Assessment of noise in GPS coordinate time series: Methodology and results

A. R. Amiri-Simkooei,[1,2] C. C. J. M. Tiberius,[1] and P. J. G. Teunissen[1]

[1]   We propose a methodology to assess the noise characteristics in time series of position estimates for permanent Global Positioning System (GPS) stations. Least squares variance component estimation (LS-VCE) is adopted to cope with any type of noise in the data. LS-VCE inherently provides the precision of (co)variance estimators. One can also apply statistical hypothesis testing in conjunction with LS-VCE. Using the w-test statistic, a combination of white noise and flicker noise turns out in general to best characterize the noise in all three position components. An interpretation for the colored noise of the series is given. Unmodelled periodic effects in the data will be captured by a set of harmonic functions for which we rely on the least squares harmonic estimation (LS-HE) method and parameter significance testing developed in the same framework as LS-VCE. Having included harmonic functions into the model, practically only white noise can be shown to remain in the data. Remaining time correlation, present only at very high frequencies (spanning a few days only), is expressed as a first-order autoregressive noise process. It can be caused by common and well-known sources of errors like atmospheric effects as well as satellite orbit errors. The autoregressive noise should be included in the stochastic model to avoid the overestimation (upward bias) of power law noise. The results confirm the presence of annual and semiannual signals in the series. We observed also significant periodic patterns with periods of 350 days and its fractions $350/n$, $n = 2, \ldots, 8$ that resemble the repeat time of the GPS constellation. Neglecting these harmonic signals in the functional model can seriously overestimate the rate uncertainty.

1.   Introduction

[2]   Continuous Global Positioning System (GPS) measurements have been used now nearly 15 years for estimation of crustal deformation. Station positions are determined with respect to an earth-fixed terrestrial reference system. Geophysical studies using geodetic measurements of surface displacement or strain require not only accurate estimates of these parameters but also accurate error estimates. The precision of these estimates is often assessed by their repeatability defined by the mean squared error of individual coordinate components (i.e., north, east, and vertical) about a linear trend. Except for the significant episodic deformation, such as large earthquakes, a linear trend can be a good representative of the deformation behavior. The site velocities are usually determined by linear regression of individual coordinate components. The least squares technique is used to estimate the line parameters, i.e., the intercept and the slope (site velocity).

[3]   In the ideal case, it is desired that the time series possess only white noise and all functional effects are fully understood. The noise in GPS coordinate time series turns out not to be white. Several geodetic data sets have provided evidence for error sources that introduce large temporal correlations into the data. The ultimate goal of noise studies is to come up with a stochastic model that allows one to process the coordinate time series such that the "best" solution (most precise solution together with proper precision description) of the station positions and site velocities can be determined. An intermediate goal is therefore to better understand and to identify the various noise components of the stochastic model.

[4]   Two techniques have generally been employed to assess the noise characteristics of geodetic time series, namely, the power spectral method and the maximum likelihood estimation (MLE) method. The former is aimed to examine the data in the frequency domain while the latter is used to examine the data covariance matrix in the time (space) domain. The MLE can estimate the parameters of a noise model effectively in contrast to the classical power spectra techniques. In this contribution, we will not make

---

[1]Delft Institute of Earth Observation and Space systems (DEOS), Delft University of Technology, Delft, Netherlands.
[2]Department of Surveying Engineering, Faculty of Engineering, The University of Isfahan, Isfahan, Iran.

use of the spectral techniques. The MLE method is generally used to compute the amount of white noise, flicker noise, and random walk noise in the time series [see, e.g., *Zhang et al.*, 1997; *Langbein and Johnson*, 1997; *Mao et al.*, 1999; *Williams et al.*, 2004; *Langbein*, 2004]. In this paper, we introduce and use a different variance component estimation method based on the least squares principle. Our motivation is given in the next section.

2. Previous Work and Outline

[5] *Zhang et al.* [1997] processed 19 months of continuous GPS coordinates from 10 sites in southern California. Using MLE with integer spectral indices, they found that the noise in the GPS time series was best described as a combination of white noise and flicker noise. This combination suggested that the velocity uncertainties should be three to six times larger than those obtained from a pure white noise model. Using the power spectra, the noise was characterized by a fractal noise process with a spectral index of -0.4. Neglecting this fractal white noise model, site velocity uncertainties could be underestimated by a factor of $2-4$. In an analogous way, *Calais* [1999], *Mao et al.* [1999], and *Williams et al.* [2004] found that GPS position time series best fitted a noise model consisting of both white noise and flicker noise. Higher frequency $(1-30 \text{ s})$ GPS position time series have also been shown to contain white plus flicker noise [*Bock et al.*, 2000; *Langbein and Bock*, 2004].

[6] Several studies have also recognized random walk noise in geodetic data. Random walk noise was detected in continuous measurements of strainmeters as well as very short baseline GPS data at Piñon Flat Observatory in southern California [*Wyatt*, 1982, 1989; *Wyatt et al.*, 1989; *Johnson and Agnew*, 2000]. *Langbein and Johnson* [1995, 1997] showed that the noise in the electronic distance measuring (EDM) data is well characterized by a combination of white and random walk noise. The random walk amplitude for a very short baseline at Piñon Flat Observatory is only 0.4 mm/year$^{1/2}$ [*Johnson and Agnew*, 2000]. *Beavan* [2005] shows that the noise properties of GPS time series for concrete pillar monuments are very similar to those of deeply drilled braced monuments. Using two-color EDM measurements in California, *Langbein* [2004] shows that the random walk noise model is valid for about 30% of the data. In some cases, a combination of random walk and band-pass-filtered noise best characterizes the data.

[7] Site position time series obtained from continuous GPS arrays show significant seasonal variations with annual and semiannual periods. Such seasonal deformation is present in both global and regional GPS coordinate time series [see *van Dam et al.*, 2001; *Dong et al.*, 2002]. *Kusche and Schrama* [2005] show that after removing the atmospheric pressure loading effect, estimated annual variations of continental scale mass redistribution exhibit patterns similar to those obtained with Gravity Recovery and Climate Experiment (GRACE). *Ding et al.* [2005] used time series of daily positions of eight colocated GPS and very long baseline interferometry (VLBI) stations to assess the seasonal signals using the wavelet transform. *Blewitt and Lavallée* [2002] showed that annual signals can significantly bias the site velocity if they are not estimated in the model. Another important systematic error in GPS time series is the presence

of offsets (jumps). *Williams* [2003b] and *Perfetti* [2006] discuss offset detection and estimation strategies. *Kenyeres and Bruyninx* [2004] estimate offsets for coordinate time series in the EUREF permanent network.

[8] This study differs in several ways from previous work. We use the least squares variance component estimation (LS-VCE) method, which has attractive and unique features that we point out in this paper. First, LS-VCE is generally applicable and can cope with any type of noise (and with any number of noise components) in the data series. The method can be implemented in a relatively simple and straightforward manner. Second, using LS-VCE one can obtain the covariance matrix of the estimators describing the uncertainty of the (co)variance components. Third, the general formulation of LS-VCE is applied to a special case to estimate time correlation assuming that the time series are stationary in time. Fourth, we introduce the *w*-test statistic with which one can simply test the "contribution" of single noise components. One can determine which noise combination best characterizes the noise of GPS position time series. In the same framework as LS-VCE, we then introduce the least squares harmonic estimation (LS-HE) method. The goal is to introduce harmonic functions to capture unmodelled effects in the time series. It is shown that practically only white noise remains, which is attractive from the data processing point of view. Such a duality between the stochastic and the functional model is useful to be able to correctly judge on the amount and the behavior of noise.

3. Variance Component Estimation (VCE)

[9] In this section, we briefly explain the power spectrum of a power law noise process, the maximum likelihood estimation (MLE) of variance components, the least squares variance component estimation (LS-VCE), and the modeling approach and misspecification of a noise process.

3.1. Introduction to Noise Process

[10] The power spectra, P_y, of many geophysical phenomena, including the noise in GPS position time series, are well approximated by a power law process [*Mandelbrot*, 1983; *Agnew*, 1992; *Mao et al.*, 1999; *Williams*, 2003a; *Williams et al.*, 2004]. The one-dimensional time behavior of the stochastic process is such that its power spectrum has the form

$$P_y(f) = P_0 \left(\frac{f}{f_0} \right)^{\kappa} \tag{1}$$

where f is the temporal frequency, P_0 and f_0 are the normalizing constants, and κ is the spectral index [see, e.g., *Mandelbrot and van Ness*, 1968]. Typical spectral index values lie within $[-3, 1]$; for stationary processes $-1 < \kappa < 1$ and for nonstationary processes $-3 < \kappa < -1$. A smaller spectral index implies a more correlated process and a more relative power at lower frequencies. Special cases within this stochastic process occur at the integer values for κ. Classical white noise has a spectral index of 0, flicker noise has a spectral index of -1, and random walk noise has a spectral index of -2. The power spectral method can be employed to assess the noise characteristic of GPS time series.

[11] The second way is to use (co)variance component estimation (VCE) methods. The role of the data series covariance matrix is considered to be an important element with respect to the quality criteria of the unknown parameters. Therefore VCE methods are of great importance. There are many VCE methods, a review of which is out of the scope of this contribution. For more information on VCE methods and their applications to GPS data, refer to *Crocetto et al.* [2000], *Wang et al.* [2002], *Tiberius and Kenselaar* [2003], *Amiri-Simkooei and Tiberius* [2004], *Zhang et al.* [2004], and *Bischoff et al.* [2005].

[12] The noise components of GPS coordinate time series, i.e., white noise, flicker noise, and random walk noise, are usually estimated by the MLE method, which is a well-known estimation principle. The MLE problem can be solved in several ways [see *Kubik*, 1970; *Koch*, 1986; *Ou*, 1989; *Yu*, 1996; *Grodecki*, 1999, 2001]. In time series analysis of GPS coordinates, the MLE is described by *Langbein and Johnson* [1997], *Zhang et al.* [1997], *Calais* [1999], *Mao et al.* [1999], and *Williams et al.* [2004]. They all selected the downhill simplex method developed by Nelder and Mead [*Press et al.*, 1992]. Using expressions given by *Sahin et al.* [1992], some authors used a simplified algorithm for VCE [see *Davies and Blewitt*, 2000; *Altamimi et al.*, 2002].

3.2. Least Squares Variance Component Estimation

[13] Least squares variance component estimation (LS-VCE) was developed by *Teunissen* [1988]. For a review, see *Teunissen and Amiri-Simkooei* [2006, 2007] and *Amiri-Simkooei* [2007]. Consider the linear model of observation equations with p-number of unknown (co)variance components

$$\mathsf{E}\left(\underline{y}\right) = \mathbf{A}x, \quad \mathsf{D}\left(\underline{y}\right) = \mathbf{Q}_y = \mathbf{Q}_0 + \sum_{k=1}^{p} \sigma_k \mathbf{Q}_k \qquad (2)$$

where an underscore indicates that a quantity is a random variable and E and D are the expectation and dispersion operators, respectively. In the preceding functional model, $\underline{y}$ is the m-vector of observables, x is the n-vector of parameters of interest, and the $m \times n$ design matrix $\mathbf{A}$ is to be of full column rank. The data covariance matrix $\mathbf{Q}_y$ is expressed as an unknown linear combination of known $m \times m$ cofactor matrices $\mathbf{Q}_k$'s. $\mathbf{Q}_0$ is the known part (if any) of the stochastic model; it also allows one to use a nonlinear stochastic model (see section 6.5). The unknown (co)variance components $\sigma_k, k = 1, \ldots, p$ are to be estimated.

[14] To apply the least squares method to (co)variance component model, one can reformulate the second part of equation (2) in terms of a model of observation equations as $\mathsf{E}(\mathrm{vh}(\underline{tt}^T - \mathbf{B}^T\mathbf{Q}_0\mathbf{B})) = \mathbf{A}_{\mathrm{vh}}\sigma$, where $\sigma = [\sigma_1, \cdots, \sigma_p]^T, \underline{t} = \mathbf{B}^T\underline{y}$, and $\mathbf{A}_{\mathrm{vh}} = [\mathrm{vh}(\mathbf{B}^T\mathbf{Q}_1\mathbf{B}), \ldots, \mathrm{vh}(\mathbf{B}^T\mathbf{Q}_p\mathbf{B})]$ in which vh (vector-half) operator applies to symmetric matrices and $\mathbf{B}$ is an $m \times (m - n)$ matrix of which the $m - n$ linearly independent columns span the null-space of $\mathbf{A}^T$, i.e., $\mathbf{A}^T\mathbf{B} = 0$ or $\mathbf{B}^T\mathbf{A} = 0$.

[15] The least squares estimator for the p-vector of unknown (co)variance components can then be obtained as

$$\hat{\underline{\sigma}} = \left(\mathbf{A}_{\mathrm{vh}}^T\mathbf{Q}_{\mathrm{vh}}^{-1}\mathbf{A}_{\mathrm{vh}}\right)^{-1}\mathbf{A}_{\mathrm{vh}}^T\mathbf{Q}_{\mathrm{vh}}^{-1}\mathrm{vh}\left(\underline{tt}^T - \mathbf{B}^T\mathbf{Q}_0\mathbf{B}\right) = \mathbf{N}^{-1}\underline{l} \qquad (3)$$

where $\mathbf{Q}_{\mathrm{vh}}$ is the covariance matrix of the observables vh($\underline{tt}^T$). One can show that the entries of the $p \times p$ normal matrix $\mathbf{N}$ and of the p-vector $\underline{l}$ are obtained as

$$n_{kl} = \frac{1}{2}\mathrm{tr}\left(\mathbf{Q}_y^{-1}\mathbf{P}_\mathbf{A}^\perp\mathbf{Q}_k\mathbf{Q}_y^{-1}\mathbf{P}_\mathbf{A}^\perp\mathbf{Q}_l\right), k, l = 1, 2, \ldots, p \qquad (4)$$

and

$$\underline{l}_k = \frac{1}{2}\hat{\underline{e}}^T\mathbf{Q}_y^{-1}\mathbf{Q}_k\mathbf{Q}_y^{-1}\hat{\underline{e}} - \frac{1}{2}\mathrm{tr}\left(\mathbf{Q}_k\mathbf{Q}_y^{-1}\mathbf{P}_\mathbf{A}^\perp\mathbf{Q}_0\mathbf{Q}_y^{-1}\mathbf{P}_\mathbf{A}^\perp\right) \qquad (5)$$

where $k = 1, 2, \ldots, p$ and the least squares residual vector $\hat{\underline{e}} = \mathbf{P}_\mathbf{A}^\perp\underline{y}$ in which the orthogonal projector $\mathbf{P}_\mathbf{A}^\perp$ is given as

$$\mathbf{P}_\mathbf{A}^\perp = \mathbf{I} - \mathbf{A}\left(\mathbf{A}^T\mathbf{Q}_y^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{Q}_y^{-1} \qquad (6)$$

with $\mathbf{I}$ an identity matrix. Since the estimators $\hat{\underline{\sigma}}$ are based on the least squares method, the inverse of the normal matrix $\mathbf{N}$ automatically gives the covariance matrix of the estimated (co)variance components

$$\mathbf{Q}_{\hat{\sigma}} = \mathbf{N}^{-1} \qquad (7)$$

Therefore this equation offers us measures of precision for the estimators.

[16] To implement the method, one starts with an initial guess for the (co)variance components ($\sigma_k^0, k = 1, \ldots, p$). Using these values, one computes $\mathbf{Q}_y = \mathbf{Q}_0 + \sum_{k=1}^{p}\sigma_k^0\mathbf{Q}_k$. Equation $\hat{\sigma} = \mathbf{N}^{-1}l$ with equations (4) and (5) gives estimates for the $\sigma_k, k = 1, \ldots, p$, which in the next cycle are considered as an improved initial guess for those (co)variance components. This iterative procedure is repeated until the estimated (co)variance components do not change with further iterations. In this section, we considered a linear stochastic model. LS-VCE can also be applied to a nonlinear stochastic model, namely, $\mathbf{Q}_y = \mathbf{Q}(\sigma)$. To overcome the nonlinearity, one can expand the model into a Taylor series for which one needs the initial values of the unknown vector σ, namely, σ^0. After linearization, one obtains a linear form of the (co)variance component model and thus equation (3) can be used [see *Amiri-Simkooei*, 2007].

[17] In contrast to MLE which gives biased estimators, LS-VCE provides *unbiased* and *minimum variance* estimators. The unbiasedness property is independent of the (unspecified) distribution of the data. LS-VCE is also faster than MLE since it iterates in a Newton–Raphson scheme toward a solution rather than using the downhill simplex which can be extremely slow [see *Press et al.*, 1992]. With LS-VCE one can thus efficiently incorporate any number of noise components in the stochastic model. Using hypothesis testing, one can also simply judge in an objective manner which noise components are likely to be present in the series (see section 5.2).

3.3. Misspecification in Functional and Stochastic Model

[18] Misspecifications (errors) in the functional model and/or stochastic model will, in general, affect the optimality properties of the estimators. This also holds true for

variance component estimation. It is therefore of interest to understand how such misspecifications affect the result of estimation. Without treating this topic in detail, it is relevant to briefly mention some of these effects. Concerning the number of parameters in the model $\mathsf{E}(\underline{y}) = \mathbf{A}x$ and $\mathbf{Q}_y = \mathbf{Q}_0 + \sum_{k=1}^{p} \sigma_k \mathbf{Q}_k$, two types of misspecifications can occur: overparametrization and underparametrization.

[19] Underparametrization in the functional model $\mathsf{E}(\underline{y}) = \mathbf{A}x$ will generally lead to biases in the estimation of x and thus also in the results of variance component estimation (aliasing). Such biases in the results of VCE may have the side effect, that they become misinterpreted as an underparametrization of the stochastic model (see section 5.3). In contrast to underparametrization, overparametrization in the functional model does not lead to biases in the estimation of x. Here, however, one has to be aware that overparametrization reduces the redundancy and therefore also the precision with which the results can be obtained. This is true for the estimation of x, as well as for the estimation of the (co)variance components.

[20] Misspecifications in the stochastic model will not lead to biases in the estimation of x. However, underparametrization in the stochastic model will lead to biases in the estimation of the (co)variance components, as a consequence of which an incorrect precision description is obtained for the estimator of x. To discuss this effect, let $\mathbf{Q}_y$ be the correct covariance matrix and $\mathbf{Q}'_y$ be the incorrect one. The least squares estimator of x, based on $\mathbf{Q}'_y$, is then given as $\hat{x} = (\mathbf{A}^T \mathbf{Q}'^{-1}_y \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}'^{-1}_y \underline{y}$. This estimator is still unbiased [see *Teunissen et al.*, 2005]. Its covariance matrix follows from an application of the error propagation law as

$$\mathbf{Q}_{\hat{x}} = \left(\mathbf{A}^T \mathbf{Q}'^{-1}_y \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{Q}'^{-1}_y \mathbf{Q}_y \mathbf{Q}'^{-1}_y \mathbf{A} \left(\mathbf{A}^T \mathbf{Q}'^{-1}_y \mathbf{A}\right)^{-1} \quad (8)$$

If one believes that $\mathbf{Q}'_y$ is the correct covariance matrix, while it is not, one will use the matrix $\mathbf{Q}'_x = (\mathbf{A}^T \mathbf{Q}'^{-1}_y \mathbf{A})^{-1}$ to describe the precision of $\hat{x}$. This matrix, however, gives an incorrect precision description, which can either be too optimistic ($\mathbf{Q}_{\hat{x}} \geq \mathbf{Q}'_x$) or too pessimistic ($\mathbf{Q}_{\hat{x}} \leq \mathbf{Q}'_x$). Comparisons of precision descriptions for different stochastic models are given in section 6.6.

4. GPS Coordinate Time Series

[21] This section demonstrates how to estimate the time correlation of GPS coordinate time series using LS-VCE. We rely on a commonly accepted structure of the functional and stochastic model and eventually arrive at a simple expression.

4.1. Functional Model

[22] We restrict ourselves to the problem of time correlation estimation for an individual component of GPS coordinate time series. In equation (2), y is the m-vector of time series observations, for example, daily GPS positions of one component. Hereinafter it is denoted by $y(t)$ where t refers to the time instant. When a linear trend describes the deformation behavior, the functional model will read: $\mathsf{E}(y(t)) =$

$y_0 + rt$. When there are in addition q periodic signals in the data series, the functional model is extended to

$$\mathsf{E}(y(t)) = y_0 + rt + \sum_{k=1}^{q} a_k \cos \omega_k t + b_k \sin \omega_k t \quad (9)$$

Two trigonometrical terms cos and sin together represent a sinusoidal wave with in general a nonzero initial phase. The structure introduced above has the advantage of being linear. The unknown vector x consists of the intercept y_0, the slope r, and the coefficients a_k and b_k. In case of a linear trend and annual and semiannual signals ($q = 2$), the design matrix $\mathbf{A}$ is of size $m \times 6$. Its ith row at time instant t_i is given as

$$a_i = [1 \ \ t_i \ \ \cos 2\pi t_i \ \ \sin 2\pi t_i \ \ \cos 4\pi t_i \ \ \sin 4\pi t_i] \quad (10)$$

where t_i is expressed in terms of year. In section 5.1, we show how to obtain an appropriate functional model.

4.2. Stochastic Model

[23] If the time series of GPS coordinates is composed of white noise, flicker noise, and random walk noise with variances σ^2_w, σ^2_f, and σ^2_{rw}, respectively, the covariance matrix of the time series can then be written as ($\mathbf{Q}_0 = 0$)

$$\mathbf{Q}_y = \sigma^2_w \mathbf{I} + \sigma^2_f \mathbf{Q}_f + \sigma^2_{rw} \mathbf{Q}_{rw} \quad (11)$$

where $\mathbf{I}$ is the $m \times m$ identity matrix and $\mathbf{Q}_f$ and $\mathbf{Q}_{rw}$ are the cofactor matrices relating to flicker noise and random walk noise, respectively. The structure of $\mathbf{Q}_y$ matrix is known through $\mathbf{I}$, $\mathbf{Q}_f$, and $\mathbf{Q}_{rw}$, but the contributions through σ_w, σ_f, and σ_{rw} are unknown. In section 5.2, we show how to improve an existing stochastic model.

[24] The elements of the flicker noise cofactor matrix $\mathbf{Q}_f$ can be approximated by [*Zhang et al.*, 1997]

$$q^{(f)}_{ij} = \begin{cases} \dfrac{9}{8} & \text{if} \quad \tau = 0 \\ \dfrac{9}{8}\left(1 - \dfrac{\log \tau / \log 2 + 2}{24}\right) & \text{if} \quad \tau \neq 0 \end{cases} \quad (12)$$

where $\tau = |t_j - t_i|$. For evenly spaced data, the matrix $\mathbf{Q}_f$ is a symmetric Toeplitz matrix that contains constant values along negative-sloping diagonals. It is important to note that the Hosking flicker noise covariance matrix, which was introduced and used by *Williams* [2003a], *Langbein* [2004], *Williams et al.* [2004], and *Beavan* [2005], can also be used. The main difference is a scaling of the amplitudes. Therefore the flicker noise variances we use here are roughly one half the size of those quoted in these papers.

[25] A random walk process is derived by integrating white noise. Random walk noise is supposed to be zero at initial time t_0. For evenly spaced data, the random walk cofactor matrix $\mathbf{Q}_{rw}$ is expressed as

$$\mathbf{Q}_{rw} = f^{-1}_s \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & \cdots & m \end{bmatrix}; \ f_s = \frac{m-1}{T} \quad (13)$$

where f_s is the sampling frequency in year^{-1}, and T is the total observation span [*Johnson and Wyatt*, 1994; *Zhang et al.*, 1997; *Mao et al.*, 1999].

[26] The variance components σ_w^2, σ_f^2, and σ_{rw}^2 can now be estimated using the LS-VCE method.

4.3. Estimation of Time Correlation

[27] Let us now consider a stationary noise process. This process does not contain the random walk noise, i.e., $\sigma_{rw} = 0$ in equation (11). We consider a side-diagonal structure for the covariance matrix $\mathbf{Q}_y$. This implies that the correlation between time series observations is only a function of time-lag $\tau = |t_j - t_i|$, i.e., $\sigma_{ij} = \sigma_\tau$. This structure is in fact similar to that of flicker noise introduced in equation (12). The only difference is that for flicker noise we have one variance as a scale, namely, σ_f^2, but here we employ one (unknown) covariance for each time-lag, namely, m components all together. The covariance matrix can then be written as a linear combination of m cofactor matrices

$$\mathbf{Q}_y = \sigma^2 \mathbf{I} + \sum_{\tau=1}^{m-1} \sigma_\tau \mathbf{Q}_\tau \tag{14}$$

where, for each value of τ, the $m \times m$ cofactor matrix $\mathbf{Q}_\tau$ has only two parallel side-diagonals of *ones* located on both sides of the main diagonal and τ steps away.

[28] We can now apply the general LS-VCE approach to the special case of estimating time correlation in the time series. If we measure a functionally known quantity ($\mathsf{E}(\underline{y}) = \mu_y$ in equation (2)), it can be shown that the (co)variances are simply estimated as [*Teunissen and Amiri-Simkooei*, 2007]

$$\underline{\hat{\sigma}}_\tau = \frac{\sum_{i=1}^{m-\tau} \hat{\underline{e}}_i \hat{\underline{e}}_{i+\tau}}{m - \tau}, \quad \tau = 0, 1, \ldots, m-1 \tag{15}$$

where the scalar $\hat{\underline{e}}_i$ is the ith element of the residual vector $\hat{\underline{e}} = \mathbf{P}_A^\perp \underline{y}$. When $\tau = 0$, the preceding equation gives the well-known expression for the estimator of the variance $\underline{\hat{\sigma}}_0 = \underline{\hat{\sigma}}^2$. Equation (15) is identical to the so-called unbiased estimator for the autocorrelation function (ACF) of stationary zero-mean least squares residuals [*Priestley*, 1981]. The biased estimate obtained from MLE uses m instead of $m - \tau$ in the denominator of equation (15).

[29] One can also compute the correlation coefficients that together represent the empirical autocorrelation function (ACF)

$$\hat{\underline{\rho}}_\tau = \frac{\hat{\underline{\sigma}}_\tau}{\underline{\hat{\sigma}}^2}, \quad \tau = 1, 2, \ldots, m-1 \tag{16}$$

Application of the error propagation law to the linearized form of the preceding equation gives

$$\sigma_{\hat{\rho}_\tau}^2 = \frac{1}{m - \tau} + \frac{2\rho_\tau^2}{m} \tag{17}$$

which shows that the precision of the autocorrelation function gets poorer when τ increases. This means that the

correlation of long-memory processes is poorly estimable for large time-lags τ. Therefore when there exists a predefined noise process like power law noise, one may however still prefer to use that structure which is also readily possible with LS-VCE, and the above formulation of time correlation remains to get a general impression of the noise behavior (see section 6.3.1).

5. Model Identification

5.1. Least Squares Harmonic Estimation (LS-HE)

[30] In this section, we aim to determine an adequate design matrix, $\mathbf{A}$, for the functional model through parameter significance testing. For a time series $\underline{y}^T = [\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_m]$ defined on R^m, we assume that it can be expressed as a linear trend plus a sum of q individual trigonometric terms, i.e., $\mathsf{E}(\underline{y}(t)) = y_0 + rt + \sum_{k=1}^q a_k \cos \omega_k t + b_k \sin \omega_k t$ (see equation (9)). In matrix notation, we may write

$$\mathsf{E}(\underline{y}) = \mathbf{A}x + \sum_{k=1}^q \mathbf{A}_k x_k, \quad \mathsf{D}(\underline{y}) = \mathbf{Q}_y \tag{18}$$

where the design matrix $\mathbf{A}$ contains two columns of the linear regression terms and the matrix $\mathbf{A}_k$ consists of two columns corresponding to the frequency ω_k of the sinusoidal function

$$\mathbf{A}_k = \begin{bmatrix} \cos \omega_k t_1 & \sin \omega_k t_1 \\ \cos \omega_k t_2 & \sin \omega_k t_2 \\ \vdots & \vdots \\ \cos \omega_k t_m & \sin \omega_k t_m \end{bmatrix} \text{ and } x_k = \begin{bmatrix} a_k \\ b_k \end{bmatrix} \tag{19}$$

with a_k, b_k, and ω_k being (un)known real numbers. On the one hand, if the frequencies ω_k are known, one will deal with the most popular (linear) least squares problem to estimate amplitudes a_k and b_k's. *Petrov and Ma* [2003] studied harmonic position variations of 40 VLBI stations at 32 known tidal frequencies. They found that the estimates of station displacements generally agree with the ocean loading computed on the basis of modern ocean tide models for the main diurnal and semidiurnal tides. On the other hand, if the frequencies ω_k are unknown, the problem of finding these unknown parameters is the task of least squares harmonic estimation.

[31] The problem now is to find the set of frequencies $\omega_1, \cdots, \omega_q$, and in particular the value q, in equation (18). The following null and alternative hypotheses are put forward (to start, set $i = 1$):

$$H_0 : \mathsf{E}(\underline{y}) = \mathbf{A}x + \sum_{k=1}^{i-1} \mathbf{A}_k x_k \tag{20}$$

versus

$$H_a : \mathsf{E}(\underline{y}) = \mathbf{A}x + \sum_{k=1}^{i} \mathbf{A}_k x_k \tag{21}$$

The detection and validation of ω_i is completed through the following two steps:

[32] Step I: The goal is to find the frequency ω_i (and correspondingly $\mathbf{A}_i$) by solving the following minimization problem:

$$\underline{\omega}_i = \arg\min_{\omega_j} \left\| \mathbf{P}^\perp_{[\overline{\mathbf{A}}\ \mathbf{A}_j]} \underline{y} \right\|^2_{\mathbf{Q}_y^{-1}} = \arg\min_{\omega_j} \|\underline{\hat{e}}_a\|^2_{\mathbf{Q}_y^{-1}} \qquad (22)$$

where $\|.\|^2_{\mathbf{Q}_y^{-1}} = (.)^T \mathbf{Q}_y^{-1}(.)$, $\overline{\mathbf{A}} = [\mathbf{A}\ \mathbf{A}_1 \ldots \mathbf{A}_{i-1}]$ and $\underline{\hat{e}}_a$ is the least squares residuals under the alternative hypothesis. The matrix $\mathbf{A}_j$ has the same structure as $\mathbf{A}_k$ in equation (19); the one that minimizes the preceding criterion is set to be $\mathbf{A}_i$. The above minimization problem is equivalent to the following maximization problem [Teunissen, 2000, p. 96]:

$$\underline{\omega}_i = \arg\max_{\omega_j} \left\| \mathbf{P}_{\overline{\mathbf{A}}_j} \underline{y} \right\|^2_{\mathbf{Q}_y^{-1}}, \text{ with } \overline{\mathbf{A}}_j = \mathbf{P}^\perp_{\overline{\mathbf{A}}} \mathbf{A}_j \qquad (23)$$

where $\mathbf{P}^\perp_{\overline{\mathbf{A}}} = \mathbf{I} - \overline{\mathbf{A}}(\overline{\mathbf{A}}^T \mathbf{Q}_y^{-1} \overline{\mathbf{A}})^{-1} \overline{\mathbf{A}}^T \mathbf{Q}_y^{-1}$ and $\mathbf{P}_{\overline{\mathbf{A}}_j} = \overline{\mathbf{A}}_j (\overline{\mathbf{A}}_j^T \mathbf{Q}_y^{-1} \overline{\mathbf{A}}_j)^{-1} \overline{\mathbf{A}}_j^T \mathbf{Q}_y^{-1}$. The preceding equation simplifies to

$$\underline{\omega}_i = \arg\max_{\omega_j} \underline{\hat{e}}_0^T \mathbf{Q}_y^{-1} \mathbf{A}_j \left(\mathbf{A}_j^T \mathbf{Q}_y^{-1} \mathbf{P}^\perp_{\overline{\mathbf{A}}} \mathbf{A}_j \right)^{-1} \mathbf{A}_j^T \mathbf{Q}_y^{-1} \underline{\hat{e}}_0 \qquad (24)$$

with $\underline{\hat{e}}_0 = \mathbf{P}^\perp_{\overline{\mathbf{A}}} \underline{y}$ the least squares residuals under the null hypothesis. In case that the time series contains only white noise, namely, $\mathbf{Q}_y = \sigma^2 \mathbf{I}$, it follows that

$$\underline{\omega}_i = \arg\max_{\omega_j} \underline{\hat{e}}_0^T \mathbf{A}_j \left(\mathbf{A}_j^T \mathbf{P}^\perp_{\overline{\mathbf{A}}} \mathbf{A}_j \right)^{-1} \mathbf{A}_j^T \underline{\hat{e}}_0 \qquad (25)$$

[33] Analytical evaluation of the above maximization problem is complicated. In practice, one has to be satisfied with numerical evaluation. A plot of spectral values $\|\mathbf{P}_{\overline{\mathbf{A}}_j} \underline{y}\|^2_{\mathbf{Q}_y^{-1}}$ versus a set of discrete values for ω_j can be used as a tool to investigate the contribution of different frequencies in the construction of the original time series. That is, we can compute the spectral values for different frequencies using equations (24) or (25). The frequency at which $\|\mathbf{P}_{\overline{\mathbf{A}}_j} \underline{y}\|^2_{\mathbf{Q}_y^{-1}}$ achieves its maximum value is used to construct $\mathbf{A}_i$.

[34] Step II: To test H_0 against H_a in equation (21), we consider $\mathbf{Q}_y = \sigma^2 \mathbf{I}$, with σ^2 unknown. The following test statistic can be used [see Teunissen et al., 2005]

$$\underline{T}_2 = \frac{\left\| \mathbf{P}_{\overline{\mathbf{A}}_i} \underline{y} \right\|^2_{\mathbf{Q}_y^{-1}}}{2\underline{\hat{\sigma}}_a^2} = \frac{\underline{\hat{e}}_0^T \mathbf{A}_i \left(\mathbf{A}_i^T \mathbf{P}^\perp_{\overline{\mathbf{A}}} \mathbf{A}_i \right)^{-1} \mathbf{A}_i^T \underline{\hat{e}}_0}{2\underline{\hat{\sigma}}_a^2} \qquad (26)$$

where $\overline{\mathbf{A}}_i = \mathbf{P}^\perp_{\overline{\mathbf{A}}} \mathbf{A}_i$ and the estimator for the variance, $\underline{\hat{\sigma}}_a^2$, has to be computed under the alternative hypothesis. Under H_0, the test statistic has a central Fisher distribution

$$\underline{T}_2 \sim \mathbf{F}(2, m-n-2i) \qquad (27)$$

The above hypothesis testing is in fact the parameter significance test because the test statistic $\underline{T}_2$ can also be expressed in terms of $\hat{x}_i$ in equation (21) and its covariance matrix. If the null hypothesis is rejected, we can increase i by one step and perform the same procedure for finding yet another frequency. As a generalization of the Fourier spectral analysis, the method is neither limited to evenly spaced data nor to integer frequencies.

[35] When measuring a functionally known quantity, equation (25) reads $\omega_i = \arg\max_{\omega_j} \underline{\hat{e}}_0^T \mathbf{A}_j (\mathbf{A}_j^T \mathbf{A}_j)^{-1} \mathbf{A}_j^T \underline{\hat{e}}_0$, which is identical to the least squares spectral analysis (LSSA) method developed by Vaníček [1969]. For some applications, refer to Craymer [1998], Abbasi [1999], Asgari and Harmel [2005], and Amiri-Simkooei [2005]. Amiri-Simkooei and Tiberius [2007] assess the noise characteristics of GPS receivers using zero-baseline time series and arrive at the same conclusions using short baseline time series when multipath effects were captured by a set of harmonic functions. With equations (18)–(21) and (24), we now have a formulation for LSSA even when an initial design matrix $\mathbf{A}$ is present in the model and the covariance matrix $\mathbf{Q}_y$, in general, is not a scaled identity matrix.

[36] Our application of harmonic estimation, in the first place, is to find any potential periodicities in the series. The remaining unmodelled effects (for example, power law noise) will also be interpreted and captured by a set of harmonic functions. Once we compensate for these effects in the functional model, the remaining noise characteristics of the series will be assessed. A nearly white noise combined with autoregressive noise can be shown to remain in the data series.

5.2. The w-Test Statistic

[37] Here we aim to determine the appropriate covariance matrix $\mathbf{Q}_y$ through significance testing of the stochastic model. One advantage of LS-VCE over other methods is that one can use statistical hypothesis testing in the stochastic model (similar as done with the functional model). When there is no misspecification in the functional part of the model $\mathsf{E}(\underline{y}) = \mathbf{A}x$, the following two hypotheses, as an example, are considered:

$$\mathrm{H}_0: \mathbf{Q}_y = \sigma_w^2 \mathbf{I} \quad \text{versus} \quad \mathrm{H}_a: \mathbf{Q}_y = \sigma_w^2 \mathbf{I} + \mathbf{C}_y \delta \qquad (28)$$

where $\mathbf{C}_y$ is a known cofactor matrix, for example, $\mathbf{Q}_f$ or $\mathbf{Q}_{rw}$, and δ is an unknown (co)variance parameter. We can use the generalized likelihood ratio test for testing H_0 against H_a. If we do so, the following w-test statistic can be obtained [Amiri-Simkooei, 2007]:

$$\underline{w} = \frac{b\,\underline{\hat{e}}^T \mathbf{C}_y \underline{\hat{e}} - \mathrm{tr}(\mathbf{C}_y \mathbf{P}^\perp_{\mathbf{A}}) \underline{\hat{e}}^T \underline{\hat{e}}}{\sigma_w^2 \left[2b^2\, \mathrm{tr}(\mathbf{C}_y \mathbf{P}^\perp_{\mathbf{A}} \mathbf{C}_y \mathbf{P}^\perp_{\mathbf{A}}) - 2b\, \mathrm{tr}(\mathbf{C}_y \mathbf{P}^\perp_{\mathbf{A}})^2 \right]^{1/2}} \qquad (29)$$

with $b = m - n$ the redundancy of the functional model and $\underline{\hat{e}}$ the least squares residuals under the null hypothesis. The orthogonal projector $\mathbf{P}^\perp_{\mathbf{A}}$ is also given under the null hypothesis.

[38] The expectation and the variance of the w-test statistic are 0 and 1, respectively. The distribution of this statistic, for large m, can be approximated by the standard normal distribution. The goal now is to compute the w-test statistic values for different alternative hypotheses, i.e., different $\mathbf{C}_y$'s in the preceding equation, and to select the one that gives the maximum value for the w-test. In fact,
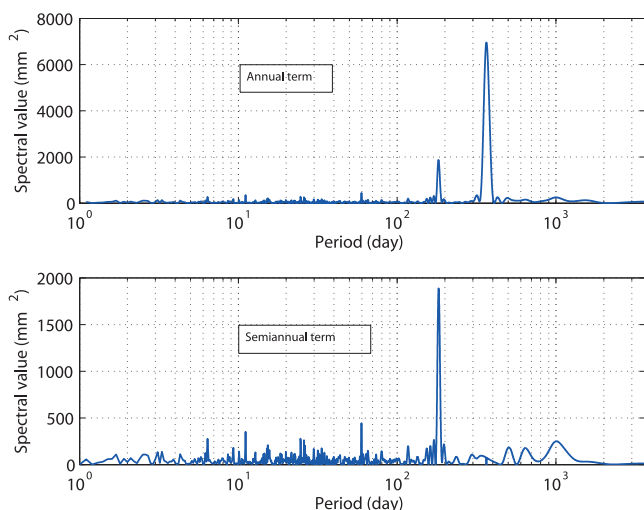
Figure 1. Estimated spectral values, equation (25), of simulated data series to detect annual and semiannual signals; top gives annual term (364.7 days), bottom semiannual term (182.4 days).

equation (29) provides us with an objective measure to judge, whether or not (or which), additional noise processes are likely to be present in the data at hand. Because of the special structure of the above hypotheses, the numerical evaluation of the preceding test statistic is very simple. We do not need for instance to invert a full covariance matrix since it is diagonal under the (assumed) null hypothesis, namely, $\mathbf{Q}_y = \sigma_w^2 \mathbf{I}$.

5.3. Demonstration Using Simulated Data

[39] To illustrate how the proposed LS-HE and LS-VCE work, we simulated a 10-year time series (daily samples) containing only white noise with a standard deviation of 5 mm. Two sinusoidal functions with amplitude of 2 and 1 mm, respectively, for the annual and semiannual term, have then been added to the data. We now use LS-HE to find the frequencies (or periods) of these signals. This was repeated 100 times and it follows that the empirical standard deviation of the detected periods is 1.4 and 0.7 days for the annual and semiannual term, respectively. Figure 1 shows one typical example of application of the method to find the periods of harmonic functions. In the first step, the annual term is detected and in the second step the semiannual term.

[40] A correlogram portrays the autocorrelation versus time-lag τ. Figure 2 shows the typical example of the simulated data corresponding to Figure 1. In each graph, the top window is the time series itself, the middle shows the running averages, and the bottom gives the autocorrelation coefficients obtained from LS-VCE using equation (16). In case of pure white noise, the autocorrelation function behaves randomly around zero. When both annual and semiannual terms are added, the autocorrelation function (ACF) shows a periodic behavior that resembles the periodicity of the annual signal. This makes sense because the ACF of a sinusoidal wave is again a sinusoidal wave with the same frequency. But the amplitude is proportional to the square of the amplitude of the original signal [*Priestley*, 1981]. If one removes the annual term, the ACF will still

show a periodicity which is due to the presence of the semiannual signal. When one also removes this signal, the ACF becomes very similar to the case of pure white noise.

[41] One can also compute the values of the w-test statistic using equation (29) for the different cases mentioned above. The cofactor matrix is chosen as that of flicker noise $\mathbf{C}_y = \mathbf{Q}_f$. Based on the simulation of 20 data sets, the w-test values on average become as follows: in presence of annual and semiannual signals $w = 15.3$; removed annual signal $w = 1.9$; and removed both annual and semiannual signals $w = 0.3$. Using LS-VCE, white and flicker noise amplitudes were estimated. The amplitudes on average are as follows: in presence of both signals $\sigma_w = 4.85$ mm and $\sigma_f = 2.84$ mm; removed annual term $\sigma_w = 4.98$ mm and $\sigma_f = 1.35$ mm; and removed both terms $\sigma_w = 5.00$ mm and $\sigma_f = 0.07$ mm. These all together simply express that if there exist unmodelled effects in data, they can be misinterpreted as time correlation (here flicker noise). One should therefore take care of these signals in the functional model.

6. Numerical Results and Discussions

6.1. Data and Model Description

[42] Global time series of site positions are supposed to have more noise than those from a regional solution [*Williams et al.*, 2004]. In this study, the daily GPS global solutions of different stations processed by the GPS Analysis Center at JPL are adopted. The data were processed using the precise point positioning method in the GIPSY software [*Zumberge et al.*, 1997]. The satellite orbits, satellite clocks, and Earth rotation parameters (ERP) used for the daily solutions were estimated with data from 42 globally distributed IGS tracking stations [see *Beutler et al.*, 1999]. In addition, corrections for geophysical effects such as pole and ocean tide effects have been applied. The reader is referred to the JPL Web site [http://sideshow.jpl.nasa.gov/mbh/series.html].

[43] The estimated coordinates of a site are uncorrelated with those of the other sites if the effects of the common errors in the satellite orbits and clocks and ERP on the estimated coordinates are insignificant. To make a proper statement, one can rely on multivariate time series analysis methods. The time series are processed on a component-by-component basis in this study.

[44] Most of the results given are based on five stations, namely, KOSG, WSRT, ONSA, GRAZ, and ALGO. Four stations are in Europe of which KOSG and WSRT in the Netherlands, ONSA in Sweden, and GRAZ in Austria. ALGO is in Canada. We have used 10 years of daily solutions for all sites except WSRT, which covers only 6.5 years. To justify some of the statements that we will make, 71 globally distributed GPS stations were also processed. Our point of departure is the original time series and its linear model of observation equations $y_t = y_0 + rt$. In most cases, the annual and semiannual signals have been considered as well. At times, we have included a set of harmonic functions to compensate for (parts of) unmodelled effects in the series.

6.2. Variance Component Analysis

[45] Three stochastic models have been chosen to describe the noise characteristics of GPS coordinate time
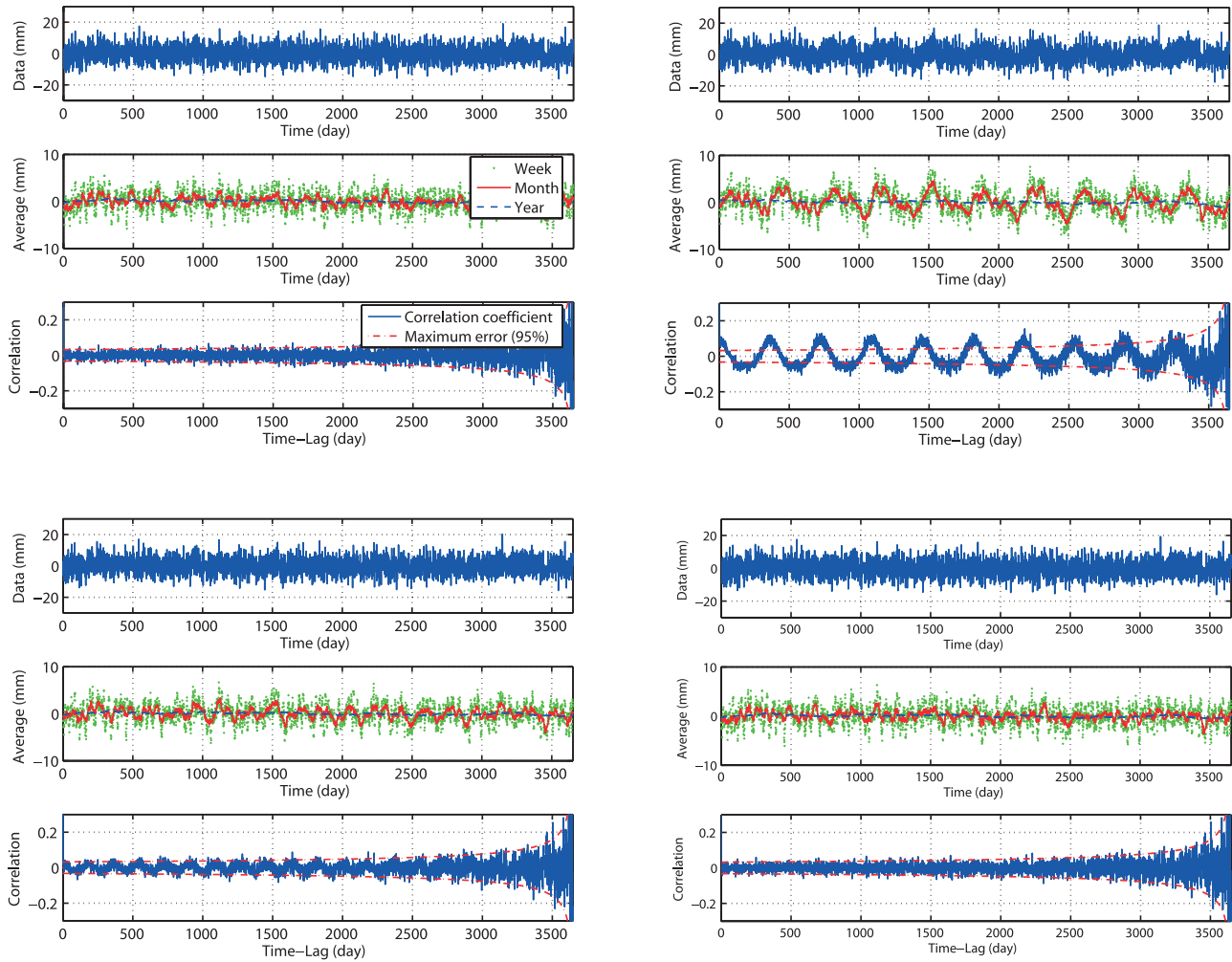
Figure 2. Simulated data series with their running averages (over 1 week, 1 month, and 1 year) as well as their autocorrelation coefficients for white noise only (top-left); added two sinusoidal functions with annual and semiannual periods and amplitudes of 2 and 1 mm, respectively (top-right); removed annual term obtained from LS-HE method (bottom-left); and removed both annual and semiannual terms obtained from LS-HE method (bottom-right).

series. They include the pure white noise model (I), the white plus flicker noise model (IIa), and the white plus random walk noise model (III). We employed LS-VCE to estimate the white noise, flicker noise, and random walk noise amplitudes (see equation (11)). Williams' investigations (S. D. P. Williams, Proudman Oceanographic Laboratory, personal communication, 2006) show that LS-VCE gives the same results as MLE. This holds in fact if $m \gg n$,

Table 1. White Noise, Flicker Noise, and Random Walk Noise Amplitude Estimates for North, East, and Vertical Components of Site Time Series in Three Different Stochastic Models (White Noise Only (Model I); White Noise Plus Flicker Noise (Model IIa); and White Noise Plus Random Walk Noise (Model III)); Functional Model Is Linear Regression Model

	I: WN (mm)			IIa: WN (mm) + FN (mm)						III: WN (mm) + RW (mm/year$^{1/2}$)					
Site Code	N	E	V	N	E	V	N	E	V	N	E	V	N	E	V
KOSG	3.34	3.44	7.45	2.45	2.54	5.21	3.41	3.67	8.82	2.79	2.89	6.17	7.96	9.06	23.4
StD[a]	0.04	0.04	0.09	0.05	0.05	0.12	0.18	0.18	0.40	0.04	0.04	0.08	0.51	0.56	1.31
WSRT	2.76	2.82	7.12	2.12	2.35	5.08	2.58	2.37	8.24	2.32	2.52	5.95	6.21	5.11	21.3
StD[a]	0.04	0.04	0.11	0.05	0.05	0.14	0.18	0.19	0.47	0.04	0.04	0.10	0.51	0.48	1.52
ONSA	3.35	3.65	7.85	2.54	2.70	5.28	3.39	3.45	9.62	2.85	3.02	6.29	8.25	7.50	27.6
StD[a]	0.04	0.04	0.10	0.05	0.05	0.13	0.18	0.19	0.42	0.04	0.04	0.09	0.53	0.52	1.45
GRAZ	3.74	4.75	9.12	2.81	3.62	6.53	3.62	4.16	9.90	3.14	3.97	7.51	8.26	9.12	25.7
StD[a]	0.05	0.06	0.11	0.06	0.07	0.15	0.20	0.25	0.49	0.04	0.05	0.10	0.55	0.65	1.52
ALGO	3.62	3.60	8.22	2.32	2.77	5.20	3.87	3.67	9.71	2.74	3.08	6.29	9.53	9.38	25.6
StD[a]	0.04	0.04	0.10	0.05	0.06	0.13	0.17	0.19	0.40	0.04	0.04	0.09	0.55	0.58	1.37

[a]Standard deviation of estimate.

Table 2. Difference in Log-Likelihood Values for Model With White Noise Plus Flicker Noise (WN + FN) Versus White Noise Plus Random Walk Noise (WN + RW), Both Compared With Pure White Noise (Log = 0)

	Error Model					
	IIa: WN + FN			III: WN + RW		
Site Code	North	East	Vertical	North	East	Vertical
KOSG	382	326	351	332	277	277
WSRT	229	138	227	213	129	182
ONSA	310	423	396	264	392	328
GRAZ	368	399	376	331	371	320
ALGO	682	290	598	620	253	535
Mean	394	315	390	352	284	328

which is usually the case in time series analysis. Table 1 gives the noise amplitudes of different components for different stochastic models. The table also provides the precision (standard deviation) of the estimates using equation (7). We find, for different noise components, that the horizontal components are less noisy than the vertical components by a factor of 2–4. Compared to the white noise model only, the amplitude of white noise for the white plus flicker noise model is 30% smaller, while this reduction for white plus random walk noise model is about 20%.

[46] In a similar manner to *Zhang et al.* [1997] and *Williams et al.* [2004], we produced the difference in the log likelihood values for each site, each component, and each error model. The results are given in Table 2. The values given in this table are normalized such that the pure white noise model has a log likelihood of zero. These results confirm that the white noise plus flicker noise model seems to be preferred over the pure white noise model or the white noise plus random walk noise model, which coincide with the findings of Williams et al. for global solutions. We will give the results of the *w*-test statistic in section 6.4.2 and show to which extent they are different from those obtained by MLE.

[47] The number of iterations in VCE methods is in fact an indication of (the lack of) appropriateness of the selected stochastic model. Figure 3 gives typical examples of estimated variance components at each iteration step for two variance

component models. The graphs show that the flicker and random walk noise variances systematically converge to their final estimates from one side. This confirms the presence of misspecification in the model, which will in fact result in overestimated (biased upward) flicker and random walk noise variances. This coincides with the findings of *Langbein* [2004] on EDM data (conversely white noise is biased downward). This is mainly due to the presence of ignored noise at high frequencies which leaks, in a systematic way, into lower frequencies (see section 6.4).

[48] The overestimation can also be verified when we compute the position errors and compare them with the scatter of the time series themselves (see section 6.6). The overestimation of random walk noise is more significant than that of flicker noise. This also means that the white plus flicker noise model is the preferred model in these circumstances. In other words, if the correlated noise in the time series is flicker noise and one tries to estimate the amplitude of the random walk noise, then the result will be biased (Williams, personal communication, 2006; see also section 3.3).

[49] These discussions essentially mean that there are still misspecifications (in fact underparameterization) in the model. In the functional model for instance, one should take care of any potential periodicities in the series (see section 6.3). Also, a better stochastic model may include in addition to power law noise other noise models like autoregressive noise (see section 6.4). For EDM observations, *Langbein* [2004] proposed to use a combination of power law noise and band-pass-filtered noise. The upward bias of power law noise and downward bias of white noise can thus be circumvented by introducing a more sophisticated functional and stochastic model.

6.3. Functional Model
6.3.1. Simple and Intuitive Technique
[50] Seasonal variations in site positions consist of signals from various geophysical sources and systematic modeling errors [*Dong et al.*, 2002]. The weekly, monthly, and yearly mean residuals calculated from averaging the daily residuals are shown in Figure 4 (left) for ALGO. Running averages
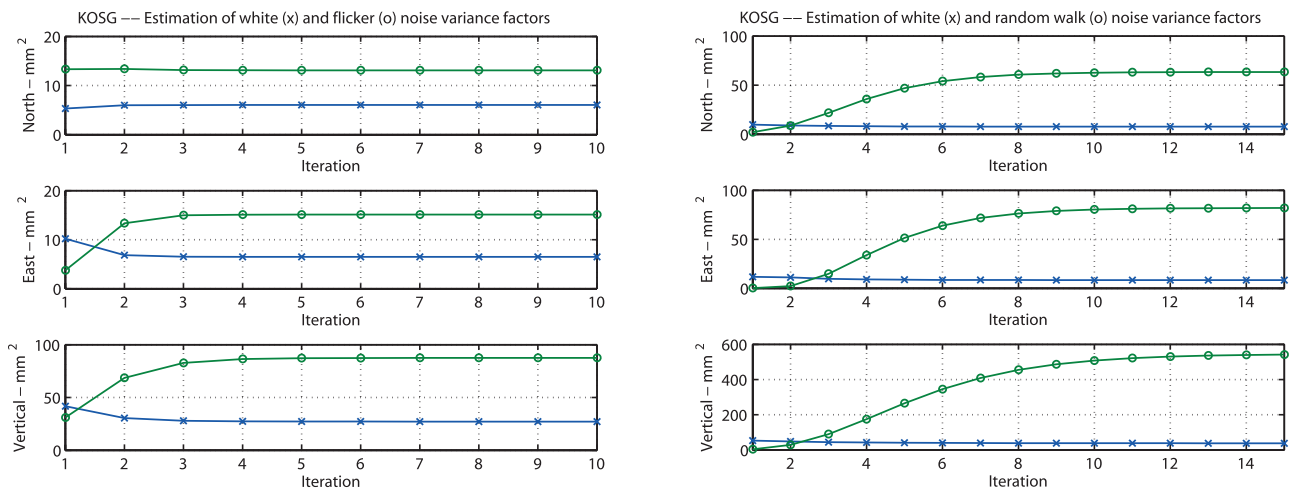

Figure 3. Estimated variance components at each iteration step for KOSG; combination of white plus flicker noise (left); and combination of white plus random walk noise (right).
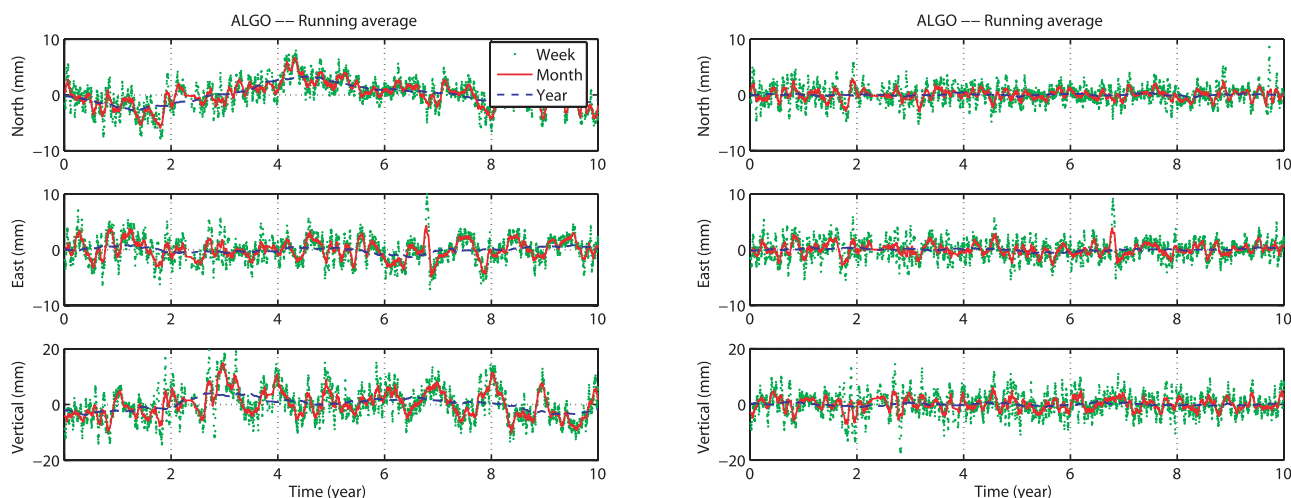
Figure 4. Weekly, monthly, and yearly mean residual series of north, east, and vertical components for two functional models (site ALGO); linear trend only (left); and linear trend plus 10 harmonic functions (right).

naturally remove the high-frequency noise and leave the lower frequency signals. Annual and seasonal variations can be observed in the running averages. For example, the vertical component shows a clear annual signal. They should be eliminated from the time series in order to obtain a more realistic assessment of the noise behavior. We can also see some high- and low-frequency fluctuations, which can likely be captured by flicker noise and random walk noise, respectively (see east and north components, respectively). When parts of these variations have a deterministic behavior, they should be compensated for in the functional rather than the stochastic model. *Ding et al.* [2005] tried to interpret this behavior as some interannual signals.

[51] We now focus on time correlation in the series and estimate one covariance for each time-lag τ using equations (15) and (16). Figure 5 (top) shows the autocorrelation coefficients for the time series of the each component of two sites. The annual and seasonal variations as well as long-term fluctuations can be seen in the correlograms. The variations are clearer here than those for the running averages in Figure 4. For example, the periodic behavior in the ALGO vertical correlogram shows the annual signal in the series. When the annual signal was included in the functional model, the annual periodicity of the correlogram disappeared. However, this was not the case for KOSG and ALGO east components, which show an annual-like signal. This implies that there might still be some hidden periodicities in the data series.

6.3.2. Harmonic Estimation

[52] Figure 6 (left) shows the test statistic values given by equation (26) to find the first 15 frequencies. The step size used for $T_j = \frac{2\pi}{\omega_j}$ is taken small at high frequencies and gets larger at lower frequencies. We can see that the value for the test statistic levels off quickly. With 6–10 harmonic functions, it gets close to the critical value. In all subsequent results, the number of harmonic functions q in equation (18) was set to 10 starting with just the offset and slope model. The combination of all 10 harmonic functions included in the functional model of the series is given in Figure 6 (right). The periods of 10 harmonic functions are given in

Table 3. Our opinion is that the periodic functions detected by the LS-HE method are due to the following four reasons:

[53] 1. Unmodelled periodic ground motion: The site is actually moving periodically in this case. Annual and semiannual signals, for instance, can be specified into this category. Except for a few components, both annual and semiannual signals can be seen in the series. A good example is the first period obtained for ALGO vertical component (366 days) that reveals the annual signal.

[54] 2. Periodic variation of the estimated time series: The site is apparently moving periodically. This is known as the aliasing effect. Unmodelled periodic systematic (for example, tidal) errors present at a station will result in spurious longer periodic systematic effects in the resultant time series [see *Penna and Stewart*, 2003; *Stewart et al.*, 2005]. A harmonic function with a period of 13.63 days is detected in the north components of WSRT and ALGO. In the east component of KOSG and WSRT, a period of 14.2 days is seen. There are also periods ranging from 170 to 180 days that coincide with those given by *Penna et al.* [2007] for unmodeled S2 ocean tide loading effect at different globally distributed sites.

[55] 3. Aliased multipath effect (still a challenging problem): We observe periodic patterns with periods of roughly 350, 175, 117, 88, 70, 59, 50, and 44 days. To justify this, the time series of 71 GPS stations were processed. Figure 7 shows the stacked (weighted) power spectra for these stations after including the annual and the semiannual signals. The peaks shown in the spectrum coincide well with the numbers given above. The set of stations was split into two parts and the same conclusion could again be drawn. The results also confirm *Ray*'s [2006] findings. Two possibilities which may lead to this effect are as follows: *Agnew and Larson* [2007] show that the repeat time of the GPS constellation, through which multipath can repeat at permanent stations, averages at 247 s less than a solar day (24 h). However, daily GPS position estimates are based on a full solar day. The difference will alias to a frequency of 0.0028565 cycles/day or 1.04333 cycles/year (350 days period). The periods found fit with this frequency and its
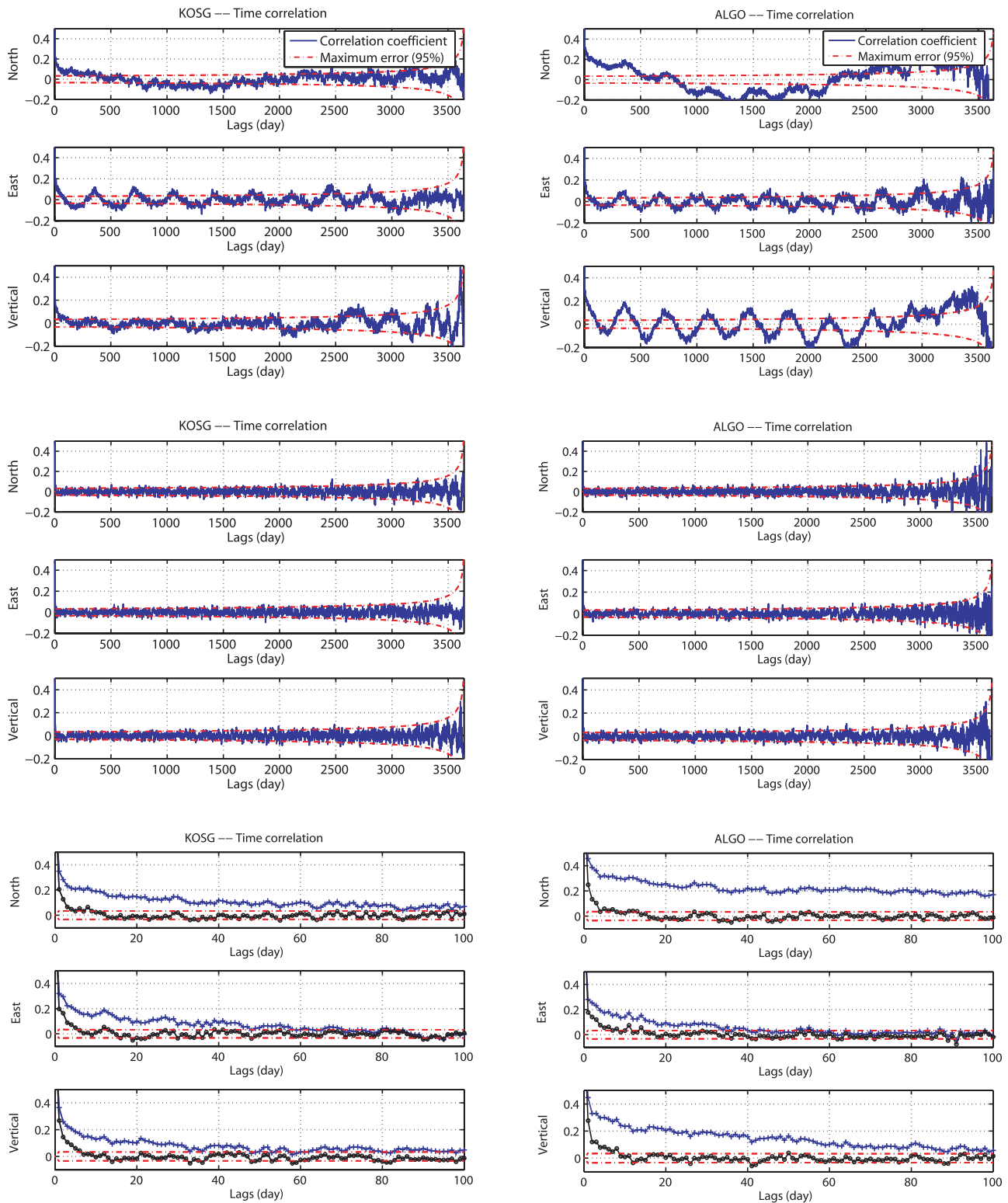
Figure 5. Autocorrelation coefficients for time series of north, east, and vertical components before (top) and after (middle) removing 10 harmonic functions; graphs in bottom provides a "zoom in" over the first 100 days for both before (+) and after (o) removing 10 harmonic functions.

harmonics. Periodic variations of the range residuals with maximum at the eclipse seasons indicate orbit modeling deficiencies for the GPS satellites [*Urschl et al.*, 2005, 2006]. The periods found above also coincide with the

period of one draconitic GPS year (about DJ = 351 days) and fractions DJ/n, n = 2, ..., 8 [see *Beutler*, 2006].

[56] 4. Presence of power law noise: There are still many numbers in the table (about 50%) that do not fit into the

Figure 6. The first 15 test statistic values obtained from equation (26) as well as its critical value (dashed line) from equation (27) with type I error of $\alpha = 0.0001$ (left); combination of 10 harmonic functions (unmodelled effects) as removed from original coordinate component time series (right).
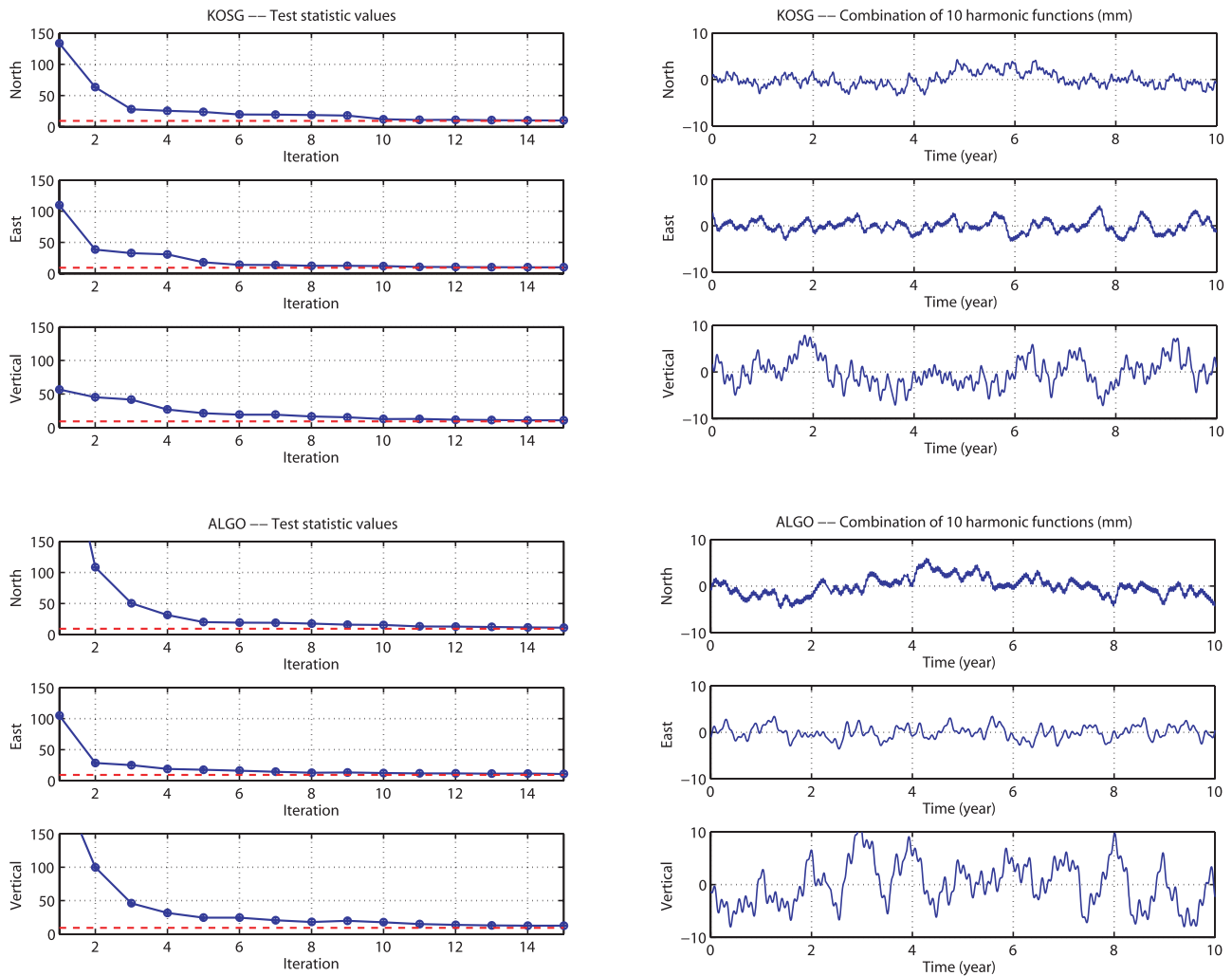
previous categories. Williams (personal communication, 2006) argues that the 10 harmonic functions uniformly distributed in log-frequency space would be sufficient to simulate power law noise. The higher frequency effects are likely due to flicker noise. Long-term period (for example, larger than 1000 days) effects are observed for most of the series. Parts of these effects can likely be considered as

random walk noise. Note also that undetected offsets in the time series can mimic random walk noise [*Williams*, 2003b]. We used so far equation (25) based on the pure white noise model to detect the frequencies. Therefore some of the detected periods are most likely due to the presence of colored noise in the data that has been ignored in the results of Table 3 and Figure 7 (left). The graph shows that white

Table 3. Periods (Days) of the First Ten Harmonic Functions Obtained from Least Squares Harmonic Estimation for North, East, and Vertical Components Using Equation (25) With $\mathbf{Q}_y = \sigma_w^2 \mathbf{I}$ and a Linear Regression Model

|	KOSG			WSRT			ONSA			GRAZ			ALGO		
Number	N	E	V	N	E	V	N	E	V	N	E	V	N	E	V
1	2620	354	885	1586	349	2035	3154	1418	337	2349	3651	910	2772	332	366
2	1256	384	377	354	168	863	1307	343	1467	175	1274	2505	378	177	3651
3	315	862	2782	171	265	712	235	2698	522	187	393	370	783	1483	1566
4	187	176	334	292	533	210	50.4	121	927	1282	792	325	1459	86.9	307
5	579	116	94.5	841	115	150	145	756	296	361	171	697	122	108	161
6	145	248	1401	71.0	59.3	114	278	260	215	387	180	117	13.6	74.5	88.3
7	69.8	97.0	87.5	84.2	188	37.4	185	116	191	50.4	506	278	200	762	43.9
8	175	70.4	37.6	190	70.2	88.7	173	70.6	41.5	144	148	88.0	86.2	527	923
9	50.4	264	503	50.2	14.2	30.3	372	394	3627	88.6	122	210	73.2	93.4	466
10	24.9	14.2	116	13.6	1694	66.6	57.4	237	94.7	822	343	37.5	182	43.9	120
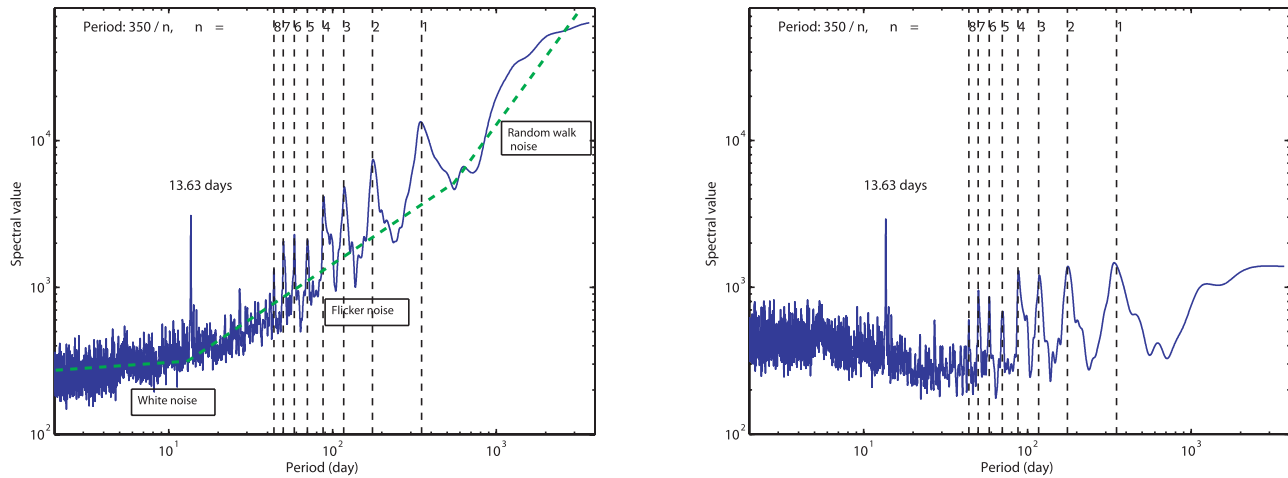
Figure 7. Stacked (weighted) power spectra of 71 permanent GPS stations (213 time series) for a linear regression model when annual and semiannual signals were removed; use of estimated pure white noise model (left); use of estimated short-memory plus flicker noise model—prewhitened time series (right). The word "weighted" also implies that the individual spectra have been scaled according to their estimated noise magnitudes (components).

noise is mainly present at high frequencies, flicker noise at medium frequencies, and random walk noise at low frequencies. To justify this, in the harmonic estimation, we used equation (24) with a more sophisticated noise model given by equation (30). Most of the lower frequency effects that were detected in the white noise model could not be detected here. Figure 7 (right) shows the stacked power spectra of 71 stations using this new stochastic model. The spectrum looks more or less flat and thus does not contradict our statement.

[57] Therefore to avoid biases in the estimate of x and also the amplitude of noise components (see section 3.3), one should take good care of any potential periodicities in the GPS position time series. We can at least mention the annual and the semiannual signals, signals with the periods of 13.66, 14.2, and 14.8 days, and most likely signals with the periods of 350 days and its fractions.

6.4. Stochastic Model

6.4.1. Simple and Intuitive Technique

[58] Figure 4 (right) shows the weekly, monthly, and yearly running average of residuals when 10 harmonic functions are included to describe unmodelled effects (site ALGO). Most of the signals and fluctuations have now been removed. One can also plot the autocorrelation coefficients for the corrected functional model. The correlograms of the time series are given in Figure 5 (middle). At first insight, they appear to represent white noise. The graphs at bottom provide a "zoom-in" on the first part of the graphs at the top and in the middle, i.e., the correlograms over the first 100 days. Unlike the original data, the autocorrelation coefficients become small already after a few days (max 10 days). Our impression is that this remaining high-frequency correlation can be caused by common and well-known sources of errors like atmospheric effects and satellite orbit errors. Table 4 provides the numerical results over the first 5 days. The correlation coefficients reduce approximately exponentially, for example, by $e^{-\alpha\tau}$, which is known as a first-order autoregressive noise process

AR(1). More results on this model are presented in the next section for the cases $\alpha = 1$ and $\alpha = 0.25$.

6.4.2. The w-Test Statistic

[59] The results of the w-test statistic are presented to find the most appropriate noise model for the global GPS coordinate time series. The larger (absolute value) the w-test statistic is, the more powerfully the null hypothesis tends to be rejected, and hence the more likely the alternative

Table 4. Estimated Time Correlation Over the First Five Time-Lags τ (Days) for Time Series of North, East, and Vertical Components Before (Left) and After (Right) Removing Harmonic Functions; Standard Deviation of all Estimates is 0.02

Site Code	τ	Correlation coefficient[a]			Correlation coefficient[b]		
		N	E	V	N	E	V
KOSG	1	0.35	0.32	0.37	**0.20**	**0.20**	**0.26**
	2	0.28	0.29	0.26	**0.12**	**0.16**	**0.14**
	3	0.23	0.22	0.23	0.06	0.09	0.10
	4	0.22	0.21	0.21	0.05	0.08	0.08
	5	0.21	0.19	0.18	0.03	0.06	0.06
WSRT	1	0.31	0.21	0.36	**0.13**	**0.07**	**0.25**
	2	0.27	0.22	0.26	**0.09**	**0.08**	**0.13**
	3	0.26	0.17	0.23	0.08	0.03	0.09
	4	0.24	0.18	0.20	0.06	0.05	0.06
	5	0.24	0.18	0.17	0.07	0.06	0.03
ONSA	1	0.31	0.33	0.37	**0.20**	**0.16**	**0.27**
	2	0.25	0.29	0.29	**0.13**	**0.11**	**0.18**
	3	0.21	0.26	0.25	0.08	0.08	0.13
	4	0.20	0.25	0.22	0.07	0.06	0.11
	5	0.18	0.23	0.20	0.04	0.04	0.08
GRAZ	1	0.32	0.32	0.35	**0.16**	**0.14**	**0.23**
	2	0.27	0.27	0.29	**0.10**	**0.08**	**0.16**
	3	0.23	0.28	0.26	0.06	0.09	0.13
	4	0.21	0.25	0.22	0.03	0.06	0.08
	5	0.23	0.24	0.20	0.06	0.05	0.07
ALGO	1	0.46	0.28	0.45	**0.25**	**0.18**	**0.27**
	2	0.39	0.25	0.33	**0.13**	**0.14**	**0.12**
	3	0.36	0.23	0.33	0.11	0.12	0.12
	4	0.31	0.19	0.30	0.04	0.07	0.07
	5	0.32	0.17	0.30	0.07	0.05	0.08

[a]Before removing harmonic functions from data.
[b]After removing 10 harmonic functions from data.

Table 5. The w-Test Statistic Values for Time Series of Position Estimate, Before (Left) and After (Right) Removing 10 Harmonic Functions Using Equation (25), for Different Alternative Hypotheses (Different $\mathbf{C}_y$'s)[a]

Site Code	Component	Before Removing Harmonic Functions						After Removing Harmonic Functions					
		w_1	w_2	w_3	w_4	w_5	w_6	w_1	w_2	w_3	w_4	w_5	w_6
KOSG	North	20.5	−8.1	**123.5**	**125.9**	27.5	40.0	12.1	−10.6	11.1	4.9	**14.7**	**16.8**
	East	18.8	−1.8	**32.9**	**9.3**	26.0	37.3	11.9	−5.1	12.9	9.1	**15.6**	**18.8**
	Vertical	21.5	0.3	**58.1**	**43.2**	27.7	37.6	15.9	0.6	10.4	1.1	**19.0**	**21.2**
WSRT	North	14.6	−11.0	**62.1**	**50.5**	20.4	31.7	6.5	−12.0	6.4	0.2	**8.4**	**11.0**
	East	10.0	−8.0	**25.4**	**10.9**	14.5	22.8	3.6	−9.8	6.4	2.4	**5.3**	**7.5**
	Vertical	17.2	−8.3	**61.2**	**58.4**	22.4	30.8	12.3	−8.5	15.1	6.4	**14.7**	**16.6**
ONSA	North	16.0	−8.2	**71.0**	**66.5**	21.3	35.7	12.1	−11.1	13.5	6.1	**14.9**	**18.0**
	East	12.3	−8.4	**67.7**	**52.8**	17.1	41.4	10.0	−11.2	11.3	3.2	**12.8**	**15.5**
	Vertical	17.7	−3.1	**43.1**	**32.7**	23.4	40.7	16.5	−4.5	16.9	4.5	**20.8**	**25.5**
GRAZ	North	18.7	−8.8	**119.8**	**119.2**	25.5	38.5	10.0	−10.6	7.1	0.3	**12.2**	**13.9**
	East	18.7	−4.5	**139.1**	**144.4**	26.0	41.0	8.5	−10.8	5.2	0.1	**10.7**	**13.5**
	Vertical	20.6	−5.9	**70.9**	**50.7**	27.8	40.2	13.7	−7.0	14.9	3.9	**17.4**	**21.5**
ALGO	North	27.5	−7.9	**285.1**	**350.1**	37.6	57.1	15.0	−8.7	13.1	3.3	**18.1**	**20.9**
	East	10.4	−5.4	**20.7**	**11.2**	14.4	34.4	10.7	−10.4	10.2	4.7	**14.2**	**18.2**
	Vertical	23.5	−2.7	**99.2**	**99.3**	31.1	52.5	16.7	−5.7	14.3	4.6	**19.4**	**22.0**

[a]w_1: $\mathbf{C}_y = \mathbf{Q}_\tau$, $\tau = 1$ contains only ones on two parallel side-diagonals next to the main diagonal; w_2: $\mathbf{C}_y = \text{diag}(\mathbf{Q}_{rw})$, only diagonal elements of $\mathbf{Q}_{rw}$ in in equation (13); w_3: $\mathbf{C}_y = \mathbf{Q}_f$, flicker noise structure introduced in equation (12); w_4: $\mathbf{C}_y = \mathbf{Q}_{rw}$, random walk noise structure introduced in equation (13); w_5: $\mathbf{C}_y = c_{ij}$ full matrix extracted from an exponential function of the form $e^{-\tau}$; w_6: $\mathbf{C}_y = c_{ij}$ full matrix extracted from an exponential function of the form $e^{-0.25\tau}$.

model will be as a candidate for the noise in the time series. Six stochastic models were tested using the hypotheses as in equation (28). The results are given in Table 5. For the original data (before removing harmonic functions), the maximum values are obtained for flicker noise and random walk noise models (columns w_3 and w_4). In addition, except for a few components (for example, ALGO north), flicker noise is preferred to random walk noise. When the values are very close (for example, ALGO vertical), both noise components (in addition to white noise) are likely to be present in the series. In other words, the power law noise is well described with a spectral index between −1 and −2.

[60] Usually the white noise along with either flicker noise or random walk noise are estimated. To confirm the w-test results, we included all three variances in the stochastic model, namely, white noise, flicker noise, and random walk noise as in equation (11). If a variance is negative, it is an indication that this noise model is most likely not the preferred model and can be excluded from the stochastic model. Table 6 shows the signs of the estimated variances using LS-VCE. In 53% of the cases, the random walk noise variance is negative. They are correspondingly related to the cases that the w-test values for flicker noise given in Table 5 are significantly larger than those for random walk noise. In 44% of the cases both flicker and random walk noise variances are positive. They are related to the case that the w-test values of flicker noise are approximately identical to those of random walk noise. Only for the north component of ALGO, the flicker noise variance is negative, which is also verified because the w-test value for flicker noise is smaller than that for the random walk noise.

[61] In columns 5 and 6, we have respectively used the following matrices for $\mathbf{C}_y$: $c_{ij} = e^{-\tau}$ and $c_{ij} = e^{-0.25\tau}$ with $i, j = 1, 2, \ldots, m$. The corresponding w-test statistic values are mostly significantly smaller than those for flicker and random walk noise. However, after removing 10 harmonic functions (individually per component) from the original data series, the largest values for w-test statistic are obtained for the $e^{-0.25\tau}$ stochastic model (column 6). Note

also that the results given in columns 5 and 6 are not much different. This therefore confirms the existence of remaining correlation at very high frequencies that is believed to be due to common sources of errors that last only a couple of successive days (see Figure 5, bottom). A relatively large value of the w-test statistic (values on the right) for flicker and random walk noise is most likely due to this correlation. A significant decrease in the w-test values for flicker and random walk noise implies that most parts of the power law noise have now been captured by the harmonic functions. This statement was also justified when the white and the flicker noise amplitudes were estimated using the extended functional model (with 10 harmonics), which led to small positive or negative values for flicker noise amplitudes.

[62] Let us now turn our attention to the second column (w_2) in Table 5. The goal is to test stationarity of the white noise amplitude in the series. For this purpose, we selected $\mathbf{C}_y = \text{diag}(\mathbf{Q}_{rw})$. Most of the w-test values are negative, implying that the white noise amplitude in the daily position estimates gets reduced toward the end of the series as $\sigma^2_{w_i} =$

Table 6. Sign of White Noise, Flicker Noise, and Random Walk Noise Variances for Different Components

Site Code	Component	Sign of Variance Component		
		WN	FN	RW
KOSG	North	+	+	+
	East	+	+	−
	Vertical	+	+	−
WSRT	North	+	+	−
	East	+	+	−
	Vertical	+	+	+
ONSA	North	+	+	+
	East	+	+	−
	Vertical	+	+	−
GRAZ	North	+	+	+
	East	+	+	+
	Vertical	+	+	−
ALGO	North	+	−	+
	East	+	+	−
	Vertical	+	+	+

$\sigma_w^2 + t_i \, \delta$ (δ negative). This reflects the improvements in analysis products (for example, satellite orbits and Earth orientation parameters), which makes sense, of course, as equipment is improving and also our knowledge about error sources like atmosphere and orbit is continuously improved. The reduction in noise amplitudes with time was shown by *Williams et al.* [2004]. They showed that such a reduction of noise also holds true for the flicker noise amplitude.

[63] A large value for the *w*-test leads to the rejection of the null hypothesis. One can obtain the *w*-test values for different alternative hypotheses. The one that gives the largest absolute value is considered as a superior candidate for describing the noise characteristics of the data. In our case, in general, the flicker noise model was preferred. After introducing 10 harmonic functions to the model, the largest values were obtained for the AR(1) noise process. The *w*-test statistic is considered to be a powerful tool to decide on the preferred noise model. Based on simulated data, Williams (personal communication, 2006) concluded that the *w*-test statistic and the difference in the log likelihood values give very similar results. Note, however, that the *w*-test statistic can simply be used while the MLE method needs successive inversion of the covariance matrix.

6.5. Remarks and Discussions

[64] We would like to point out here the duality between harmonic functions and time correlated noise. If one could compensate for all unmodelled effects in the functional model, this would be the best way to do so. Otherwise, they will be misinterpreted as if the data were time correlated; see section 3.3 and the simulated example in section 5.3. Examples of such hidden periodicities are the annual and the semiannual signals, signals with the periods of 13.66, 14.2, and 14.8 days, and likely signals with the periods of $350/n$ ($n = 1, \ldots, 8$).

[65] If the unmodelled variations cannot be considered as deterministic signals to be compensated for in the functional model (for example, by harmonic functions), they can be captured as power law noise process (for example, flicker noise or random walk noise) through the stochastic model. Although not fully physically justified, we captured unmodelled effects by a set of harmonic functions (in this study 10 individual functions for each series). We observed a duality between the functional and the stochastic model. When unmodelled effects were removed by these harmonic functions, a short-memory process is left in the time series. In other words, most parts of the power law noise have been captured by the harmonic functions. However, this does not necessarily mean that there exist 10 individual hidden periodic functions in each series.

[66] The remaining minor time correlation, as a short-memory process, exponentially vanishes within a few days and can be expressed for instance as an AR(1) noise process. This essentially means that to avoid biases in the power law noise amplitude due to underparameterization, one will have to include also the AR(1) noise process $\sigma_a^2 \mathbf{Q}_a$ in equation (11), namely, $\mathbf{Q}_y = \sigma_w^2 \mathbf{I} + \sigma_a^2 \mathbf{Q}_a + \sigma_f^2 \mathbf{Q}_f$ (when one ignores random walk noise). This holds indeed also for any potential periodicities in the functional part of the model.

[67] Our investigations show that the time series are not yet long enough to separately estimate one variance for each
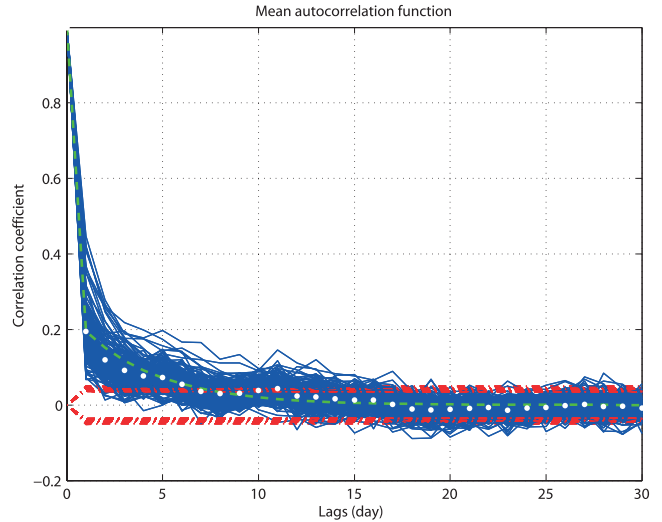

Figure 8. Individual autocorrelation functions (solid line) as well as mean autocorrelation function (white circles) of 71 permanent GPS stations (213 time series) for a linear regression model with 10 harmonic functions; light dashed line represents the estimated short-memory (SM = WN + AR(1)) noise process.

noise component. Therefore first the LS-HE method was employed to include a set of harmonic functions to compensate for power law noise model. The remaining noise is now expressed as a combination of white and autoregressive noise. The unknowns in this case are the amplitudes of white and autoregressive noise (σ_w and σ_a) and the timescale α of the noise process. In other words, the short-memory process is expressed as $\mathbf{Q}_y = \mathbf{Q}(\alpha; \sigma_w^2, \sigma_a^2) = \sigma_w^2 \mathbf{I} + \sigma_a^2 \mathbf{Q}_a$ where $q_{ij}^a = e^{-\alpha\tau}$ and $\tau = |t_j - t_i|$. This is in fact a nonlinear stochastic problem that can again be solved by LS-VCE.

[68] The method was applied to 71 globally distributed GPS stations. The average value for the timescale is $\alpha \approx 0.25$. The mean amplitude of white noise and autoregressive noise is $\sigma_w = 2.3, 3.3,$ and 6.3 and $\sigma_a = 1.3, 1.8,$ and 4.0 (all in mm) for north, east, and vertical components, respectively. Suppose now that we do not include the 10 harmonic functions to compensate for flicker noise. The covariance matrix is again of the form $\mathbf{Q}_y = \sigma_w^2 \mathbf{I} + \sigma_a^2 \mathbf{Q}_a + \sigma_f^2 \mathbf{Q}_f$. In practice, it is more convenient to combine white noise and autoregressive noise into one "short-memory" process using the average values obtained above. Based on these results, if one assumes that the timescale α and also the relative magnitude of noise components σ_a/σ_w is known, the covariance matrix $\mathbf{Q}_y = \sigma_w^2 \mathbf{I} + \sigma_a^2 \mathbf{Q}_a + \sigma_f^2 \mathbf{Q}_f$ can be reformulated as

$$\mathbf{Q}_y = \sigma_s^2 \mathbf{Q}_s + \sigma_f^2 \mathbf{Q}_f \qquad (30)$$

where $\sigma_s^2 = \sigma_a^2 + \sigma_w^2$ is the variance of the short-memory noise process and $\mathbf{Q}_s$ is given as

$$q_{ij}^s = \begin{cases} 1 & \text{if} \quad \tau = 0 \\ \beta e^{-\alpha\tau} & \text{if} \quad \tau \neq 0 \end{cases}. \qquad (31)$$

with $\alpha \approx 0.25$ and $\beta = \sigma_a^2 / (\sigma_w^2 + \sigma_a^2) \approx 0.25$. Figure 8 shows the weighted mean autocorrelation function of

Table 7. Short-Memory (WN plus AR(1)) Noise and Flicker Noise Amplitude Estimates for North, East, and Vertical Components in Modified Functional and Stochastic Model

	Short-Memory and Flicker Noise (Model IIb)					
	WN + AR(1) (mm)			FN (mm)		
Site Code	North	East	Vertical	North	East	Vertical
KOSG	2.85	3.06	6.36	2.23	1.61	5.21
StD[a]	0.05	0.05	0.12	0.25	0.27	0.57
WSRT	2.43	2.67	6.03	1.45	1.03	5.66
StD[a]	0.05	0.05	0.15	0.28	0.31	0.69
ONSA	2.96	3.10	6.32	2.06	2.21	6.99
StD[a]	0.05	0.06	0.14	0.26	0.27	0.57
GRAZ	3.26	4.09	7.75	2.02	2.72	6.07
StD[a]	0.06	0.07	0.15	0.29	0.34	0.69
ALGO	2.76	3.25	6.43	2.69	1.91	5.98
StD[a]	0.06	0.05	0.13	0.24	0.28	0.57

[a]Standard deviation of estimate.

71 permanent GPS stations and its approximation based on equation (31).

[69] Stochastic model (30) is referred to as the short-memory noise and flicker noise model (model IIb), for which two variance components σ_s^2 and σ_f^2 need to be estimated by LS-VCE. We now consider this equation to estimate the magnitude of short-memory (combined WN and AR(1)) and flicker noise process. A correct functional model consisting of annual and semiannual signals, a period of 13.66 days, and a period of 350 days and its fractions $350/n$ ($n = 2, \ldots, 8$) was also used. The results are given in Table 7. Compared to the results given in Table 1 for the white plus flicker noise model (model IIa), flicker noise shows a reduction of 40% whereas white (in fact short-memory) noise increases by about 20%. The differences in the log likelihood values have also been computed which show an increase of about 10% compared to the values given in Table 2 for model IIa. With this strategy, not only can one obtain a better precision for the parameters of interest, but also one will be able to increase the log likelihood values.

6.6. Error Estimate of Parameters

[70] The goal here is to estimate and compare the error of four parameters of interest, namely, intercept, slope (rate), position in the middle, and position at the far end of the time series, for different stochastic models. Model I, IIa, IIb, and III are the pure white noise, white noise plus flicker noise, short-memory (combined white and autoregressive) noise plus flicker noise, and white noise plus random walk noise, respectively. Model IIb also includes the annual and the semiannual signals, a period of 13.66 days, and periods of $350/n$ days, $n = 1, \ldots, 8$ in the functional model.

[71] We show how an incorrect stochastic model will result in a too optimistic (or too pessimistic) precision description of the parameters of interest (see section 3.3). The results given in Table 8 are based on $\mathbf{Q}'_x = (\mathbf{A}^T \mathbf{Q}'^{-1}_y \mathbf{A})^{-1}$, where $\mathbf{Q}'_y$ is a (in)correct covariance matrix obtained from model I, IIa, IIb, or III. Model I gives the most optimistic results and model III generally gives the least precise results. The error of parameters for different models compared to those for the pure white noise model is larger by the coefficients given in Table 9. For example, if a white plus flicker noise model is used instead of a pure white noise model, the velocity error obtained can be larger by factors of 9–16. Among different models, the error estimates of site velocity and position for model III are considerably larger (about one order of magnitude) than those for other models.

[72] Compared to the white noise magnitude of the series (scatter of the series), the standard deviations of positions in the middle of the time series are 2%, 80%, 50%, and 400% for models I, IIa, IIb, and III, respectively. These values increase at the end of the series to 4%, 90%, 55%, and 800%. For all models, except model III, the minimum error estimate of the position is obtained in the middle and the largest values are given at both ends of the series. The results of model I are too optimistic. The error contribution of the intercept and the slope on the position seems to be the same (2% in the middle and 4% at both ends). The results obtained from models IIa and IIb appear to be more realistic. In model IIa, the slope has only a contribution of 10% on the

Table 8. Error Estimate (Formal Standard Deviation) of Slope, Intercept, and Position for Different Stochastic Models[a]

		KOSG			WSRT			ONSA			GRAZ			ALGO		
Error in	Model	N	E	V	N	E	V	N	E	V	N	E	V	N	E	V
Intercept (mm)	I	0.12	0.12	0.26	0.12	0.13	0.32	0.12	0.13	0.28	0.13	0.17	0.32	0.13	0.13	0.29
	IIa	2.87	3.08	7.41	2.22	2.03	7.07	2.85	2.90	8.07	3.04	3.46	8.35	3.25	3.09	8.15
	IIb	1.89	1.37	4.40	1.30	0.92	4.90	1.75	1.87	5.90	1.72	2.16	5.13	2.27	1.62	5.05
	III	1.13	1.23	2.90	0.90	0.85	2.70	1.15	1.13	3.20	1.20	1.42	3.32	1.42	1.49	3.58
Slope (mm/year)	I	0.02	0.02	0.04	0.03	0.03	0.08	0.02	0.02	0.05	0.02	0.03	0.06	0.02	0.02	0.05
	IIa	0.25	0.27	0.65	0.29	0.27	0.93	0.25	0.25	0.71	0.27	0.31	0.73	0.29	0.27	0.71
	IIb	0.17	0.12	0.39	0.18	0.13	0.66	0.15	0.17	0.52	0.15	0.19	0.46	0.20	0.15	0.45
	III	2.52	2.87	7.39	2.44	2.01	8.37	2.61	2.38	8.75	2.62	2.89	8.12	3.02	2.97	8.11
Position[b] (mm)	I	0.06	0.06	0.13	0.06	0.06	0.15	0.06	0.06	0.14	0.06	0.08	0.16	0.06	0.06	0.14
	IIa	2.58	2.77	6.66	2.00	1.83	6.38	2.56	2.61	7.25	2.73	3.13	7.50	2.92	2.77	7.33
	IIb	1.69	1.22	3.94	1.16	0.81	4.39	1.56	1.67	5.29	1.53	1.93	4.59	2.03	1.45	4.52
	III	12.6	14.3	37.0	7.95	6.54	27.3	13.1	11.9	43.8	13.1	14.5	40.6	15.1	14.9	40.5
Position[c] (mm)	I	0.12	0.12	0.25	0.12	0.12	0.30	0.11	0.12	0.27	0.13	0.16	0.31	0.12	0.12	0.28
	IIa	2.87	3.08	7.41	2.22	2.03	7.06	2.85	2.90	8.07	3.04	3.49	8.35	3.25	3.08	8.15
	IIb	1.88	1.37	4.40	1.29	0.91	4.89	1.75	1.87	5.90	1.72	2.16	5.13	2.27	1.62	5.05
	III	25.2	28.7	73.9	15.9	13.1	54.5	26.1	23.7	87.5	26.1	28.9	81.2	30.1	29.7	81.0

[a]I: White noise only; IIa: white noise plus flicker noise; IIb: short-memory noise plus flicker noise with proper functional model; and III: white noise plus random walk noise.
[b]Error position in the middle of time series.
[c]Error position at the end of time series.

Table 9. Minimum and Maximum Error Estimate Coefficients of Different Parameters Compared to Those of White Noise Only Model (I)

	Error Model		
Error	IIa	IIb	III
Intercept	16−29	7−21	7−12
Slope	9−16	5−10	67−185
Position[a]	31−52	13−38	109−313
Position[b]	17−30	8−22	109−324

[a]In the middle of time series.
[b]At the end of time series.

position error estimate; the error in the intercept plays the main role. This holds also for model IIb, but with an improvement by a factor of 1.6. The behavior of model III is somehow different. The results are too pessimistic and the error in the slope plays the main role.

[73] In geophysical literature, the site velocity uncertainty is usually of interest and not directly the position error. However, in geodetic applications (for example, realization of ITRF), the goal of the site velocity studies is both the intercept and the slope, from which the position and its uncertainty can be directly propagated. Figure 9 illustrates the effect of site velocity only and site velocity plus intercept on the position error for different stochastic models. The plot shows the importance of the propagation of errors correctly on topics like realization of ITRF. When interested in position error and relying only on the site velocity error, it seems that models IIa and IIb are more appropriate for long-term accuracy but give optimistic results over short periods. However, if one propagates both the slope and the intercept errors, these models will always be preferred. On the other hand, model III is likely suitable for short-term accuracy but yields pessimistic results over long periods.

7. Conclusions and Recommendations

[74] We assessed the noise characteristics in global time series of daily position estimates by LS-VCE. The method is easily understood, generally applicable, and very flexible. The LS-VCE estimators are unbiased and of minimum variance. This method provides the precision of the (co)-variance estimators. Based on the results given, the following conclusions can be drawn:

[75] The w-test statistic is a powerful tool to recognize the data noise characteristics in order to construct an appropriate stochastic model. Using the w-test, a combination of short-memory (white plus autoregressive) noise and flicker noise was in general found to best describe the noise characteristics of the position components; we hardly observed that the strict white plus random walk was the preferred noise model. These results have also been verified using correlograms of the time series, the frequencies of harmonic functions, and the signs of estimated flicker and random walk noise variance components.

[76] The least squares harmonic estimation method was used to find and consequently remove a set of harmonic functions from the data. These harmonic functions captured unmodelled effects. The results confirm the presence of annual and semiannual signals in the series. We could also observe other periodic effects; for example, a period of 13.66, 14.2, and 14.8 days. We observed also significant periodic patterns with a period of 350 days and its fractions $350/n$, $n = 2, \ldots, 8$, which are likely due to aliased multipath effects in permanent stations. When such variations have underlying physical phenomena (or modeling error), their effects can be considered as systematic periodic signals. It may not be appropriate to capture their effects by a power law noise process in the stochastic model. They may mistakenly mimic flicker or random walk noise if we neglect them in the functional model. Therefore neglecting such effects, which may be best described by a deterministic model rather than a power law noise model, can seriously affect the error estimate of the site velocity and the position.

[77] There are also some effects in the series that are not of periodic nature. They can most likely be expressed as power law noise. We however employed the harmonic estimation method to find more frequencies in the series.
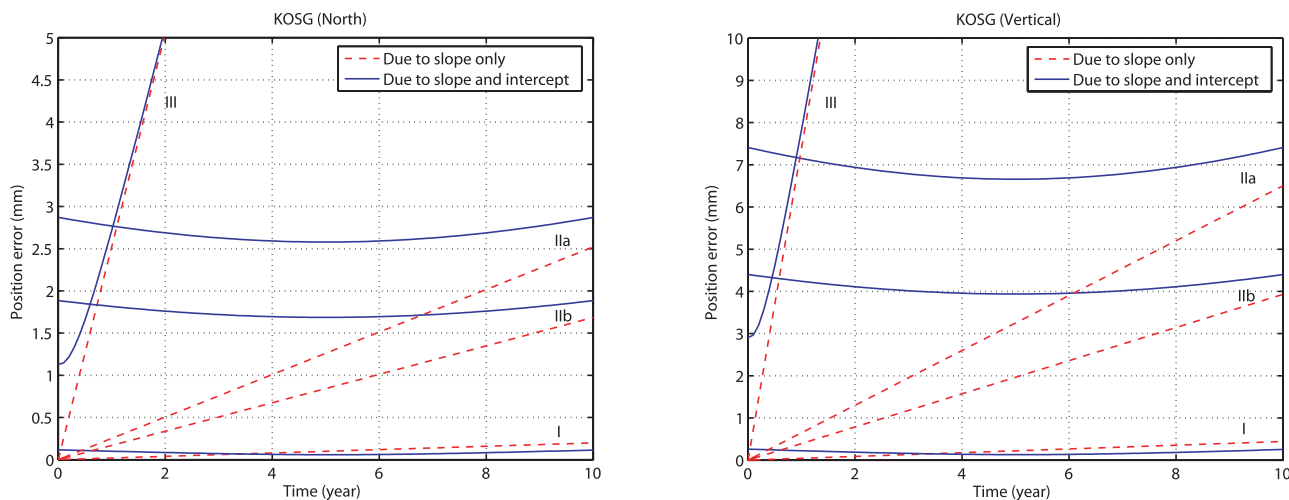

Figure 9. Position error (standard deviation) of different stochastic models as a function of time for north and vertical component of KOSG caused by rate uncertainty only (dashed lines) and full covariance matrix of intercept and site velocity (solid lines); WN only (I), WN + FN (IIa), SM (short-memory) + FN (IIb), and WN + RW (III).

A significant decrease in the *w*-test values for power law noise implies that most parts of this noise are captured by the harmonic functions. This led us to see a duality between the stochastic and the functional model. In fact, what is not captured in the functional model is captured in the stochastic model. When we include the harmonic functions, almost exclusively white noise remains in the data. Only at very high frequencies a significant time correlation appeared to be present, which can be expressed as an exponential function (for example, a first-order autoregressive noise process). This noise can be caused by common sources of errors like atmospheric effects as well as satellite orbit errors that last over only a few successive days. Based on the value of the timescale $\alpha \approx 0.25$, Williams (personal communication, 2006) suggested atmospheric loading as a candidate for this noise process.

[78] The overestimation of the power law noise was due to the presence of the autoregressive noise and also the justified hidden periodic effects in the series. This means that neither the white noise plus flicker noise model nor the white noise plus random walk noise model is the preferred model. The best model includes in addition to power law noise also other noise models like autoregressive noise or as *Langbein* [2004] used band-pass-filtered noise on EDM data. Instead of a strict white noise model, a short-memory noise process was introduced which led to the reduction of the flicker noise magnitude.

References

Abbasi, M. (1999), Comparison of Fourier, least-squares and wavelet spectral analysis methods, tested on Persian Gulf tidal data, M.Sc. thesis, Surveying Engineering Department, K.N. Toosi University of Technology, Tehran, Iran.

Agnew, D. (1992), The time domain behavior of power law noises, *Geophys. Res. Lett.*, 19, 333–336.

Agnew, D. C., and K. M. Larson (2007), Finding the repeat times of the GPS constellation, *GPS Solut.*, 11(1), 71–76.

Altamimi, Z., P. Sillard, and C. Boucher (2002), ITRF2000: A new release of the International Terrestrial Reference Frame for earth science applications, *J. Geophys. Res.*, 107(B10), 2214, doi:10.1029/2001JB000561.

Amiri-Simkooei, A. R. (2005), Separating receiver noise and multipath effects in time series of GPS baselines using harmonic functions, in *ION GNSS 18th International Technical Meeting of the Satellite Division*, pp. 1034–1045, Long Beach, Calif., 13–16 Sept. (sponsored student paper).

Amiri-Simkooei, A. R. (2007), Least-squares variance component estimation: Theory and GPS applications, Ph.D. thesis, Delft University of Technology, Publication on Geodesy, 64, Netherlands Geodetic Commission, Delft.

Amiri-Simkooei, A. R., and C. C. J. M. Tiberius (2004), Testing of high-end GNSS receivers, in *2nd ESA/Estec Workshop on Satellite Navigation User Equipment Technologies NaviTec'2004*, edited by G. S. Granados, WPP-239, Noordwijk, Netherlands, 8–10 Dec.

Amiri-Simkooei, A. R., and C. C. J. M. Tiberius (2007), Assessing receiver noise using GPS short baseline time series, *GPS Solut.*, 11(1), 21–35.

Asgari, J., and A. Harmel (2005), Least squares spectral analysis of GPS derived ionospheric data, in *The ION National Technical Meeting*, San Diego, Calif., 24–26 Jan.

Beavan, J. (2005), Noise properties of continuous GPS data from concrete pillar geodetic monuments in New Zealand and comparison with data from U.S. deep drilled braced monuments, *J. Geophys. Res.*, 110, B08410, doi:10.1029/2005JB003642.

Beutler, G. (2006), Geodetic monitoring of the Earth and the role of geodetic theory in the age of space geodesy, in *VI Hotine−Marussi Symposium of Theoretical and Computational Geodesy: Challenge and Role of Modern Geodesy*, Wuhan University, China, 29 May−2 June.

Beutler, G., M. Rothacher, S. Schaer, T. A. Springer, J. Kouba, and R. E. Neilan (1999), The International GPS Service (IGS): An interdisciplinary service in support of Earth sciences, *Adv. Space Res.*, 23(4), 631–635.

Bischoff, W., B. Heck, J. Howind, and A. Teusch (2005), A procedure for testing the assumption of homoscedasticity in least squares residuals: A case study of GPS carrier-phase observations, *J. Geod.*, 78, 397–404.

Blewitt, G., and D. Lavallée (2002), Effect of annual signals on geodetic velocity, *J. Geophys. Res.*, 107(B7), 2145, doi:10.1029/2001JB000570.

Bock, Y., R. M. Nikolaidis, P. J. de Jonge, and M. Bevis (2000), Instantaneous geodetic positioning at medium distances with the Global Positioning System, *J. Geophys. Res.*, 105(B12), 28,223–28,253.

Calais, E. (1999), Continuous GPS measurements across the Western Alps, 1996−1998, *Geophys. J. Int.*, 138, 221–230.

Craymer, M. R. (1998), The least-squares spectrum, its inverse transform and autocorrelation function: Theory and some applications in geodesy, Ph.D. thesis, Department of Civil Engineering, University of Toronto, Toronto, Canada.

Crocetto, N., M. Gatti, and P. Russo (2000), Simplified formulae for the BIQUE estimation of variance components in disjunctive observation groups, *J. Geod.*, 74, 447–457.

Davies, P., and G. Blewitt (2000), Methodology for global geodetic time series estimation: A new tool for geodynamics, *J. Geophys. Res.*, 105(B5), 11,083–11,100.

Ding, X. L., D. W. Zheng, D. N. Dong, C. Ma, Y. Q. Chen, and G. L. Wang (2005), Seasonal and secular positional variations at eight co-located GPS and VLBI stations, *J. Geod.*, 79(1−3), 71–81, doi:10.1007/s00190-005-0444-3.

Dong, D., P. Fang, Y. Bock, M. K. Cheng, and S. Miyazaki (2002), Anatomy of apparent seasonal variations from GPS-derived site position time series, *J. Geophys. Res.*, 107(B4), 2075, doi:10.1029/2001JB000573.

Grodecki, J. (1999), Generalized maximum-likelihood estimation of variance components with inverted gamma prior, *J. Geod.*, 73, 367–374.

Grodecki, J. (2001), Generalized maximum-likelihood estimation of variance−covariance components with non-informative prior, *J. Geod.*, 75, 157–163.

Johnson, H. O., and D. C. Agnew (2000), Correlated noise in geodetic time series, in *U.S. Geol. Surv. Final Tech. Rep., FTR-1434-HQ-97-GR-03155*.

Johnson, H. O., and F. K. Wyatt (1994), Geodetic network design for fault-mechanics studies, *Manuscr. Geod.*, 19, 309–323.

Kenyeres, A., and C. Bruyninx (2004), EPN coordinate time series monitoring for reference frame maintenance, *GPS Solut.*, 8, 200–209, doi:10.1007/s10291-004-0104-8.

Koch, K. R. (1986), Maximum likelihood estimate of variance components, *Bull. Geod.*, 60, 329–338. (ideas by A.J. Pope).

Kubik, K. (1970), The estimation of the weights of measured quantities within the method of least squares, *Bull. Geod.*, 95, 21–40.

Kusche, J., and J. O. Schrama (2005), Surface mass redistribution inversion from global GPS deformation and Gravity Recovery and Climate Experiment (GRACE) gravity data, *J. Geophys. Res.*, 110, B09409, doi:10.1029/2004JB003556.

Langbein, J. (2004), Noise in two-color electronic distance meter measurements revisited, *J. Geophys. Res.*, 109, B04406, doi:10.1029/2003JB002819.

Langbein, J., and Y. Bock (2004), High-rate real-time GPS network at Parkfield: Utility for detecting fault slip and seismic displacements, *Geophys. Res. Lett.*, 31, L15S20, doi:10.1029/2003GL019408.

Langbein, J., and H. Johnson (1995), Noise level of geodetic monuments, *Eos Trans. AGU*, 76(46), Fall Meet. Suppl., F142.

Langbein, J., and H. Johnson (1997), Correlated errors in geodetic time series: Implications for time-dependent deformation, *J. Geophys. Res.*, 102(B1), 591–603.

Mandelbrot, B. (1983), *The Fractal Geometry of Nature*, W. H. Freeman, New York.

Mandelbrot, B., and J. van Ness (1968), Fractional Brownian motions, fractional noises, and applications, *SIAM Rev.*, 10, 422–439.

Mao, A., C. G. A. Harrison, and T. H. Dixon (1999), Noise in GPS coordinate time series, *J. Geophys. Res.*, 104(B2), 2797–2816.

Ou, Z. (1989), Estimation of variance and covariance components, *Bull. Geod.*, 63, 139–148.

Penna, N. T., and M. P. Stewart (2003), Aliased tidal signatures in continuous GPS height time series, *Geophys. Res. Lett.*, 30(23), 2184, doi:10.1029/2003GL018828.

Penna, N. T., M. A. King, and M. P. Stewart (2007), GPS height time series: Short-period origins of spurious long-period signals, *J. Geophys. Res.*, 112, B02402, doi:10.1029/2005JB004047.

Perfetti, N. (2006), Detection of station coordinate discontinuities within the Italian GPS Fiducial Network, *J. Geod.*, 80(7), 381–396, doi:10.1007/s00190-006-0080-6.

Petrov, L., and C. Ma (2003), Study of harmonic site position variations determined by very long baseline interferometry, *J. Geophys. Res.*, *108*(B4), 2190, doi:10.1029/2002JB001801.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1992), *Numerical Recipes*, 818 pp., Cambridge Univ. Press, New York.

Priestley, M. B. (1981), Spectral Analysis and Time Series, in Probability and Mathematical Statistics, Academic Press, San Diego.

Ray, J. (2006), Systematic errors in GPS position estimates, in *IGS Workshop 2006*, ESOC Darmstadt, Germany, session 11, 8–12 May.

Sahin, M., P. A. Cross, and P. C. Sellers (1992), Variance component estimation applied to satellite laser ranging, *Bull. Geod.*, *66*, 284–295.

Stewart, M. P., N. T. Penna, and D. D. Lichti (2005), Investigating the propagation mechanism of unmodelled systematic errors on coordinate time series estimated using least squares, *J. Geod.*, *79*, 479–489, doi:10.1007/s00190-005-0478-6.

Teunissen, P. J. G. (1988), Towards a least-squares framework for adjusting and testing of both functional and stochastic model, *Internal research memo*, Geodetic Computing Centre, Delft. (A reprint of original 1988 report is also available in 2004, No. 26, at http://www.lr.tudelft.nl/mgp)

Teunissen, P. J. G. (2000), Adjustment theory: An introduction, in *Series on Mathematical Geodesy and Positioning*, Delft University Press. (Available at http://www.vssd.nl)

Teunissen, P. J. G., and A. R. Amiri-Simkooei (2006), Variance component estimation by the method of least-squares, in *VI Hotine–Marussi Symposium of Theoretical and Computational Geodesy: Challenge and Role of Modern Geodesy*, 29 May–2 Jun, Wuhan University, China.

Teunissen, P. J. G., and A. R. Amiri-Simkooei (2007), Least-squares variance component estimation, *J. Geod.*, doi:10.1007/s00190-007-0157-x, in press.

Teunissen, P. J. G., D. G. Simons, and C. C. J. M. Tiberius (2005), *Probability and observation theory*, Faculty of Aerospace Engineering, Delft University, Delft University of Technology. (lecture notes AE2-E01).

Tiberius, C. C. J. M., and F. Kenselaar (2003), Variance component estimation and precise GPS positioning: Case study, *J. Surv. Eng.*, *129*(1), 11–18.

Urschl, C., W. Gurtner, U. Hugentobler, S. Schaer, and G. Beutler (2005), Validation of GNSS orbits using SLR observations, *Adv. Space Res.*, *36*(3), 412–417.

Urschl, C., G. B. W. Gurtner, U. Hugentobler, and S. Schaer (2006), Validation of GNSS orbits using SLR observations, in *IGS Workshop 2006*, ESOC Darmstadt, Germany, session 11, 8–12 May.

van Dam, T., J. Wahr, P. C. D. Milly, A. B. Shmakin, G. Blewitt, D. Lavallée, and K. M. Larson (2001), Crustal displacements due to continental water loading, *Geophys. Res. Lett.*, *28*(4), 651–654.

Vaníček, P. (1969), Approximate spectral analysis by least-squares fit, *Astrophys. Space Sci.*, *4*, 387–391.

Wang, J., C. Satirapod, and C. Rizos (2002), Stochastic assessment of GPS carrier phase measurements for precise static relative positioning, *J. Geod.*, *76*, 95–104, doi:10.1007/s00190-001-0225-6.

Williams, S. D. P. (2003a), The effect of coloured noise on the uncertainties of rates estimated from geodetic time series, *J. Geod.*, *76*, 483–494.

Williams, S. D. P. (2003b), Offsets in Global Positioning System time series, *J. Geophys. Res.*, *108*(B6), 2310, doi:10.1029/2002JB002156.

Williams, S. D. P., Y. Bock, P. Fang, P. Jamason, R. M. Nikolaidis, L. Prawirodirdjo, M. Miller, and D. J. Johnson (2004), Error analysis of continuous GPS position time series, *J. Geophys. Res.*, *109*, B03412, doi:10.1029/2003JB002741.

Wyatt, F. K. (1982), Displacement of surface monuments: Horizontal motion, *J. Geophys. Res.*, *87*, 979–989.

Wyatt, F. K. (1989), Displacement of surface monuments: Vertical motion, *J. Geophys. Res.*, *94*, 1655–1664.

Wyatt, F. K., H. Bolton, S. Bralla, and D. C. Agnew (1989), New designs of geodetic monuments for use with GPS, *Eos Trans. AGU*, *70*, 1054–1055.

Yu, Z. (1996), A universal formula of maximum likelihood estimation of variance–covariance components, *J. Geod.*, *70*, 233–240.

Zhang, J., Y. Bock, H. Johnson, P. Fang, S. Williams, J. Genrich, S. Wdowinski, and J. Behr (1997), Southern California Permanent GPS Geodetic Array: Error analysis of daily position estimates and site velocities, *J. Geophys. Res.*, *102*, 18,035–18,055.

Zhang, C. D., X. P. Wu, J. M. Hao, H. B. He, and D. M. Zhao (2004), ISO2002: An analytical stochastic model of multi-difference GPS carrier-phase data, *J. Geod.*, *78*(3), 192–200.

Zumberge, J. F., M. B. Heflin, D. C. Jefferson, M. M. Watkins, and F. H. Webb (1997), Precise point positioning for the efficient and robust analysis of GPS data from large networks, *J. Geophys. Res.*, *102*, 5005–5017.

———————————

A. R. Amiri-Simkooei, P. J. G. Teunissen, and C. C. J. M. Tiberius, Delft Institute of Earth Observation and Space systems (DEOS), Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, 2629 HS, Delft, Netherlands. (a.amirisimkooei@tudelft.nl)