The effect of correlation on uncertainty estimates – with GPS examples

Research Article

P. Jansson[1]*, C.-G. Persson[1, 2]

1   Division of Geodesy & Geoinformatics, Royal Institute of Technology, SE-100 44 Stockholm, Sweden
2   Lantmäteriet (Swedish Mapping, Cadastral and Land Registration Authority), SE-80182 Gävle, Sweden

Abstract:
This article deals with the effect of correlation on the estimates of measurement uncertainty, with some focus on Global Positioning Satellite (GPS) time series analysis. Analytical derivations and Monte Carlo simulations form the theoretical basis, which shows that uncompensated correlation produces unrealistic uncertainty estimates. Tools for handling correlation in connection with estimation of uncertainty, construction of confidence intervals, hypothesis testing, design of measurement strategies, and development of tolerances are outlined and demonstrated. The GPS observation time series used in the article has a short to medium range correlation, and can therefore be handled with the presented tools – based on a simple Location-model and stationary stochastic processes.

1.   Introduction

In this paper, we will give a general insight into the effect of correlation in the study of measurement uncertainty. We will analyse the consequences of neglecting correlation when estimating measurement uncertainty and we will briefly outline how correlation might be compensated for. Our goal is not to provide perfect estimates, but to avoid uncertainty measures that are 100% wrong. Using Global Positioning System (GPS) is only one of many possible applications where this is valid, and is used as illustration in this paper.

Multiple studies have shown GPS errors to be time-correlated (El-Rabbany, 1994; Howind et al., 1999; Bona, 2000). Time-correlated errors appear due to multipath, antenna effects, atmospheric errors including ionospheric and tropospheric, and slowly changing satellite constellations (Olynik, 2002).

*E-mail: Patric.jansson@abe.kth.se;
Tel.: +46–8–790–7376; Fax: +46–8–790–8580

A basic assumption in both least squares and Kalman filtering, used in the estimation of position, is that each new observation is completely independent from the others. If they are not, correlation exists and this must be accounted for. This is a problem because temporal correlation is generally ignored by modern RTK manufacturers. The correlation is less important for least squares parameter estimation than for their quality/uncertainty estimates. Treating time-correlated observations as independent can result in an over-optimistically estimated positional accuracy, in particular for a short time of observation, and too much confidence being placed in the results by the surveyors.

The importance of handling correlation between geodetic observations was originally pointed out by Tienstra (1947). El-Rabbany (1994) presented an algorithm for accounting for temporal correlation using a modified least squares algorithm. The method incorporates a fully populated enlarged variance- covariance matrix where each new measurement results in an additional row and column being added to the matrix. However, the matrix would grow to immense proportions especially in RTK applications where a 1 Hz data rate is typically used.

VERSITA

A different approach is accounting for temporal correlation by time-differencing the observations in a Kalman filter. This method differences observations across adjacent epochs in order to cancel the time-correlated errors. The method was originally presented by Bryson and Henrikson (1968), but has been modified by Petovello et al. (2009) in order to eliminate the issue of time-latency. Miller et al. (2010) implemented this alternative approach, as presented by Petovello et al. (2009), and evaluated its performance using real GPS data. The author concludes that for surveying applications, this modified Kalman filter is most beneficial during static data collection periods and not when the receiver is moving.

Bona (2000) studied precision, cross- and temporal-correlation of the GPS code, and phase observables. He concluded that noise is seldom white; in other words, time correlation is present. Further, the stochastic aspects of the observations are receiver specific, so for proper modelling, the characteristics have to be determined for all receivers involved in a survey. Along the same line of thought, Amiri-Simkooei and Tiberius (2007) evaluated the GPS receiver-induced measurement noise and multipath effects using harmonic functions via the estimated baseline time series. Borre and Tiberius (2000) studied the time behaviour of the GPS observations' noise, using standard techniques from modern time series analysis. An attempt was made to capture the time correlation by a first-order auto-regressive noise process. The agreement between the data and this model was found to be poor for one receiver make, but good for another.

In the context of the possible adverse effects of ignoring correlation between the GPS-derived position time series, Williams (2003) studied the rate uncertainty using different stochastic models, in other words, the white-noise and coloured-noise process. Khodabandeh et al. (2012) showed the procedure of periodic effects' identification can completely go wrong, should one choose the white-noise structure for a coloured-noise process.

Odolinski (2011) studied two different methods for estimating correlation lengths in data sets from network RTK positioning: the composite first-order Gauss-Markov model, and the one-component first-order Gauss-Markov model. He concluded that the composite model showed a significantly better least-squared fit to data compared to the one-component model. Further, it also showed good capability to model the autocovariance function for network RTK positioning.

Vollath et al. (2002) evaluated time-correlated errors in network RTK in perspective of quantifying the improvements of the measurement errors by distinguishing the errors based on their statistical properties: time-correlated errors, uncorrelated errors, and biases. The major finding was the presence of very high correlation times in both the ionospheric and the geometric errors in the range of one hour. The consequence is that simple averaging techniques, taken over a time span of one hour, for instance, are quite insufficient for productive RTK positioning because the user expects a time-to-fix in the tens of seconds, to a few minutes.

The main purpose of this study is to analyse the effect of correlation in the study of measurement uncertainty. An important difference with respect to previous studies is the emphasis on estimating correlation lengths and outlining how correlation might be compensated for. Many of the studies mentioned above did not directly intend to quantify and estimate these kinds of correlation lengths. The structure of this paper is organized as follows. The functional and stochastic models are defined, and the basic formulas are derived in Section 2. These establish the theoretical platform for the derivations and analyses in Section 3, which are applied to GPS time series analysis in Section 4. Finally, the results and findings are summarized in Section 5.

In the present paper, positive correlation functions have been used throughout. They are intuitive, they might be given a geometrical interpretation, and they are commonly used. Note also that estimated correlation functions might be negative and/or oscillating even if the "true" functions are strictly positive. Further, it is not uncommon that the oscillations are non-significant in a statistical sense.

2. Basic formulas

In Section 2.1, the functional and the stochastic models are introduced respectively. In Section 2.2, we introduce the location-model, and in Section 2.3 we discuss stochastic processes, covariance functions, and correlation functions.

2.1. Functional/stochastic model and least squares estimates

Our *functional model* is the following (Bjerhammar, 1973, p. 124):

$$\mathbf{Ax} = \mathbf{l} - \varepsilon \qquad (1)$$

where $\mathbf{A}$ is the *design matrix*, $\mathbf{x}$ is the vector of *unknowns,* and $\mathbf{l}$ is the vector of *observations*, with *errors* ε. The *stochastic model* is

$$V\{\mathbf{l}\} = E\{\varepsilon\varepsilon^{\mathsf{T}}\} = \sigma^2\mathbf{Q} = \mathbf{C}, \qquad (2)$$

where $V\{\ \}$ denotes *variance* and $E\{\ \}$ denotes *expectation*. σ^2 is the *variance factor*, $\mathbf{Q}$ the *correlation matrix,* and $\mathbf{C}$ the *variance–covariance matrix* of the observations $\mathbf{l}$.

As limited knowledge of correlation is part of our problem, we use "un-weighted" *least squares estimation* in our analyses. Note, however, that this is not a new stochastic model; it is only a way to get information about variance-covariance relations. The formulas for this estimation are:

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathsf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{T}}\mathbf{l} = \mathbf{A}^{+}\mathbf{l}, \qquad (3)$$

and for the errors:

$$\hat{\varepsilon} = \mathbf{l} - \mathbf{A}\hat{\mathbf{x}} = (\mathbf{I} - \mathbf{A}(\mathbf{A}^{\mathsf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{T}})\mathbf{l} = (\mathbf{I} - \mathbf{A}^{\mathsf{o}})\mathbf{l} = (\mathbf{I} - \mathbf{A}^{\mathsf{o}})\varepsilon, \qquad (4)$$

where $(A^TA)^{-1}A^T = A^+$, and $A(A^TA)^{-1}A^T = AA^+ = A^o$. The expression $V\{FY\} = F \cdot V\{Y\} \cdot F^T$ yields:

$$V\{\hat{x}\} = C_{\hat{x}\hat{x}} = (A^TA)^{-1}A^TCA(A^TA)^{-1}$$
$$= A^+C(A^+)^T = \sigma^2 A^+Q(A^+)^T, \qquad (5)$$

and further:

$$V\{\hat{\varepsilon}\} = C_{\hat{\varepsilon}\hat{\varepsilon}} = E\{\hat{\varepsilon}\hat{\varepsilon}^T\}$$
$$= (I - A^o)C(I - A^o) = \sigma^2(I - A^o)Q(I - A^o) \qquad (6)$$

because the matrix $I - A^o$ is symmetric. This means that if we can estimate C and/or Q , we can use these estimates for uncertainty analyses afterwards.

2.2. Specific formulas for the actual application

In our study, the *mean μ* is the only unknown parameter, which yields the *location–model*:

$$A\mu = I - \varepsilon$$
$$A\hat{\mu} = A\bar{I} = I - \hat{\varepsilon} \qquad (7)$$

This model is simple and easy to handle, but of course it does not cover all geodetic applications. Our results are, therefore, not generally applicable.
The expression of the location-model could be simplified with the use of the auxiliary matrices:

$$e^T_n = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}; \quad (1, n), \qquad (8)$$

and

$$E_n = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}; \quad (n, n), \qquad (9)$$

with the properties:

$$e_n e_n^T = E_n, \qquad (10)$$

and

$$e_n^T e_n = n, \qquad (11)$$

thus:

$$e_n^T E_n = n e_n^T \qquad (12)$$

and

$$E_n E_n = n E_n, \qquad (13)$$

which yields:

$$e_n^T C e_n = tr\{E_n C\} = \sum_i^n \sum_j^n C_{ij}, \qquad (14)$$

where $tr\{\}$ denotes the *trace* of a (square) matrix – that is, the sum of its diagonal elements. In our application, we note that:

$$A = e_n \qquad (15)$$

which yields:

$$A^+ = (e_n^T e_n)^{-1} e_n^T = \frac{1}{n} e_n^T \qquad (16)$$

and

$$A^o = e_n(e_n^T e_n)^{-1} e_n^T = \frac{1}{n} e_n e_n^T = \frac{1}{n} E_n. \qquad (17)$$

2.3. Stochastic processes, covariance functions and correlation functions

Our stochastic model might also be applied to *stochastic processes*, which usually are assumed to be *stationary*. Then the variance-covariance matrix C is computed with the use of the (auto) *covariance function $r(\tau)$* in the following way:

$$C_{ij} = r(\tau_{ij}), \qquad (18)$$

where τ_{ij} denotes the distance, in time or space, between observation i and j; and $r(0)$ corresponds to the variance factor σ^2 in Eq. (2). The correlation matrix Q is computed from the (auto) *correlation function $\rho(\tau)$* in a corresponding way:

$$Q_{ij} = \rho(\tau_{ij}) = r(\tau_{ij})/r(0) \Leftrightarrow C_{ij} = \sigma^2 Q_{ij}. \qquad (19)$$

Two examples of correlation functions are shown in Figure 1.

3. Uncertainty estimates

3.1. Variance factor and variance-covariance matrix known

Derivation and use of Ω

If the observations are correlated, the general formula for the variance of the mean $\hat{\mu} = \bar{l}$ is:

$$\sigma^2_{\bar{l}} = V\{\bar{l}\} = \sigma^2 A^+Q(A^+)^T = \frac{\sigma^2}{n^2} e_n^T Q e_n = \frac{\sigma^2}{n^2} \sum_{i=1}^n \sum_{j=1}^n Q_{ij} \qquad (20)$$
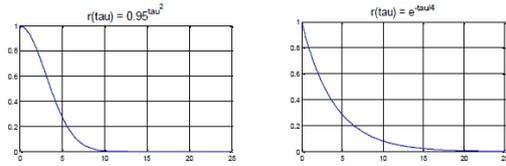
Figure 1. "Gaussian correlation function" (left) and "First-order Gauss-Markov correlation function"(right).

If the observations are uncorrelated, we get the well-known, special case of:

$$\sigma^2{}_{\bar{l}} = \frac{\sigma^2}{n^2} n = \frac{\sigma^2}{n} \Leftrightarrow \sigma_{\bar{l}} = \frac{\sigma}{\sqrt{n}}. \qquad (21)$$

One way to estimate the consequences of correlation is to calculate the number of uncorrelated observations corresponding to Eq. (20). This *effective number of observations* n_o is given by:

$$n_0 = n^2 / \left(\sum_i \sum_j Q_{ij} \right), \qquad (22)$$

because:

$$\frac{\sigma^2}{n_0} = \frac{\sigma^2}{n^2} \sum_i \sum_j Q_{ij} = \sigma^2{}_{\bar{l}}. \qquad (23)$$

The ratio Ω (omega) between n and n_o:

$$\Omega = n/n_0 = \frac{1}{n} \sum_i \sum_j Q_{ij}, \qquad (24)$$

is a sort of *correlation length* – that is, the (time) distance required in order to eliminate the effect of correlation.
Next,

$$\frac{1}{\Omega} = \frac{n_0}{n} \qquad (25)$$

is the portion of the original observations that give "full effect" when estimating the uncertainty of the mean. Note that $\Omega = 1$ means uncorrelated observations, where $n_o = n$.
Furthermore,

$$\sigma^2{}_{\bar{l}} = \frac{\sigma^2}{n_0} = \frac{\sigma^2}{n} \cdot \frac{n}{n_0} = \frac{\sigma^2}{n} \cdot \Omega \Leftrightarrow \sigma_{\bar{l}} = \frac{\sigma}{\sqrt{n}} \sqrt{\Omega}. \qquad (26)$$

The Gaussian type of correlation function from Figure 1:

$$\rho(\tau) = u^{\tau^2} = e^{-\tau^2 |\ln u|}; \quad 0 < u < 1, \qquad (27)$$

Table 1. The effect of correlation: confidence intervals are enlarged, and the number of effective observations is reduced. Stated u-values refer to the Gaussian-type correlation function in Eq. (27).

u–value	Ω_{appr}	$\sqrt{\Omega_{appr}} \cdot k$	$n_o \approx (1/\Omega_{appr}) \cdot n$
0.99	17.7	$4{,}2 \cdot k$	$0.057 \cdot n$
0.90	5.5	$2{,}3 \cdot k$	$0.183 \cdot n$
0.70	3.0	$1{,}7 \cdot k$	$0.337 \cdot n$
0.40	1.9	$1{,}4 \cdot k$	$0.540 \cdot n$

has similarities with the frequency function of a normally distributed stochastic variable. This leads to the approximation:

$$\Omega \approx \Omega_{appr} = \sqrt{\pi / \ln |u|} \qquad (28)$$

from which an estimate of $n_o \approx (1/\Omega_{appr}) \cdot n$ can be derived. One consequence of Eq. (26) is a "magnification" of the *confidence intervals*

$$\text{from}: \ |\mu| \leq k\sigma/\sqrt{n} \ \text{to} \ |\mu| \leq \sqrt{\Omega} \cdot k\sigma/\sqrt{n}. \qquad (29)$$

In Table 1, these correlation effects are shown for different u-values.

Geometric interpretation of Ω

The previous definition of Ω is *analytical/quantitative*. It is also possible to interpret Ω in a more *geometric* way.
For large matrices and a limited correlation length, the following approximation is valid:

$$\Omega = \frac{1}{n} \sum_i \sum_j Q_{ij} \approx \frac{1}{n} n \sum_j Q_j = \sum_j Q_j \qquad (30)$$

that is, $\Omega \approx$ the sum of a row in the correlation matrix.
The following schematic correlation function, see Figure 2, is illustrative:

$$\rho(i) = 0, \ |i| \geqslant n_{corr}$$
$$\rho(i) = \frac{n_{corr} - |i|}{n_{corr}}, \ 0 < |i| < n_{corr}$$
$$\rho(0) = 1 \qquad (31)$$

This function generates the following sum of the elements in an infinitely long row in the variance- covariance matrix:

$$\sum_j Q_j \approx \sum_{-\infty}^{\infty} \rho(j) = \int_{-\infty}^{\infty} \rho(\tau) \, d\tau = n_{corr}, \qquad (32)$$

because the integral is equal to the area of the triangle in Figure 2. In addition, therefore, the geometric interpretation is that $\Omega \approx$ the distance n_{corr} where correlation ceases, that is, the correlation length.
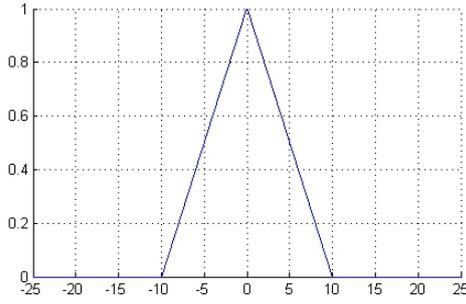
Figure 2. Schematic, triangular correlation function ($n_{corr} = 10$).

3.2. Variance factor unknown but correlation matrix known

An unbiased estimator of the variance factor and its uncertainty

For uncorrelated observations we have:

$$E\{\hat{\varepsilon}^{\mathsf{T}}\hat{\varepsilon}\} = E\{\varepsilon^{\mathsf{T}}(I - A^o)^{\mathsf{T}}(I - A^o)\varepsilon\}$$
$$= tr\{E\{\varepsilon^{\mathsf{T}}(I - A^o)\varepsilon\}\} = tr\{(I - A^o)E\{\varepsilon\varepsilon^{\mathsf{T}}\}\}$$
$$= tr\{(I - A^o)C\} = tr\{C\} - tr\{A^oC\}$$
$$= n\sigma^2 - \frac{1}{n}tr\{E_nC\} = n\sigma^2 - \frac{1}{n}\sum_i\sum_j C_{ij}$$
$$= n\sigma^2 - \sigma^2\frac{1}{n}\sum_i\sum_j Q_{ij} = \left\{\begin{array}{c}uncorr.\\obs.\end{array}\right\} = n\sigma^2 - \sigma^2$$
$$= (n-1)\sigma^2, \tag{33}$$

which means that the usual estimator $\hat{\sigma}^2 = \hat{\varepsilon}^{\mathsf{T}}\hat{\varepsilon}/(n-1)$ of σ^2 is unbiased. For correlated observations we can compute a similar estimator. Eq. (33) yields:

$$E\{\hat{\varepsilon}^{\mathsf{T}}\hat{\varepsilon}\} = n\sigma^2 - \sigma^2\frac{1}{n}\sum_i\sum_j Q_{ij} = \left\{\begin{array}{c}corr.\\obs.\end{array}\right\} = n\sigma^2 - \Omega\sigma^2$$
$$= (n - \Omega)\sigma^2, \tag{34}$$

which means, in the correlated case:

$$\hat{\sigma_0}^2 = \hat{\varepsilon}^{\mathsf{T}}\hat{\varepsilon}/(n - \Omega) \tag{35}$$

is the unbiased estimator of σ^2. For an arbitrary symmetric matrix we have (cf. Persson, 1981):

$$V\{\varepsilon^{\mathsf{T}}M\varepsilon\} = 2tr\{MCMC\}, \tag{36}$$

if ε is normally distributed with mean 0 and variance-covariance matrix C , $N(0, C)$. This yields:

$$V\{\hat{\varepsilon}^{\mathsf{T}}\hat{\varepsilon}\} = V\{\varepsilon^{\mathsf{T}}(I - A^o)\varepsilon\} = 2tr\{(I - A^o)C(I - A^o)C\}$$
$$= 2tr\{(C - A^oC)(C - A^oC)\}$$
$$= 2tr\{C^2 - A^oC^2 - CA^oC + A^oCA^oC\}$$
$$= 2tr\{C^2\} - 4tr\{A^oC^2\} + 2tr\{A^oCA^oC\}$$
$$= 2\sum_i\sum_j C_{ij}^2 - \frac{4}{n}\sum_i\sum_j C_{ij}^2 + \frac{2}{n^2}\left(\sum_i\sum_j C_{ij}\right)^2$$
$$= \frac{2(n-2)\sigma^4}{n}\sum_i\sum_j Q_{ij}^2 + \frac{2\sigma^4}{n^2}\left(\sum_i\sum_j Q_{ij}\right)^2$$
$$= 2\sigma^4\left(\frac{(n-2)}{n}\sum_i\sum_j Q_{ij}^2 + \Omega^2\right), \tag{37}$$

that is:

$$V\left\{\frac{\hat{\varepsilon}^{\mathsf{T}}\hat{\varepsilon}}{n - \Omega}\right\} = \frac{V\{\hat{\varepsilon}^{\mathsf{T}}\hat{\varepsilon}\}}{(n - \Omega)^2}$$
$$= \frac{2\sigma^4}{n - \Omega}\left[\frac{\Omega^2}{n - \Omega} + \frac{(n-2)}{n(n - \Omega)}\sum_i\sum_j Q_{ij}^2\right]. \tag{38}$$

For uncorrelated observations, $\Omega = 1$ and $\sum_i\sum_j Q_{ij}^2 = n$, we have:

$$V\left\{\frac{\hat{\varepsilon}^{\mathsf{T}}\hat{\varepsilon}}{n - 1}\right\} = \frac{2\sigma^4}{n - 1}\left[\frac{1}{n - 1} + \frac{n - 2}{n - 1}\right] = \frac{2\sigma^4}{n - 1}, \tag{39}$$

which explains the way the last membra of Eq. (38) is expressed. The expression in brackets is a further effect of correlation – except that $n - 1$ is replaced by $n - \Omega$. Its value is $\geqslant 1$, and the variance – the uncertainty of the estimator – increases with increasing correlation (higher Ω -values).

Approximate confidence intervals

It is tempting to try to treat $n - \Omega$ as a sort of *degrees of freedom* and use it to construct confidence intervals – despite the fact that this quantity usually has a non-integer value. Then, the traditional approach would yield the following (approximate) *confidence in–terval for the mean* (cf. Eq. (29):

$$P\left\{\mu \in \bar{l} \pm t_{\alpha/2}(n - \Omega) \cdot \frac{\hat{\sigma}_0}{\sqrt{n}}\sqrt{\Omega}\right\} \approx 1 - \alpha. \tag{40}$$

$P\{\}$ denotes probability, α is the *risk level,* and $t_{\alpha/2}(n - \Omega)$ is the critical value from a Student's t-distribution, corresponding to that risk level and $n - \Omega$ degrees of freedom.

Simulation studies – applied to the three different correlation functions in Figure 1-2 – show that this approximation is quite good. However, approximate confidence intervals for the estimated variance factor are more cumbersome to derive. According to standard procedures, one would like to try:

$$P\left\{ \frac{(n-\Omega)\,\hat{\sigma_0}^2}{\chi^2_{\alpha/2}\,(n-\Omega)} \leqslant \sigma^2 \leqslant \frac{(n-\Omega)\,\hat{\sigma_0}^2}{\chi^2_{1-\alpha/2}\,(n-\Omega)} \right\} \approx 1-\alpha, \tag{41}$$

where $\chi^2_{\alpha/2}\,(n-\Omega)$ and $\chi^2_{1-\alpha/2}\,(n-\Omega)$ are critical values from a χ^2-distribution.

However, this does not work equally smoothly! Equation (38) shows an obvious deviation from the "χ^2-behaviour" that would make Eq. (41) useful. Instead, we propose a Monte Carlo simulation.

Monte Carlo simulated confidence intervals

A time series ε_Q of n observations with correlation matrix Q is produced in the following way:

– Cholesky-factorize the correlation matrix. This yields a (lower) triangular matrix T with the property $TT^T = Q$.

– Generate a vector ε with elements $\varepsilon_i \in N(0,1)$, $i = 1, \cdots, n; E\{\varepsilon \cdot \varepsilon^T\} = I$

– Generate ε_Q from $\varepsilon_Q = T \cdot \varepsilon$; $E\{\varepsilon_Q \varepsilon_Q^T\} = E\{(T \cdot \varepsilon)(T \cdot \varepsilon)^T\} = T \cdot E\{\varepsilon \varepsilon^T\} \cdot T^T = TT^T = Q$.

This vector could be further transformed to $\varepsilon_Q{}^* = \mu \cdot \mathbf{e_n} + \sigma \cdot \varepsilon_Q$ if $\mu \neq 0$ and/or $\sigma \neq 1$. For instance, the critical values of the confidence intervals could be determined in a trial-and-error process where the values that give the requested coverage probability are chosen.

In a Monte Carlo simulation, this procedure is repeated several thousands of times and different characteristics – as mean values or variance-covariance estimates, for example – are calculated for each sample run. The distribution of these could then be studied in detail at the end of the process.

Comparison of estimation strategies

In Table 2, three different strategies for constructing confidence intervals are compared:

I. The traditional approach with the number of degrees of freedom equal to $n-1$.

II. An approximate approach where this number is set to $n-\Omega$ in order to compensate for correlation.

III. An approach based on trial-and-error Monte Carlo simulations.

The correlation function is of Gauss Markov type: e^{-tau/τ_0} (theoretical variance factor $= 1$)[1].

Except for some odd behaviour for small n-values, the following conclusions could be drawn from that table:

- The traditional approach (I) gives unrealistic uncertainty estimates for correlated observations.

- In that case, method II works quite well for estimating the uncertainty of the mean, but not for the standard deviation.

- Method III is the only method of the three that handles correlation correctly.

For uncorrelated observations (small τ_o-values), all three methods give equal and correct results.

3.3. Variance factor and correlation matrix unknown

Traditional covariance function estimation

If the correlation matrix is also unknown, a natural approach would be to start with an estimation of the covariance function using the traditional formula (consult any text-book on time series or correlation analysis):

$$\hat{r}(k) = \frac{1}{n-k} \sum_{i=1}^{n-k} \left(l_i - \bar{l}\right)\left(l_{i+k} - \bar{l}\right)$$
$$= \frac{1}{n-k} \sum_{i=1}^{n-k} \hat{\varepsilon}_i \hat{\varepsilon}_{i+k}; \quad k = 0, 1, \cdots, m, \tag{42}$$

and try to compute the correlation matrix and Ω from that estimate.

However, the use of $\hat{\varepsilon}$ instead of the true ε in the estimation makes derivation of Ω difficult, or even impossible. Equation (42) gives biased estimates if the condition $m \ll n$ is not fulfilled (see Figure 3). Furthermore, without extra-ordinary actions, it produces a singular correlation matrix and a zero- estimate of Ω!

Such an extra-ordinary action is to fit a suitable function to the "raw" $\hat{r}(k)$ values. This function should be *positive definite*, to ensure that it produces a positive definite correlation matrix. The three types of correlation functions in Figure 1-2 have this property. Another reason for this fitting is that the $\hat{r}(k)$ values in the "tail" oscillate heavily, and are usually non-significant (see Figure 6).

---

[1] *For us, τ_0 is just a parameter in the correlation function $f(\tau) = e^{-\tau/\tau_0}$. It defines the point where the correlation reaches 1/e, and is sometimes denoted "correlation length". However, we use Ω as the definition of correlation length. Thus, $\tau_o = 0$ corresponds to $\Omega = 1$, otherwise $\Omega \approx \tau_o/2$, because of the different definitions of correlation length.*

Table 2. Simulation study of different methods for the construction of confidence intervals for the mean and standard deviation, expressed as multiples (k) of the standard deviation (s). Nominal coverage probability: 95%; established probability in parenthesis. 100,000 sample runs in simulations.

Simulation parameters			I Traditional ($n-1$) (established coverage probability)		II Approximate ($n - \Omega$) (established coverage probability)		III Simulated critical values (established coverage probability)	
n	τ_o	Ω	Mean $\bar{l} \pm k \cdot s$	Std. dev. $[k_{low}, k_{high}] \cdot s$	Mean $\bar{l} \pm k \cdot s$	Std. dev. $[k_{low}, k_{high}] \cdot s$	Mean $\bar{l} \pm k \cdot s$	Std. dev. $[k_{low}, k_{high}] \cdot s$
10	3.8	5	0.72 ·s ($\approx$ 51 %)	[0.69,1.83] ·s ($\approx$72 %)	2.45 ·s ($\approx$ 95 %)	[0.84,3.30] ·s ($\approx$ 95 %)	2.39 ·s ($\approx$ 95 %)	[0.78,2.97] ·s ($\approx$ 95 %)
10	1.8	3	0.72 ·s ($\approx$ 71 %)	[0.69,1.83] ·s ($\approx$ 89 %)	1.47 ·s ($\approx$ 95 %)	[0.75,2.31] ·s ($\approx$ 94 %)	1.49 ·s ($\approx$ 95 %)	[0.71,2.25] ·s ($\approx$ 95 %)
10	10^{-5}	1	0.72 ·s ($\approx$ 95 %)	[0.69,1.83] ·s ($\approx$ 95 %)	0.72 ·s ($\approx$ 95 %)	[0.69,1.83] ·s ($\approx$ 95 %)	0.72 ·s ($\approx$ 95 %)	[0.69,1.83] ·s ($\approx$ 95 %)
50	5.6	10	0.28 ·s ($\approx$ 42 %)	[0.84,1.25] ·s ($\approx$ 56 %)	1.00 ·s ($\approx$ 93 %)	[0.91,1.42] ·s ($\approx$ 69 %)	1.08 ·s ($\approx$ 95 %)	[0.76,1.75] ·s ($\approx$ 95 %)
50	2.6	5	0.28 ·s ($\approx$ 59 %)	[0.84,1.25] ·s ($\approx$ 77 %)	0.66 ·s ($\approx$ 94 %)	[0.87,1.31] ·s ($\approx$ 80 %)	0.69 ·s ($\approx$ 95 %)	[0.78,1.46] ·s ($\approx$ 95 %)
50	10^{-5}	1	0.28 ·s ($\approx$ 95 %)	[0.84,1.25] ·s ($\approx$ 95 %)	0.28 ·s ($\approx$ 95 %)	[0.84,1.25] ·s ($\approx$ 95 %)	0.28 ·s ($\approx$ 95 %)	[0.84,1.25] ·s ($\approx$ 95 %)
200	5.1	10	0.14 ·s ($\approx$ 46 %)	[0.91,1.11] ·s ($\approx$ 61 %)	0.45 ·s ($\approx$ 94 %)	[0.93,1.14] ·s ($\approx$ 64 %)	0.46 ·s ($\approx$ 95 %)	[0.83,1.29] ·s ($\approx$ 95 %)
200	2.5	5	0.14 ·s ($\approx$ 62 %)	[0.91,1.11] ·s ($\approx$ 78 %)	0.32 ·s ($\approx$ 95 %)	[0.92,1.12] ·s ($\approx$ 78 %)	0.32 ·s ($\approx$ 95 %)	[0.87,1.19] ·s ($\approx$ 95 %)
200	10^{-5}	1	0.14 ·s ($\approx$ 95 %)	[0.91,1.11] ·s ($\approx$ 95 %)	0.14 ·s ($\approx$ 95 %)	[0.91,1.11] ·s ($\approx$ 95 %)	0.14 ·s ($\approx$ 95 %)	[0.91,1.11] ·s ($\approx$ 95 %)

The semivariogram

An alternative – and in the authors' opinion better – method is to estimate the *semivariogram*:

$$\hat{\gamma}(k) = \frac{1}{2(n-k)} \sum_{i=1}^{n-k} (l_i - l_{i+k})^2$$

$$= \frac{1}{2(n-k)} \sum_{i=1}^{n-k} ((\mu + \varepsilon_i) - (\mu + \varepsilon_{i+k}))^2$$

$$= \frac{1}{2(n-k)} \sum_{i=1}^{n-k} (\varepsilon_i - \varepsilon_{i+k})^2 \qquad (43)$$

used in connection with the interpolation method $Kriging^2$.

The following expressions define the relation between the **true** semivariogram and the **true** covariance or correlation functions:

$$r(k) = r(0) - \gamma(k) \; ; \; \rho(k) = \frac{r(k)}{r(0)} = 1 - \frac{\gamma(k)}{r(0)}. \qquad (44)$$

(However, usually, these expressions are not valid for the corresponding **estimates**, Eq. (42) and (43) respectively.)

This means that the semivariogram is "reversed" compared to the covariance or correlation functions – although with the same (geometric) correlation length. A comparison shows that Eq. (43) uses the real ε , not the estimates $\hat{\varepsilon}$ as in Eq. (42). Therefore, the semivariogram produces covariance/correlation function estimates that are more close to the theoretical ones. However, the problem with an oscillating tail remains. Therefore, there is still a need to cut the tail and fit a suitable function to the raw $\hat{\gamma}(k)$ estimates – if the complete correlation function/matrix is of interest or if an **analytical** derivation of Ω (according to Eq. (24)) is of interest. But often a good enough, **geometric** estimate of Ω is sufficient.

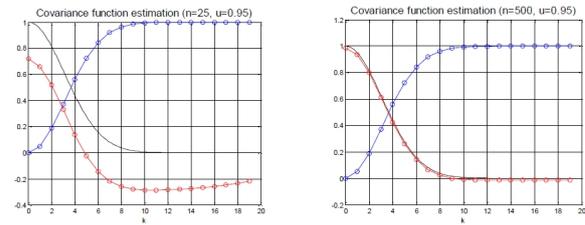[2] *Wikipedia gives quite a good description of Kriging and the semivariogram.*

Figure 3. Simulation study of the traditional covariance function estimate and the semivariogram. For small n-values the traditional approach (left) produces biased covariance estimates and an underestimated Ω -value, but the semivariogram (right) has – if "reversed" – a better correspondence with the true covariance function. 100.000 sample runs. Red: Mean value of the traditional covariance function estimates. Blue: Mean value of the estimated semivariograms. Black: True covariance function, $r(\tau) = 0.95^{\tau^2}$.

The height the semivariogram reaches when it levels off is called the *sill,* and the distance at which the semivariogram levels off to the sill is called the *range* (see Figure 4). It is possible to use the sill directly as an estimate of the variance factor. However, we propose the following strategy which gives us more opportunities:

● Construct the semivariogram (PC programs are available) → determine the range → set $\Omega \approx$ range.

This approximate Ω could then be used to estimate, for example, the standard deviation and a confidence interval for the mean using the formulas derived in Section 3.2. However, the trial-and-error simulation process for estimating critical values requires the complete function.

4. GPS examples

As can be seen in Figure 5, the behaviour of high frequency GPS observations differ from "white noise". They are correlated, which means that they could be analysed with the tools discussed in this article. In Figure 6, a representative GPS covariance function estimation is demonstrated. The oscillations at the tail are evident, and most of the raw $\hat{r}(k)$ values are non-significant because they
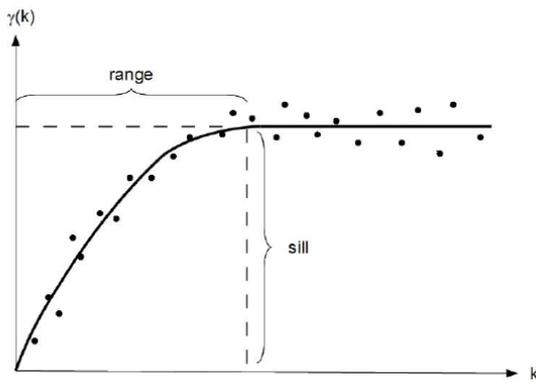
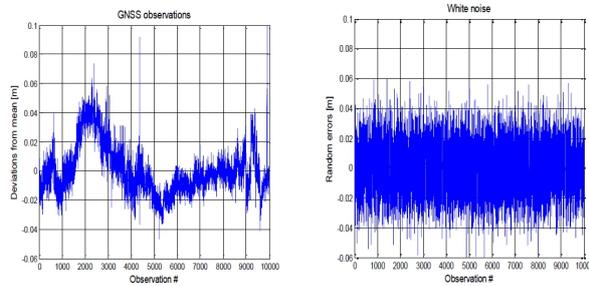**Figure 4.** A semivariogram with sill and range.

Figure 5. Real 1-second GPS height observations (Stora Skuggan, Stockholm, 2011, Leica 1200 GPS receiver) (left) vs. white noise (simulated, uncorrelated, normally distributed random errors, with the same variance as the real data) (right).

lie within the 2D confidence interval. Therefore, as an example, a first-order Gauss-Markov covariance function is fitted to these values (the number of observations is judged to be large enough to "legitimise" this approach).

In Table 3, the same type of study as in Table 2 is presented, but now with real GPS data. The behaviour of the methods is the same. Method III gives confidence intervals that are nearly 2-3 times longer than for the other, overly optimistic methods!

4.1. Uncorrelated additive noise

The GPS observations presented in Figure 7 indicate a new phenomenon: The semivariogram does not reach the origin! The reason is an *additive uncorrelated noise* at the distance $k = 0$. This "displacement" is called the "nugget effect" (see Figure 8). In the presence of such a noise, the covariance function is slightly changed to (cf. Vollat et al., 2002):

$$r_n(0) = \sigma_n^2 + \sigma_s^2$$
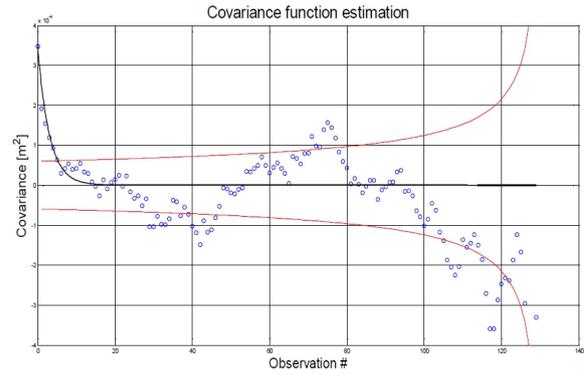$$r_n(\tau_{ij}) = r(\tau_{ij}); i \neq j \qquad (45)$$

Figure 6. Covariance function determined from real GPS height observations, one registration per minute (Tantolunden, Stockholm, 2011, Leica 1200 GPS receiver).
Black: First-order fitted Gauss-Markov covariance function.
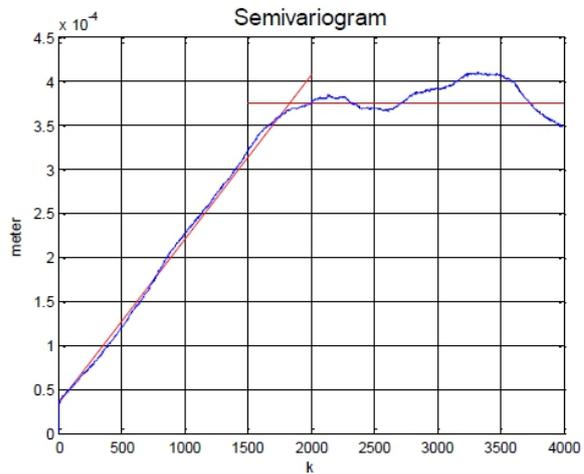Red: 95 % confidence level. Blue: Raw, empirical $\hat{r}(k)$ values.

Figure 7. The estimation of sill and range for real GPS height observations (Stora Skuggan, Stockholm, 2011, Leica 1200 GPS receiver). Sill $\approx 3.75 \cdot 10^{-4}$ m^2(375 mm^2). Range ≈ 1827 seconds (≈ 30 minutes). The noise (the "nugget", see below) has an approximate standard deviation of 6 mm ($s^2 = 35$ mm^2).

where σ_n^2 is the variance of the uncorrelated noise, and $\sigma_s^2 = r(0)$ the variance of the (correlated) "signal". This function generates the variance covariance matrix:

$$\mathbf{C_n} = \sigma_n^2 \mathbf{I} + \sigma_s^2 \mathbf{Q}. \qquad (46)$$

If the *relative noise* is denoted:

$$\beta = \frac{\sigma_n^2}{\sigma_n^2 + \sigma_s^2} = \frac{\sigma_n^2}{\sigma^2} = \frac{nugget}{still}, \qquad (47)$$

Table 3. A study of different methods for the construction of confidence intervals for the mean and standard deviation, using real GPS observations (the same methods as in Table 2, and the same data as in Figure 6). The number of observations is $n = 130$, and the mean value is $\bar{x} = 5.2208$ m. $\hat{\tau}_0 = 3$.

Parameter values	I Traditional approach, using $n-1$	II Approximate, using $n - \Omega$	III Simulated values
Omega (Ω)	1	5.9	5.9
Degrees of freedom	$n - 1 = 129$	$n - \Omega = 124.1$	$n - \Omega = 124.1$
Standard deviation	$\hat{\sigma} = 18.7$ mm	$\hat{\sigma}_o = 19.1$ mm	$\hat{\sigma}_o = 19.1$ mm
Nominal 95% interval ($\bar{x}$) (coverage probability)	5.2208 ± 0.0032 m (≈ 57 %)	5.2208 ± 0.0081 m (≈ 95 %)	5.2208 ± 0.0082 m (≈ 95 %)
Nominal 95% interval (s) (coverage probability)	$I_\sigma = [16.7, 21.3]$ mm (≈ 74 %)	$I_\sigma = [17.0, 21.8]$ mm (≈ 76 %)	$I_\sigma = [15.6, 23.8]$ mm (≈ 95 %)
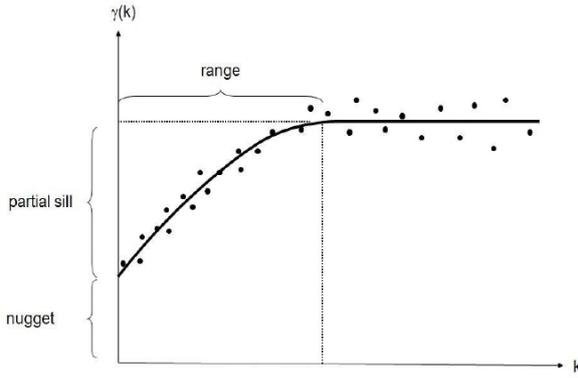
Figure 8. The sill is often composed of two parts: a discontinuity at the origin, denoted *nugget*, and the *partial sill*; added together, these give the sill.

where $\sigma^2 = \sigma_n^2 + \sigma_s^2$ is the variance factor, we get:

$$\mathbf{Q_n} = \beta \mathbf{I} + (1 - \beta)\, \mathbf{Q} \Leftrightarrow \mathbf{C_n} = \sigma^2 \mathbf{Q_n} \qquad (48)$$

which produces a modified Ω with the use of Eq. (24):

$$\Omega_n = \frac{1}{n} \sum \sum Q_n = \beta + (1 - \beta)\, \Omega \qquad (49)$$

With $\beta = 1$ (uncorrelated observations, $\sigma_s^2 = 0$) it yields $\Omega_n = 1$, and for $\beta = 0$ (no noise, $\sigma_n^2 = 0$), we get $\Omega_n = \Omega$. Furthermore: $\sigma_n^2 = \beta \cdot \sigma^2$ and $\sigma_s^2 = (1 - \beta) \cdot \sigma^2$.

For the observations of height in Figure 7 we have:
$n = 8998 \quad \hat{\varepsilon}^T \hat{\varepsilon} = 2.5 \,\mathrm{m}^2 \quad \Omega = 1827$
$nugget = 3.5 \,\mathrm{mm}^2 \quad sill = 375 \,\mathrm{mm}^2,$
which yields:
$\beta = \frac{3.5}{375} \approx 0.09333\ldots \quad \Omega_n = \beta + (1 - \beta) \cdot 1827 \approx 1657 \quad \hat{\sigma} = \sqrt{\frac{2.5}{8998 - 1657}} \approx 18.5$ mm,
and
$\hat{\sigma}_n = \hat{\sigma}\sqrt{\beta} \approx 5.6$ mm $\quad \hat{\sigma}_s = \hat{\sigma}\sqrt{1 - \beta} \approx 17.6$ mm.
The result is only of theoretical interest because the noise is very small in this example. Thus, Ω and $\hat{\sigma}$ are only marginally reduced, but one thing is obvious: The geometrical interpretation of Ω (that $\Omega \approx$ range) is not valid in the presence of noise. The graphically estimated value then has to be reduced according to Eq. (49).

Finally, let us have some discussion about the use of Ω. Before we jump to conclusions, let us also analyse the behaviour of the Northing and Easting co-ordinates.

For Northing we get:
$\beta = 0.1166\ldots \quad \Omega_n \approx 1767$ seconds $\quad \hat{\sigma}_0 \approx 9.9$ mm
$\hat{\sigma}_n \approx 3.4$ mm $\quad \hat{\sigma}_s \approx 9.3$ mm,
and for the Easting co-ordinate:
$\beta = 0.26 \quad \Omega_n \approx 1480$ seconds $\quad \hat{\sigma}_0 \approx 5.0$ mm
$\hat{\sigma}_n \approx 2.5$ mm $\quad \hat{\sigma}_s \approx 4.3$ mm.
The Ω estimates are consistent, with a mean *correlation time* of 27 minutes. This value is an indication of the minimum time span needed between repeated GPS observations to avoid correlation effects. It agrees well with the results of similar studies such as those by Odolinski (2011), and Borre et al. (2000), for example.

The additive noise (3-5 mm) is "background noise" (e.g., ionospheric effects) whereas correlation is a consequence of unmodeled multipath and atmospheric errors in combination with a slowly changing satellite constellation.

5. Summary and Conclusions

Uncompensated correlation results in unrealistic uncertainty measures, however, in this article some methods are presented for compensating correlation in connection with estimation of uncertainty in measurement. Analytical derivations and Monte Carlo simulations form the theoretical platform.

Short to medium range correlation can be handled with these tools, based on the location-model of a stationary time series. Examples of applications are:

- Estimation of GPS observation uncertainty.

- Construction of confidence intervals for the uncertainty estimates and hypothesis testing against theoretical values for these measures.

VERSITA

- The use of this knowledge to construct confidence intervals for observed quantities or functions of these, for example, the distances from GPS positions.

- Hypothesis testing of observations against known values, for example, in GPS calibration or control measurements.

- The development of tolerances for the above-mentioned tests and controls.

The key to these analyses is the quantity " Ω ", the correlation time. Thus, Ω could also be used to design measurement strategies, for example, the minimum time span between repeated GPS observations for control assurance purposes. In some cases, one also has to deal with uncorrelated additive noise, which reduces the correlation effects.

Do not neglect correlation – there are simple tools to handle it!

References

Amiri-Simkooei A. R. and Tiberius C. C. J. M., 2007, Assessing receiver noise using GPS short baseline time series, GPS Solutions 11, 21-35.

Bjerhammar A., 1973, Theory of errors and generalized matrix inverses, Elsevier Scientific Publishing Company, Amsterdam, 420 pages.

Bona P., 2000, Precision, cross correlation, and time correlation of GPS phase and code observations, GPS Solutions 4, 3-13.

Borre K. and Tiberius C., 2000, Time series analysis of GPS observables, In: Proceedings of The 13th International Technical Meeting of the Satellite Division of the Institute of Navigation GPS 2000, September 19-22, 2000, Salt Lake City, Utah.

Bryson A. E. and Henrikson L. J., 1967, Estimation using sampled-data containing sequentially correlated noise, Technical report no. 533, Harvard University, USA.

El Rabbany A. E-S., 1994, The effect of physical correlations on the ambiguity resolution and accuracy estimation in GPS differential positioning, Technical report no. 170, University of New Brunnswick, Canada.

Khodabandeh A., Amiri-Simkooei A. R. and Sharifi M. A., 2012, GPS position time-series analysis based on asymptotic normality of M-estimation, J. Geod. 86, 15-33.

Miller C., O'Keefe K. and Gao Y., 2010, Operational performance of RTK positioning when accounting for the time correlated nature of GNSS phase errors, ION GNSS 2010, September 21 -24, USA.

Odolinski R., 2011, Temporal correlation for network RTK positioning, GPS Solutions 16, 147-155.

Olynik M. C., 2002, Temporal characteristics of GPS error sources and their impact on relative positioning, UCGE reports no. 20162, University of Calgary, Canada.

Persson C-G., 1981, On the estimation of variance components in linear models – with special reference to geodetic applications, Division of Geodesy, Royal Institute of Technology, Stockholm, Sweden.

Petovello M., O'Keefe K., Lachapelle G. and Cannon M., 2009, Consideration of time-correlated errors in a Kalman filter applicable to GNSS, J. Geod. 83, 1, 51–56.

Tienstra J. M., 1947, An extension of the technique of the methods of least squares to correlated observaions, Bulletin géodésique, vol. 6, pp. 301-335.

Vollat U., Landau H. Chen X., Doucet K. and Pagels C., 2002, Network RTK versus single base RTK – understanding the error characteristics, ION GPS 2002, 24-27 September 2002, Portland, USA.

Williams S. D. P., 2003, Offsets in global positioning system time, J. Geophys. Res., 108(B6), 2310, doi:10.1029/2002JB002156, 2003.