)\,(kh_x, lh_y)\right] \tag{2c}$$

The symbol $\sigma_v^2(\cdot)$ denotes the noise variance at a single data point, whereas $\sigma_v\left[(\cdot)\,(\cdot)\right]$ corresponds to the noise covariance (CV) between two data points. The Fourier transform of the noiseless signal grid will be denoted by $\overline{G}(\omega_x, \omega_y)$ and it is given by the summation formula (DUDGEON AND MERSEREAU, 1984)

$$\overline{G}(\omega_x, \omega_y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} g(nh_x, mh_y)\, e^{-i(nh_x\omega_x + mh_y\omega_y)} \tag{3}$$

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

$$\overline{V}(\omega_x, \omega_y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} v(nh_x, mh_y)\, e^{-i(nh_x\omega_x + mh_y\omega_y)} \tag{4}$$

Two special 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 that is independent of the reference system used to describe the position of the data points. Note that the shift-invariance condition has often been applied in the theoretical formulation of optimal estimation methods using errorless data (SANSO, 1980; KOTSAKIS, 2000a), although its justification is not altered by the noise presence in the observations. Taking into account these two assumptions, the signal estimation formula will have the general convolution-type expression

$$\hat{g}(x, y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} d(nh_x, mh_y)\, \xi_h(x - nh_x, y - mh_y) \tag{5}$$

where $\xi_h(x, y)$ is a 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 data resolution level $h_x \times h_y$.

The signal error produced by the filtering formula in Eq. (5) can be decomposed into two components, i.e.

$$e(x, y) = g(x, y) - \hat{g}(x, y) = e_h(x, y) + e_v(x, y) \tag{6}$$

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. In the absence of noise from the input data, the best we can obtain is just an *interpolated model* $\tilde{g}(x, y)$ 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:

$$\tilde{g}(x, y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} g(nh_x, mh_y)\, \varphi_h(x - nh_x, y - mh_y) \tag{7}$$

where $\varphi_h(x,y)$ is some interpolating kernel that generally depends on the data resolution. The noise-dependent estimation error will be measured with respect to such a linear interpolating model for the unknown field, i.e.

$$e_v(x,y) = \tilde{g}(x,y) - \hat{g}(x,y) \tag{8a}$$

whereas the (pure) aliasing error is

$$e_h(x,y) = g(x,y) - \tilde{g}(x,y) \tag{8b}$$

The actual choice of the interpolating kernel $\varphi_h(x,y)$ is irrelevant for the purpose of this paper, and it can be optimized by minimizing a suitable functional of the aliasing error component $e_h(x,y)$; see KOTSAKIS (2000a, b). The unknown filtering kernel in Eq. (5) will be determined through the familiar Wiener-Kolmogorov *mean-square-error* (MSE) criterion

$$P_{e_v}(\omega_x,\omega_y) = E\left\{ \left| E_v(\omega_x,\omega_y) \right|^2 \right\} = \text{minimum} \tag{9}$$

where $E_v(\omega_x,\omega_y)$ is the Fourier transform of the noise-dependent error term $e_v(x,y)$. It is easy to show that the corresponding optimal estimation filter will finally be given by the formula

$$\Xi_h(\omega_x,\omega_y) = \frac{\left|\overline{G}(\omega_x,\omega_y)\right|^2}{\left|\overline{G}(\omega_x,\omega_y)\right|^2 + \overline{P}_v(\omega_x,\omega_y)}\, \Phi_h(\omega_x,\omega_y) = \overline{W}(\omega_x,\omega_y)\; \Phi_h(\omega_x,\omega_y) \tag{10}$$

where $\Xi_h(\omega_x,\omega_y)$ and $\Phi_h(\omega_x,\omega_y)$ are the Fourier transforms of the filtering kernel $\xi_h(x,y)$ and the 'reference' interpolating kernel $\varphi_h(x,y)$, respectively. The auxiliary term $\overline{P}_v(\omega_x,\omega_y)$ shown in the last equation corresponds to the 'power spectral density (PSD)' noise quantity

$$\overline{P}_v(\omega_x,\omega_y) = E\left\{ \overline{V}(\omega_x,\omega_y)\, \overline{V}^*(\omega_x,\omega_y) \right\} = E\left\{ \left| \overline{V}(\omega_x,\omega_y) \right|^2 \right\} \tag{11}$$

where E denotes the usual expectation operator. The result in Eq. (10) indicates that the estimation procedure can be decomposed into two individual steps which are connected in a linear cascading manner (see Figure 1). The first step, expressed by the filter component $\overline{W}(\omega_x,\omega_y)$, has the role of 'de-noising' the discrete data using information on the average behaviour of the input noise and the unknown field at the given resolution level. The second filter component $\Phi_h(\omega_x,\omega_y)$ is solely used to obtain a continuous representation for the output signal $\hat{g}(x,y)$ based on an a-priori selected interpolating/modelling kernel $\varphi_h(x,y)$. For more mathematical details and derivations, see KOTSAKIS AND SIDERIS (2001).

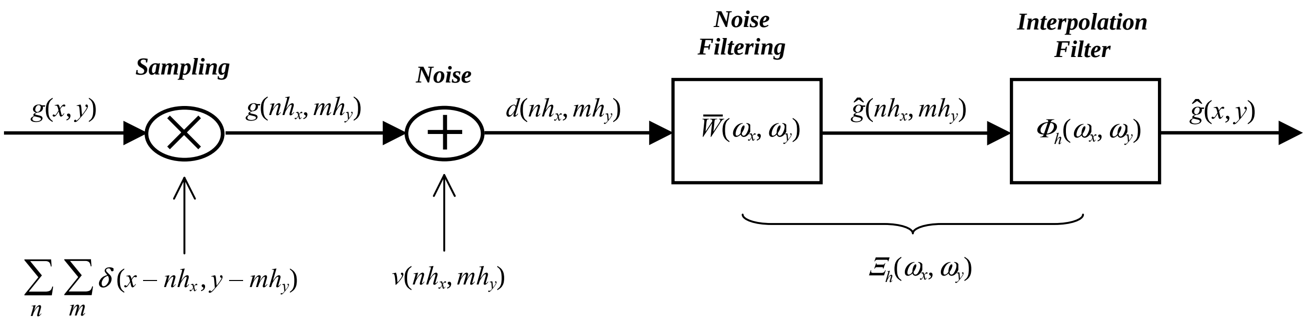

Figure 1. The cascading structure of the optimal linear estimation filter.

3. The Wiener-like structure of the optimal estimation filter

The structure of the optimal estimation filter in Eq. (10) is very similar to the classic Wiener filter, since 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 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. Thus, it is important to clarify what is the exact meaning of the two frequency-domain terms that appear in the expression of our SNR-type optimal noise filter $\overline{W}(\omega_x,\omega_y)$. From Eq. (10), we have that

$$\overline{W}(\omega_x,\omega_y) = \frac{\frac{1}{NM}\left|\overline{G}(\omega_x,\omega_y)\right|^2}{\frac{1}{NM}\left|\overline{G}(\omega_x,\omega_y)\right|^2 + \frac{1}{NM}\overline{P}_v(\omega_x,\omega_y)} = \frac{\overline{A}(\omega_x,\omega_y)}{\overline{A}(\omega_x,\omega_y) + \overline{B}(\omega_x,\omega_y)} \tag{12}$$

where NM is the total number of points in the input data grid. The two auxiliary functions $\overline{A}(\omega_x,\omega_y)$ and $\overline{B}(\omega_x,\omega_y)$, in the last equation, correspond to the Fourier transforms of two associated sequences which have the CV-like expressions (for a proof see KOTSAKIS AND SIDERIS, 2001)

$$a(nh_x, mh_y) = \frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} g\big(kh_x, lh_y\big)\; g\big((k+n)h_x, (l+m)h_y\big) \tag{13a}$$

and

$$b(nh_x, mh_y) = \frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} \sigma_v\Big[\big(kh_x, lh_y\big)\;\big((k+n)h_x, (l+m)h_y\big)\Big] \tag{13b}$$

The sequence in Eq. (13a) is the discrete (spatio-statistical) CV function of the true deterministic signal at the given data resolution level, and thus the term $\overline{A}(\omega_x,\omega_y)$ in Eq. (12) is just the power spectrum of the true signal values $g(nh_x,mh_y)$. The sequence in Eq. (13b), on the other hand, does not exactly correspond to the noise CV function and, as a result, the quantity $\overline{B}(\omega_x,\omega_y)$ in Eq. (12) should not in general be viewed as the data noise PSD (such an interpretation is possible only when the data noise is stationary). The sequence $b(nh_x, mh_y)$ can be perceived 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.

$$b(0,0) = \frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} \sigma_v\big[(kh_x, lh_y)\,(kh_x, lh_y)\big] = \frac{1}{NM}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1} \sigma_v^2(kh_x, lh_y) \tag{14}$$

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_x,mh_y). The numerical evaluation of the optimal noise filter $\overline{W}(\omega_x,\omega_y)$ 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 $b(nh_x, mh_y)$; more details on the properties and practical implementation of the optimal noise filter can be found in KOTSAKIS AND SIDERIS (2001).

4. Numerical experiment

A numerical experiment was performed to test the noise filtering component of the optimal estimation kernel that was previously derived. A deterministic signal $g(x, y)$, assumed to represent some local gravity anomaly field, was initially synthesized using a truncated Fourier series expansion with a record length of 200 × 200 km. The continuous signal was sampled at various resolution levels to obtain noiseless gridded values

$g(nh_x, mh_y)$. Four different sampling resolutions were selected, namely 0.5 × 0.5, 1 × 1, 2.5 × 2.5 and 5 × 5 km. All signal grids at each resolution level were partitioned into four equal blocks/quadrants, labeled as northwest (NW), northeast (NE), southwest (SW) and southeast (SE). The simulated data noise, which is going to be added to the true signal values, will have different statistical behaviour in each of the four grid quadrants.

A zero-mean noise sequence was added to the samples of the true signal in order to create the input data $d(nh_x, mh_y)$ at every resolution level. 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 was constant within each quadrant (NW, NE, SW and SE) of the data grids, with its values set to 144 mGals^2, 9 mGals^2, 144 mGals^2 and 49 mGals^2, respectively. The sample statistics of the simulated noise values in the four different quadrants of the data grids are shown in Table 1, for some selective resolution levels.

Table 1. Statistics of the simulated noise values in the four quadrants of the data grids (in mGals).

Data resolution (in Km)	**0.5 × 0.5**				**1.0 × 1.0**			
Grid quadrants	NW	SW	NE	SE	NW	SW	NE	SE
Max	49.61	55.64	12.47	27.37	53.15	43.04	11.45	25.88
Mean	-0.04	0.02	0.02	-0.01	0.03	0.01	0.02	-0.08
Min	-52.61	-46.34	-12.14	-27.14	-52.71	-47.87	-10.24	-28.58
Std	11.99	12.05	3.00	6.98	12.07	12.08	2.95	7.01

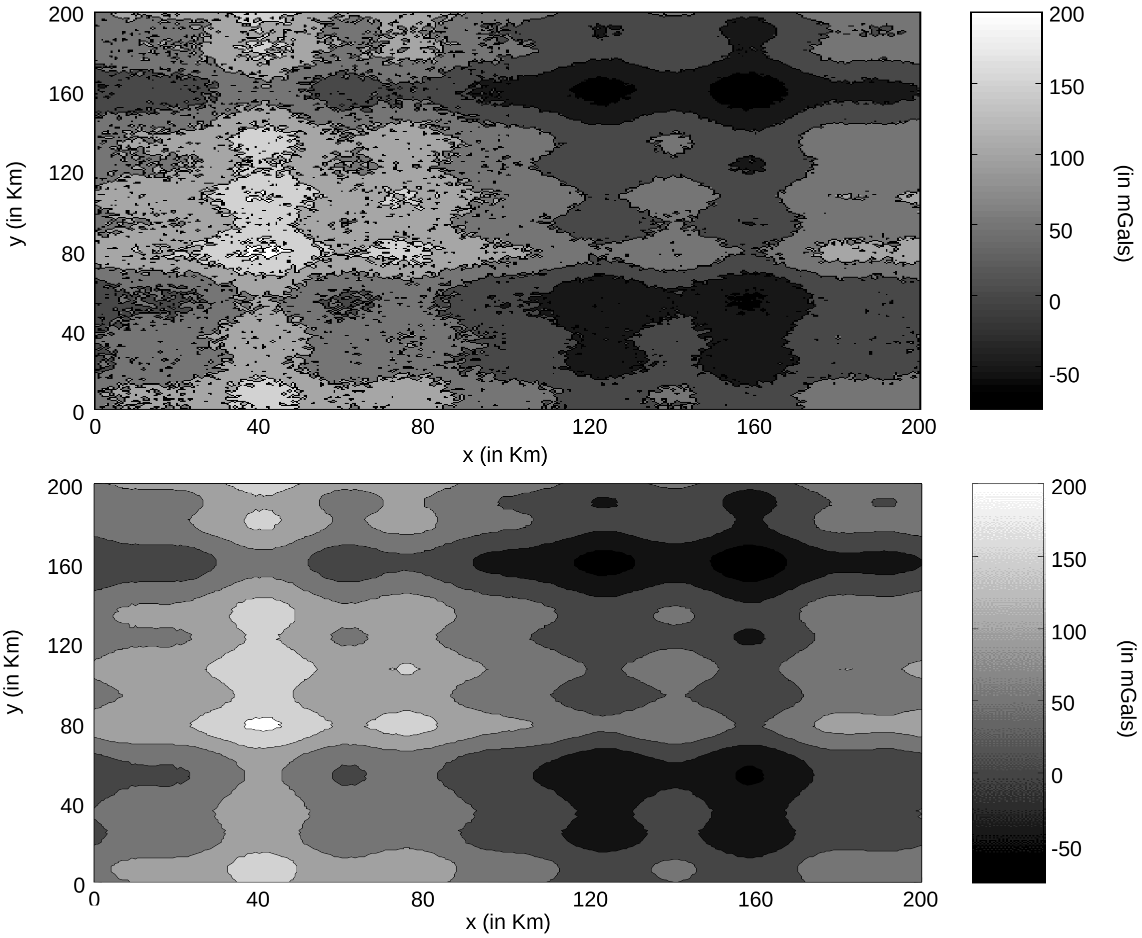

Figure 2. Noisy gravity anomaly values (top plot) and estimated/filtered gravity anomaly values (bottom plot), for the data resolution level (1 × 1) km. Observe that the input noise is much stronger in the western parts of the data grids.

The optimal noise filter $\overline{W}(\omega_x,\omega_y)$ was computed through a fast Fourier transform (FFT) algorithm at each resolution level $h_x \times h_y$, according to the SNR expression given in Eq. (12). It was then multiplied by the FFT of the noisy gridded data $d(nh_x, mh_y)$ and the result was finally transformed back to the space domain as an estimated ('de-noised') signal sequence $\hat{g}(nh_x, mh_y)$. An example of the original noisy data grid and the filtered signal values is given in Figure 2. The statistics of the differences between the true signal samples and the estimated signal values, for every resolution level used, are also given in Table 2 below.

Table 2. Statistics of the differences between the true and the filtered signal values (in mGals).

Data resolution (in Km)	**0.5 × 0.5**	**1.0 × 1.0**	**2.5 × 2.5**	**5.0 × 5.0**
Max	3.09	3.51	5.53	6.26
Mean	0.00	0.01	-0.05	0.10
Min	-3.25	-4.00	-4.86	-8.01
Std	0.66	0.91	1.43	2.26

It is interesting to observe that the output estimation error of the filtered signal values is decreasing, as the data resolution increases. 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.

5. Conclusions

We have presented a modification of the classic Wiener filtering algorithm which allows us to work with deterministic fields that are masked by additive non-stationary noise. The informal similarities of our estimation framework with the Wiener filtering formalism stem from the initial assumption in Eq. (5) 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 estimation error into an aliasing component and a noise-dependent component. 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 AND SIDERIS (2001). The presented methodology can be proven a useful tool in various geodetic estimation problems of local and/or regional scale, such as the optimal spectral geoid determination from noisy gridded gravity data.

References

DUDGEON D.E., MERSEREAU R.M. (1984) Multidimensional digital signal processing. Prentice-Hall, Englewood Cliffs.

KAILATH T. (1974) A view of three decades of linear filtering theory. IEEE Trans Info Theory, **IT-20**(2): 146-181.

KOTSAKIS C. (2000a) The multiresolution character of collocation. J Geod, **74**(3-4): 275-290.

KOTSAKIS C. (2000b) Multiresolution aspects of linear approximation methods in Hilbert spaces using gridded data. PhD thesis, Dept. of Geomatics Engineering, University of Calgary, UCGE Report No. 20138, Calgary, Alberta.

KOTSAKIS C., SIDERIS M.G. (2001) A modified Wiener-type filter for geodetic estimation problems with non-stationary noise. J Geod, **75**(11-12) (to appear).

MORITZ H. (1980) Advanced Physical Geodesy. Herbert Wichmann Verlag, Karlsruhe.

PAWLOWSKI R.S., HANSEN R.O. (1990) Gravity anomaly separation by Wiener filtering. Geoph, **55**(5): 539-548.

SANSO 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.

SANSO F., SIDERIS M.G. (1997) On the similarities and differences between systems theory and least-squares collocation in physical geodesy. Boll Geod Sci Affi, **54**(2): 173-206.

SIDERIS M.G. (1996) On the use of heterogeneous noisy data in spectral gravity field modelling methods. J Geod, **70**(8): 470-479.

TZIAVOS I.N., SIDERIS M.G., LI J. (1996) Optimal spectral combination of satellite altimetry and marine gravity data. Report of the Finnish Geodetic Institute, **96**(2): 41-56, Presented at the XXI EGS General Assembly, Hague, May 6-10.