Karlsruhe Institute of Technology

Stochastic modelling of GNSS phase observations with focus on correlations

Inaugural dissertation

for the fulfillment of the requirements
for the academic degree of

Doctor of Engineering (Dr.-Ing.)

Accepted by
the Department of Civil Engineering,
Geo and Environmental Sciences of the
Karlsruhe Institute of Technology (KIT)

Submitted by
Gael Kermarrec
Ingénieur des Travaux Géographiques et Cartographiques de l'Etat
From Brest, France

Day of examination: 22.11.2017
Main referee: Prof.Dr.-Ing. Dr. h.c. Bernhard Heck
Co-referee: Prof.Dr.-Ing. habil. Lambert Wanninger

Karlsruhe 2017

Members of the doctoral committee:

Prof. Dr.-Ing. Dr. h.c. Bernhard Heck
Karlsruhe Institute of Technology
Karlsruhe, Germany
Prof. Dr.-Ing. habil. Lambert Wanninger
Technische Universität
Dresden, Germany
Prof. Dr. rer. nat. Martin Breunig
Karlsruhe Institute of Technology
Karlsruhe, Germany
Prof. Dr.-Ing. Maria Hennes
Karlsruhe Institute of Technology
Karlsruhe, Germany
Prof. Dr.-Ing. Stefan Hinz
Karlsruhe Institute of Technology
Karlsruhe, Germany
PD Dr.-Ing. Boris Jutzi
Karlsruhe Institute of Technology
Karlsruhe, Germany

Kermarrec Gael
Stochastic modelling of GNSS phase observations with a focus on correlations

Leibniz University Hanover

Keywords:
Covariance function, stochastic modelling, Matérn covariance function, Least-Squares

“There is no progress toward ultimate freedom without transformation, and this is the key issue in all lives.”

— B.K.S. Iyengar, *Light on Life*

Abstract

The best unbiased estimates of some unknown parameters have the smallest expected mean-squared errors. This condition is only fulfilled when the residuals are weighted with their true variance-covariance matrix. Applied to positioning with Global Navigation Satellite Systems (GNSS) using a Least-Squares estimator, this means that the physical relationship between observations should be known as accurately as possible. However, this condition is never met in real applications. As a consequence, the Least-Squares solution is not trustworthy, the parameter precision estimates being over-optimistic.

Although correlations between GNSS observations were empirically evidenced, they remain disregarded and diagonal variance-covariance matrices describing only the elevation dependency of the observations are still preferred. However, besides computationally demanding procedures by means of iterative operations on observation residuals or empirical fitting with exponential functions, correlations can be physically modeled. This strategy leads to a better understanding and acceptance of the corresponding covariance function, allowing a wider use and as a consequence, a more trustworthy positioning.

Built on a covariance function for GPS phase observations due to tropospheric refractivity fluctuations, this work proposes an improved correlation function for GNSS phase observations. The elevation-weighted Matérn covariance function allows to easily compute fully populated variance-covariance matrices (VCM) that have a physical signification. The correlation structure is controlled by two model parameters: the smoothness and the correlation length. Hence, the impact on the Least-Squares solution and test statistics of wrongly assumed correlation structures up to their omission can be studied.

It is shown that correlations -if present- should not be neglected in order to obtain a realistic precision and a trustworthy solution. When ambiguities are fixed either before or with enough confidence during the final adjustment, taking correlations into account leads to negligible improvements at the estimate level compared with omitting them, particularly for long sessions of observations. However, for short sessions, the impact of a more accurate stochastic model is stronger; the test statistics being less biased allow for an improved least-squares solution. In such cases, fully populated VCM represent additional parameters, like e.g. tropospheric zenith wet delays, that cannot be estimated with enough reliability and without loss of data strength. These conclusions, based both on simulations and case studies, evidence in which cases taking correlations is relevant for an improved Least-Squares solution. Thus, in order to reduce the computational burden associated with correlations, an equivalent diagonal model is additionally proposed that reduces fully populated VCM to diagonal matrices for GNSS positioning scenarios, allowing from a GNSS user perspective a better stochastic model to be applied more easily.

Zusammenfassung

Parameterschätzwerte aus einer Ausgleichung nach der Methode der kleinsten Quadrate sind fehlerfrei mit minimaler Varianz, wenn das funktionale und stochastische Model der Beobachtungen korrekt ist. Für die Positionierung mit Globalen Satellitennavigationssystemen heißt dies, dass eine verlässliche Positionierung mit realistischen Genauigkeitsmaßen nur gewährleistet ist, wenn die stochastischen Eigenschaften zwischen den Beobachtungen perfekt modelliert werden können. Bei realen Anwendungen ist die stochastische Beschreibung der GNSS-Beobachtungen unbekannt und kann nur auf Grundlage einer vorherigen Abhängigkeitsanalyse abgeschätzt werden. Infolgedessen sind die Ergebnisse der Kleinsten-Quadrate-Schätzung nicht fehlerfrei: sowohl die Teststatistiken (Ausreißer- oder Globaltests) als auch die Schätzungen selbst sind beeinflusst.

Obwohl Korrelationen zwischen Beobachtungen empirisch nachgewiesen wurden, beruht die Beschreibung des stochastischen Modells fast ausschließlich auf einer diagonalen Varianz-Kovarianzmatrix, bei der die Korrelationen zwischen Beobachtungen nicht beachtet werden. Die Elevationsabhängigkeit der Varianz wird z.B. durch Sinus- oder Exponentialfunktionen modelliert. Darin benötigte Modellparameter müssen zusätzlich bestimmt werden. Korrelationen können durch rechnerisch anspruchsvolle Verfahren berücksichtigt werden, obwohl ihre Berücksichtigung durch eine geeignete Modellierung einfacher ist. Auf der Basis eines Modells für GPS-Phasen-Korrelationen aufgrund von Indexvariationen in der Troposphäre wird in dieser Arbeit eine physikalische Beschreibung der Korrelationsstruktur von GPS-Beobachtungen vorgeschlagen. Sie ermöglicht vollbesetzte Varianz-Kovarianzmatrizen (VCM) zu berechnen, die auf einer gewichteten Matérn-Kovarianz-Funktion aufgebaut sind. Die Korrelationsstruktur wird durch zwei Parameter bestimmt, die variiert werden können. Durch diese Variation werden die Auswirkungen der fehlerhaft angenommenen Korrelationsstrukturen auf die Kleinste Quadrate-Lösung und die Teststatistiken untersucht.

Es wird gezeigt, dass Korrelationen für eine realistische Lösung nicht vernachlässigt werden sollten. Wenn Ambiguitäten im Vorfeld der Ausgleichung oder während dieser auf eine ganze Zahl fixiert werden können, führt die Berücksichtigung von Korrelationen zu vernachlässigbar kleinen Verbesserungen für die Positionierung. Falls die Float-Ambiguität benutzt und zusammen mit der Position geschätzt wird, werden die Teststatistiken durch beispielsweise den aposteriori-Varianzfaktor und damit die Kleinste Quadrate-Lösung stark verbessert. Die Wirkung ist besonders für sehr kurze Beobachtungszeiten nachvollziehbar, wo vollbesetze VCM zusätzliche Parameter wie troposphärische Zenitlaufzeitverzögerungen auffangen, die aufgrund der kurzen Beobachtungszeit nicht mehr zuverlässig geschätzt werden können. Der Einfluss von VCM sollte unabhängig von der Sessionlänge nicht unterschätzt werden, da genauere und weniger verzerrte Teststatistiken zur Verbesserung der Wiederholbarkeit der Position und ihrer Präzision führen können. Dank eines äquivalenten Diagonalmodells kann die vollbesetzte VCM in eine Diagonalmatrix für GNSS-Positionierungsszenarien überführt werden. Dadurch wird die Rechenlast deutlich reduziert und trotzdem ein adäquates stochastisches Modell verwendet.

Preface

This article-based cumulative thesis consists of 6 chapters including an introduction and a conclusion. It is associated with the following peer-reviewed publications which are given in the appendix:

Kermarrec G, Schön S (2014) On the Matérn covariance family: a proposal for modelling temporal correlations based on turbulence theory. Journal of Geodesy 88(11):1061-1079

Author's contributions. The first author derived the simplification based on a model from the second author (Schön and Brunner 2008). She designed and carried out the numerical investigations, and wrote the manuscript for the article. The second author discussed the results and contributed to the improvement of the manuscript by suitable comments and corrections.

Kermarrec G, Schön S (2016) Taking correlation in GPS Least-Squares adjustments into account with a diagonal covariance matrix. Journal of Geodesy 90(9):793-805

Author's contributions. The idea of the equivalent diagonal model is from the first author, based on the results of Luati and Proietti for the regression case. She applied these results to the GPS positioning cases and carried out numerical examples. The manuscript was written by the first author. The second author discussed the results and contributed to the improvement of the manuscript by suitable comments and corrections.

Kermarrec G, Schön S (2017a) Apriori fully populated covariance matrices in Least-Squares adjustment – case study: GPS relative positioning. Journal of Geodesy 91(5):465-484

Author's contributions. The first author derived the function for elevation-dependent correlations. She carried out the simulations and the sensitivity analysis based on her methodology to study the impact of wrongly specifying the correlation structure. She applied the results to several baselines and wrote the manuscript. The second author discussed the results and contributed to the improvement of the manuscript by suitable comments and corrections.

Kermarrec G, Schön S, Kreinovich V (2017b) Possible explanation of empirical values of the Matérn smoothness parameter for the temporal covariance of GPS measurements. Applied Mathematical Science 11(35):1733-1737

Author's contributions. The first author discussed and debated results on the Matérn model with the third author who wrote the introduction and the conclusion. The first author wrote parts of the physical explanations and contributed to the improvement of the manuscript.

Kermarrec G, Schön S (2017c) On modelling GPS phase correlations: a parametric model. Acta Geophysica et Geodaetica https://doi.org/10.1007/s40328-017-0209-5

The first author derived the function for elevation-dependent correlations. She carried out the example to study the impact of fully populated matrices with respect to diagonal VCM. The second author discussed the results and contributed to the improvement of the manuscript by suitable comments and corrections.

Kermarrec G, Schön S (2017d) Taking correlations into account: a diagonal correlation model. GPS solution 21(4):1895-1906

Author's contributions. The first author derived the diagonal correlation model. She carried out the simulations and the case study to analyse the sensitivity analysis of the exponential factor on the Least-Squares results. She showed in which cases the model is an alternative to fully populated VCM and wrote the manuscript. The second author discussed the results and contributed to the improvement of the manuscript by suitable comments and corrections.

Kermarrec G, Schön S (2017e) Fully populated VCM or the hidden parameter. Journal of Geodetic Science 7(1):151-161

Author's contributions. The first author derived the concept of the hidden parameter, made the simulations and the data analysis and wrote the paper. The second author discussed the results and contributed to the improvement of the manuscript by suitable comments and corrections.

Table of contents

Acronyms

ADOP	Ambiguity Dilution of Precision
DCM	Diagonal Correlation Model
DGLSE	Diagonal Global Least-Squares Estimator
FGLSE	Feasible Global Least-Squares Estimator
GLSE	Global Least-Squares Estimator
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
IPP	Ionospheric Pearce Point
LS	Least-Squares
MLE	Maximum Likelihood Estimation
MSE	Mean Square Error
OLSE	Ordinary Least-Squares Estimator
OMC	Observed Minus Computed
RMS	Root Mean Square
VCM	Variance-Covariance Matrix
VLBI	Very Long Baseline Interferometry

1. Introduction

1.1. Motivation

Estimates such as position, tropospheric or clock parameters are obtained from GNSS observations as part of the parameter estimation using a Least-Squares adjustment. These parameters and their associated precision are expected to be as accurate and trustworthy as possible. This statement holds true if the observations are free of errors and the model that describes the relationship between the observations and the estimates is perfect.

However, while the functional model of GNSS observations can be considered as well-known, this is only partially true for the stochastic model which remains improvable. The variance-covariance matrices of the observations are usually reduced to a diagonal form, sometimes even taken to identity. Because of a lack of understanding of the impact of correlations, they remain neglected, leading to an untrustworthy Least-Squares solution as well as biased significance tests. Additional reasons for this misspecification are unsatisfactory descriptions of the covariance of GNSS observations, besides a computational problem due to the inversion of fully populated matrices.

1.2. Research objectives

The three main objectives of this thesis can be summarized as follows:

- To propose a covariance function that describes in a simple but physically plausible way the correlation structure of GNSS observations. This function must be easy to understand in order to motivate GNSS users to take correlations into account, while also being accurate and adaptive.
- To give a solid understanding of the effects of misspecifying the stochastic model based on simulations and case studies. The description of the errors or biases includes quantities such as the aposteriori variance factor, the precision by means of the cofactor matrix of the estimates, as well as the float solution.
- To derive a simplification of the fully populated VCM so that correlations can be resumed in a simple diagonal matrix. This way, a link between variance and covariance models can be built.

By meeting these objectives, this work aims to give the reader arguments to decide whether taking correlation into account is meaningful or not in his/her specific application, i.e. an engineer perspective based on sensitivity analysis is adopted rather than trying to assess as accurately as possible the correlation structure of GNSS observations by means of data analysis.

1.3. Outline of this thesis

The thesis is structured as follows:

Chapter 2 summarizes the main concepts for stochastic modelling with a particular focus on how misspecifications impact some carefully chosen Least-Squares quantities. Chapter 3 proposes a detailed description of the variance and covariance models used in previous research. A classification is proposed to link the different models with each other. Based on the results of other research groups, the open questions about the impact of the stochastic model are hence raised. Chapter 4 describes the proposed correlation function for GNSS phase observations. The results of the implementation of the derived fully populated VCM in a positioning Least-Squares adjustment are analysed. Simulations are used to quantify the impact of an incorrect correlation structure and a case study validates the conclusions of the sensitivity analysis. Chapter 5 provides a better understanding of how fully populated VCM act. A way to take correlations into account in a diagonal matrix is developed, allowing to draw a parallel between the commonly used exponential variance model and the proposed correlation model. Chapter 6 concludes this work by providing recommendations and perspectives for further research. The papers associated with this work are provided in the Appendix.

2. Stochastic modelling and impact of its misspecification in Least-Squares adjustments

The domain of stochastic modelling is vast and covers many different applications from financial mathematics to call-center provisioning or disease treatment options. The focus of this thesis is on providing a more trustworthy positioning with GNSS observations. After a general introduction that explains the underlying concept of stochastic modelling, this chapter provides some important mathematical definitions. The principles of Least-Squares are shortly presented, focusing on the quantification of the impact of a wrong specification of the variance-covariance matrix (VCM). Commonly used covariance functions are described. Finally, the equivalence between stochastic and augmented functional model is introduced.

2.1. Stochastic model

The word "stochastic" derives from the Greek word to aim, to guess and means "random" or "chance". The antonym is "sure", "deterministic," or "certain" (Oxford dictionary 2016). The purpose of stochastic modelling is to predict outcomes that have a certain degree of unpredictability. On the contrary, in a deterministic model, the output is fully determined by the parameter values and the initial conditions.

To be useful, a stochastic model must reflect nearly all relevant aspects of the process under study. In addition, it must be easily computable to allow the deduction of important implications about the phenomenon (Taylor and Karlin 1998). For a wide acceptance of the proposed model, the simplest mathematically and physically plausible description should be preferred.

Stochastic modelling as applied to Global Navigation Satellite Systems (GNSS) observations means being able to describe the relationships between all different observations, whether code or phase measurements. The geometry, the satellite elevation, the atmospheric conditions, the receiver noise or multipath are possible factors that are to be taken into account. Getting a simple but accurate statistical description of GNSS observations aims thus to improve the derivation of computed quantities such as the position itself and its associated precision evaluated via Least-Squares estimation. The stochastic model can be summarized in the variance-covariance matrix (VCM) of the observations which is mathematically defined in the next paragraph.

2.1.1. Variance and covariance functions

Let y_1 and y_2 be two realizations of a time dependent stochastic process Z at time t_1 and t_2 respectively. The variance of y_1 is defined as $\sigma_{y_1}^2 = E\left[\left(y_1 - E(y_1)\right)^2\right]$ where $E(\cdot)$ stands for expectation. It can be considered as the autocovariance or covariance of the vector with itself, $Cov(y_1, y_1)$ whose general formulation reads $Cov(y_1, y_2) = E\left[\left(y_1 - E(y_1)\right)\left(y_2 - E(y_2)\right)\right]$. It is thus a measure of similarity or dependence between the two realizations.

As the knowledge of the underlying probability function of the process is difficult to obtain concretely (Gaspari and Cohn 1999), a second and equivalent definition of the covariance is used instead of the theoretical one (Loeve 1963 p.466-467). It states that $Cov(y_1, y_2)$ is a covariance function if for each integer *m* and for each choice of points $y_1,...y_m$ in the set where Cov is defined (the Euclidian set for instance), the matrix $\mathbf{W}$ with elements $w_{i,j} = \{Cov(y_i, y_j)\}$ is positive semidefinite. The correlation function is a normalized covariance function, i.e. $Corr(y_1, y_2) = \frac{1}{\sigma_{y_1}\sigma_{y_2}} Cov(y_1, y_2)$.

In the time domain, a stationary process depends only on the time increment, i.e. $\tau = t_1 - t_2$. In the following, only stationary time series are considered and the covariance function simplifies to $Cov(y_1, y_2) = Cov(\tau)$.

Using this definition, as well as the Wiener-Khintchine theorem (Chatfield 1989), the covariance function can be linked with its spectral representation. This is particularly useful in spatial statistics or Kriging (Gelfand et al. 2010) as the power spectral density can be measured concretely using spectrum analysers. If $S(s)$ is the amount of power allocated on average to $e^{2\pi is\tau}$ with frequency s (Rasmussen and Williams 2006), the covariance is defined as $Cov(\tau) = \int S(s)\, e^{2\pi is\tau} ds$.

2.1.2. Mean Square differentiability of a process

The stochastic process $Z(t)$ is mean square differentiable if the limit of the covariance function of $\frac{Z(t+h)-Z(t)}{h}$ as $h \to 0$ exists and is finite. This property is related to the limit of the covariance function at the origin and thus to the smoothness of the process (Adler 1981 ch.3, Stein 1999). If we suppose that there exists a $\beta \in (0,2]$ such that the covariance function $Cov(t)$ can be approximated at the origin by $1 - Cov(t) \approx t^{\beta}$ as $t \to 0$, the corresponding process of dimension d has a fractal dimension $D = d + 1 - \beta/2$. The larger β and D (i.e. $2 - \beta/2$ for a time series), the smoother the realizations (Gelfand et al. 2010). The smoothness of a process is thus related to the behaviour of the spectral density at high frequencies.

2.2. Least-Squares adjustment, basic principles

In case of GNSS positioning, the number of observations n is larger than the number of unknown u to be estimated so that the Least-Squares method is adequate to assess the variables of interest which are the position and eventually tropospheric or clock parameters. The mathematical relationship between the observations and the parameters to be estimated is linearized, and the term “linearized Least-Squares” is thus employed (Koch 1999).

In the following, the generalized linear functional model or observation equation model is defined by:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}, \tag{1}$$

where $\mathbf{y}$ is the n-dimensional observed minus computed vector, $\mathbf{A}$ the non-stochastic $(n \times u)$ design matrix with full column $rank(\mathbf{A}) = u$, $\mathbf{x}$ the u-dimensional parameter vector to be estimated. The n-dimensional vector called $\mathbf{v}$ represents the unknown errors in the measurements.

2.2.1. Ordinary Least-Squares Estimator (OLSE)

The Ordinary Least-Squares Estimator (OLSE) assumes homoscedasticity of the errors, i.e. equal variance. If the errors are further zero-mean and statistically independent, the minimization of the sum of squared residuals leads to the OLSE given by $\hat{\mathbf{x}} = (\mathbf{A}^{\mathbf{T}}\mathbf{A})^{\mathbf{-1}}\mathbf{A}^{\mathbf{T}}\mathbf{y}$. Under the previous assumptions, the expectation of the estimator is the same as the true parameter.

This estimator has three major properties:

- Unbiasedness: the estimator is said to be unbiased when neither over- nor under-estimation of the truth occurs. Violation of this assumption can be due to a non-linearity of the model.
- Efficiency: If the errors have equal variances, the OLSE has minimum variance of all unbiased estimators, i.e. it is the most efficient estimator referring to the accuracy of the estimates produced by the estimator.
- Consistency: If the sample size increases, the estimator is consistent and converges to the true value of the parameter.

The errors are assumed to be zero-mean. If this assumption does not hold true, the Least-Squares solution may be biased (Berk 2004). A possible correction is to add a constant term to the estimates. For normally distributed errors, the OLS is the most efficient of all unbiased estimators (Koch 1999). If this property is violated, hypothesis tests may not follow the assumed t- or F-distributions, particularly for small samples (Cohen et al. 2003). However, without the assumption of normality, the OLSE remains unbiased, consistent and the most efficient in the class of *linear* unbiased estimators.

The situation changes if the homoscedasticity assumption is no longer valid and errors have different variances. As long as the errors remain independent, the OLSE will stay unbiased and consistent but not efficient anymore (Weisberg 2005). Several tests exist to confirm heteroscedasticity. We refer for example to the White Test (White 1980). In case of GNSS observations, heteroscedasticity was empirically proven by Bischoff et al. (2005).

2.2.2. Generalized Least-Squares Estimator (GLSE)

In case of heteroscedasticity and correlations between observations, the error vector can be described by $E\left(\boldsymbol{\varepsilon}\right)=0,\ E\left(\boldsymbol{\varepsilon\varepsilon}^{\mathbf{T}}\right)=\sigma^2\mathbf{W}_0$, where $\mathbf{W}_0$ is a $\left(n\times n\right)$ positive definite, symmetric fully populated variance-cofactor matrix (VCM) of the observations, σ^2 the apriori variance factor. In this case, the Aitken Theorem (Aitken 1935, Rao and Toutenburg 1999, p105) states that:

The Generalized Least-Squares Estimator (GLSE) is unbiased and is the best linear unbiased estimator (BLUE) for $\mathbf{x}$. It is given by

$$\hat{\mathbf{x}}_0=\left(\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{\mathbf{-1}}\mathbf{A}\right)^{\mathbf{-1}}\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{\mathbf{-1}}\mathbf{y}=\mathbf{K_0}\mathbf{y} \tag{2}$$

where $\mathbf{K_0}$ is the slope matrix that projects the observations to the estimate space. The cofactor matrix of the unknowns reads:

$$\mathbf{W}_{0\hat{\mathbf{x}}\hat{\mathbf{x}}}=\left(\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{\mathbf{-1}}\mathbf{A}\right)^{\mathbf{-1}} \tag{3}$$

If $\mathbf{v}_0=\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}_0$ is the $\left(n\times 1\right)$ residual vector, the following estimator for the covariance matrix of the unknowns is available, also called aposteriori estimator:

$$\hat{\mathbf{C}}_{0\hat{\mathbf{x}}\hat{\mathbf{x}}}=\hat{\sigma}^2_{\mathbf{W_0}}\left(\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{\mathbf{-1}}\mathbf{A}\right)^{\mathbf{-1}}=\hat{\sigma}^2_{\mathbf{W_0}}\mathbf{W}_{0\hat{\mathbf{x}}\hat{\mathbf{x}}}$$

$$\hat{\sigma}^2_{\mathbf{W_0}}=\frac{\left(\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}_0\right)^{\mathrm{T}}\mathbf{W}_0^{-1}\left(\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}_0\right)}{n-u}=\frac{\mathbf{v}_0^{\mathrm{T}}\mathbf{W}_0^{-1}\mathbf{v}_0}{n-u} \tag{4}$$

These estimators are unbiased i.e., $E\left(\hat{\sigma}^2_{\mathbf{W_0}}\right)=\sigma^2,\ E\left(\hat{\mathbf{W}}_{0\hat{\mathbf{x}}\hat{\mathbf{x}}}\right)=\sigma^2\mathbf{W}_{0\hat{\mathbf{x}}\hat{\mathbf{x}}}$ (Koch 1999) if the assumptions hold true.

Feasible Least-Squares Estimator (FLSE)

The VCM being unknown, the GLSE can never be reached, i.e. the variance and covariance are non-estimable. Indeed, the number of redundant observations in the stochastic model is $\frac{1}{2}(n-u)(n-u+1)$. However, the number of elements of the VCM to be estimated is $\frac{1}{2}n(n+1)$, thus not all components can be estimated (Koch 1999, Xu et al. 2007). Therefore the FLSE (Greene 2003) is used in practice where $\mathbf{W}_0$ is replaced by its estimate $\hat{\mathbf{W}}$.

Transformation of the FLSE into OLSE

The FGLSE can be transformed into an OLSE by computing a $(n \times n)$ transformation matrix $\mathbf{T}_m$ such as the covariance matrix of the observations is $\sigma^2\mathbf{I}$ where $\mathbf{I}$ is the identity matrix. This practice, also called whitening, makes use for instance of a Cholesky factorization of the Hermitian positive-definite covariance matrix $\hat{\mathbf{W}}$ (Koch 1999) where $\hat{\mathbf{W}} = \mathbf{T}_m\mathbf{T}_m^{\mathbf{T}}$ and $\mathbf{T}_m$ is a regular lower triangular matrix. By transforming $\mathbf{A}_{\mathbf{wh}} = \mathbf{T}_m^{\mathbf{T}}\mathbf{A}$, $\mathbf{y}_{\mathbf{wh}} = \mathbf{T}_m^{\mathbf{T}}\mathbf{y}$ and $\mathbf{v}_{\mathbf{wh}} = \mathbf{T}_m^{\mathbf{T}}\mathbf{v}$, an equivalent formulation to the GLSE can be obtained. The VCM of $\mathbf{v}_{\mathbf{wh}}$ is a diagonal matrix with equal weights. This procedure, as well as simplified decorrelation strategies in particular cases of autoregressive processes, was for instance proposed in Schuh et al. (2014).

2.3. Effect of misspecifying the stochastic model in the Least-Squares adjustment

The knowledge of the VCM of the observations is a central point in FGLSE. The reasons for an incorrect specification of the VCM are manifold, from a need for a simplified computation to a simple lack of knowledge of the correlation structure. In the following section, we review the effect of an approximated VCM.

The misspecification has an effect on the Least-Squares results, leading to a statistically incorrect solution. We call $\hat{\mathbf{P}} = \mathbf{P}_0 + \mathbf{\Delta P}$ the new weight matrix, where $\mathbf{P}_0$ is the original one, i.e. $\mathbf{P}_0 = \mathbf{W}_\mathbf{0}^{-1}$, $\hat{\mathbf{P}} = \hat{\mathbf{W}}^{-1}$ and $\hat{\mathbf{W}} = \mathbf{W}_0 + \mathbf{\Delta W}$.

- Aposteriori variance factor

A misspecification of the covariance matrix leads to errors in the variance estimator. The corresponding bias is given by Rao and Toutenburg (1999) or Xu (2013):

$$E\left(\hat{\sigma}^2_{\hat{\mathbf{W}}}\right) = \sigma^2 + tr\left\{\left(\mathbf{I} - \hat{\mathbf{W}}^{-1}\mathbf{A}\left(\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathbf{T}}\right)\hat{\mathbf{W}}^{-1}\mathbf{\Delta W}\right\}\frac{\sigma^2}{n-u}. \tag{5}$$

For the regression case and assuming that the regression vector forms an orthonormal set (i.e. $\mathbf{A}^{\mathbf{T}}\mathbf{A} = \mathbf{I}$), Watson (1955) obtained bounds of the variance factor showing a strong dependence on the roots of the design matrix.

- Parameters

Following Kutterer (1999) the difference in the estimated parameter vector due to $\mathbf{\Delta W}$ writes:

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}_\mathbf{0} - \left(\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{\Delta P}\mathbf{v}_\mathbf{0} = \hat{\mathbf{x}}_\mathbf{0} + \mathbf{\Delta x} \tag{6}$$

The error term $\Delta\mathbf{x}$ of Eq. (6) can be linearized. However, as it depends on the residuals and thus on the data sets used, a more general formulation of the error in the estimates cannot be given and only bounds are assessed.

- Cofactor matrix of the parameters

A formula for the difference of the cofactor matrix of the unknowns when the cofactor matrix is misspecified can be found in Kutterer (1999) or Strand (1974) and reads:

$$\hat{\mathbf{W}}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} - \hat{\mathbf{W}}_{0\hat{\mathbf{x}}\hat{\mathbf{x}}} = -\mathbf{W}_{0\hat{\mathbf{x}}\hat{\mathbf{x}}}\mathbf{A}^{\mathbf{T}}\Delta\mathbf{P}\left(\mathbf{I} + \mathbf{W}_{\mathbf{0}}^{\mathbf{-1}}\Delta\mathbf{P}\right)^{\mathbf{-1}}\mathbf{A}\mathbf{W}_{0\hat{\mathbf{x}}\hat{\mathbf{x}}} \tag{7}$$

Kutterer (1999) linearized Eq. (7). However, if $\Delta\mathbf{P}$ is assumed to be small, an accurate knowledge of the true VCM is necessary, making the bound of the error term hardly usable in practice.

- Mean-Squared Errors (MSE)

Calling $E\left(\mathbf{x}_{\mathbf{0}}\right) = \mathbf{x}_{\mathbf{ref}}$, the Mean Squared Errors of the estimates, differences can be written as: $MSE_{\hat{\mathbf{x}}\text{-}\mathbf{x}_{ref}} = MSE_{\hat{\mathbf{x}}\text{-}\mathbf{x}_0} - MSE_{\hat{\mathbf{x}}_{ref}\text{-}\mathbf{x}_0}$ or

$$MSE_{\hat{\mathbf{x}}\text{-}\mathbf{x}_{ref}} = tr\left(\left[\hat{\mathbf{W}}^{\mathbf{-1}}\mathbf{A}\left(\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{\mathbf{-1}}\mathbf{A}\right)^{-2}\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{\mathbf{-1}}\right]\mathbf{W}_0\right) - tr\left(\left(\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{\mathbf{-1}}\mathbf{A}\right)^{-1}\right) \tag{8}$$

Strand (1974) proposed an estimation of the Mean Squared Error based on the Frobenius norm of $\Delta\mathbf{W}$ and the Rayleigh quotient. The ratio $R_{MSE} = \frac{MSE_{\hat{\mathbf{x}}\text{-}\mathbf{x}_{ref}}}{MSE_{\mathbf{x}_{ref}-\hat{\mathbf{x}}_0}}$, also called relative efficiency of the estimator, reaches the null value for the true VCM and can thus be studied to quantify the effect of an incorrectly estimated VCM. By linking the MSE and the Root Mean-Square (RMS) of the estimate differences, a parallel between simulations involving a known $\mathbf{W}_0$ approximated by different $\hat{\mathbf{W}}$ and results from data analysis can be drawn.

Other effects of misspecifying $\hat{\mathbf{W}}$ were studied in a theoretical way from a geodetic perspective. We cite for example Linkwitz (1961) and Wolf (1961). Later Hahn and van Mierlo (1987) focused on statistical tests whereas other studies deal for instance with collocation (Xu 1991). The previously described effects of the apriori VCM on the Least-Squares results are summarized in Figure 1.

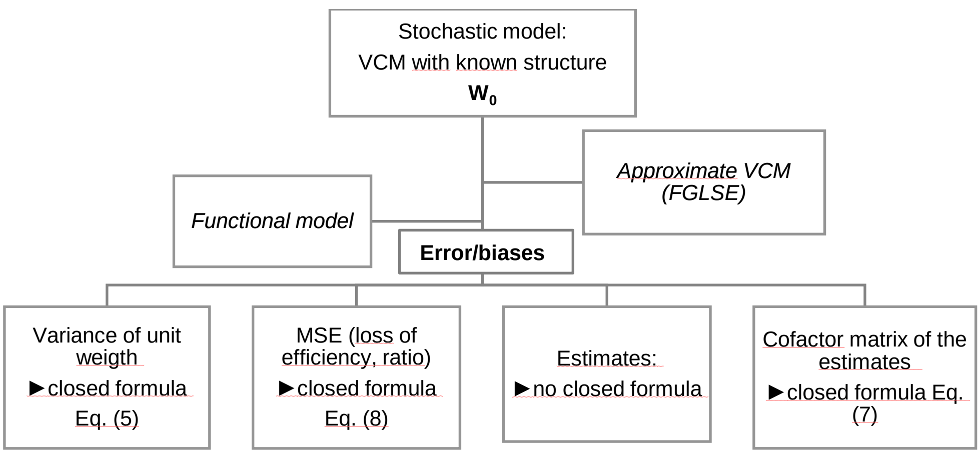


Fig.1 Effect of misspecification of the VCM on Least-Squares quantities, induced errors

2.4. Construction of fully populated VCM

A comprehensive understanding of the stochastic model is a key point in FGLSE. Some procedures have been proposed to estimate the VCM from an iterative perspective by analysing the residuals of the Least-Squares adjustment. The second strategy consists in modelling the statistical relationship between the observations, i.e. developing a stochastic model.

2.4.1. Iterative procedures

One possibility for estimating the VCM iteratively is to use variance component estimation (VCE) strategies such as MINQUE (Minimum Norm Quadratic Unbiased Estimation) for which only assumptions about the first and second order moments of the observables are to be made. BIQUE (Best Invariant Quadratic Unbiased Estimation, see Grafarend 1974, Grafarend and Awange 2014, Schaffrin 1983) which has been studied for normally distributed observations (Caspary 1987, Koch 1999), REML (Restricted Maximum Likelihood, Koch 1986, Hartley and Rao 1967) or LS-VCE (Least-Squares Variance Component Estimation, Schaffrin 1981, Teunissen and Amiri-Simkooei 2008) are possible alternatives.

Here, we briefly present two methods used in geodetic applications: MINQUE (Minimum Norm Quadratic Unbiased Estimation) as proposed in Wang et al. (1998), which is a particular case of the LS-VCE method popularized in GNSS applications by e.g. Amiri-Simkooei (2007).

MINQUE

The covariance matrix of the observations is expressed as a linear combination of *k* so-called accompanying matrices $\mathbf{T_i}$: $\hat{\mathbf{W}} = \sum_{i=1}^{k} \theta_i \mathbf{T_i}$ where $\boldsymbol{\theta} = [\theta_i, ...\theta_k]$ the estimated variance-covariance components of the measurements have to be estimated, i.e. $\hat{\boldsymbol{\theta}}_\mathbf{i} = [\hat{\theta}_i, ...\hat{\theta}_k]$ (Rao 1971). An iterative procedure gives $\hat{\boldsymbol{\theta}}^{\mathbf{j+1}} = \mathbf{S}^{-1}\left(\hat{\boldsymbol{\theta}}^{\mathbf{j}}\right)\mathbf{q}\left(\hat{\boldsymbol{\theta}}^{\mathbf{j}}\right),\ j = 1, 2, ...$ assuming an approximated value for $\hat{\boldsymbol{\theta}}_\mathbf{i}^{\mathbf{0}}$ with

$$s_{ij} = tr\left(\mathbf{R}\mathbf{T}_i\mathbf{R}\mathbf{T}_j\right), \mathbf{R} = \hat{\mathbf{W}}^{-1}\left[\mathbf{I} - \mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\hat{\mathbf{W}}^{-1}\right].$$

$$q_i = \mathbf{y}^{\mathrm{T}}\mathbf{R}\mathbf{T}_i\mathbf{R}\mathbf{y}$$

When the convergence is given, the iteration is stopped as soon as $\left|\hat{\boldsymbol{\theta}}_\mathbf{i}^{\mathbf{j+1}} - \hat{\boldsymbol{\theta}}_\mathbf{i}^{\mathbf{j}}\right| < \delta$ where δ is the convergence tolerance.

LS-VCE

Similarly to the MINQUE procedure, the LS-VCE starts with the linear model of observations equations with the decomposition $\hat{\mathbf{W}} = \sum_{i=1}^{k} \theta_i \mathbf{T_i}$. However, the unknown covariance components θ_i are estimated using LS and $\theta_i = \mathbf{N}^{-1}\mathbf{l}$ where $n_{ij} = \frac{1}{2} tr\left(\mathbf{T_i}\hat{\mathbf{W}}^{-1}\mathbf{P}^{T}\mathbf{T_j}\hat{\mathbf{W}}^{-1}\mathbf{P}\right),\ l_i = \frac{1}{2}\mathbf{y}^{\mathrm{T}}\hat{\mathbf{W}}^{-1}\mathbf{P}\mathbf{T_i}\hat{\mathbf{W}}^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{y}$ are the elements of $\mathbf{N}$ and $\mathbf{I}$, respectively, and $\mathbf{P} = \mathbf{I} - \mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\hat{\mathbf{W}}^{-1}$ is an orthogonal projector. Iterations should be followed until the convergence tolerance is reached as for the MINQUE procedure. Once more, a starting (co)variance structure has to be assumed. The estimated VCM is then approximated as the inverse of $\mathbf{N}$.

The previously described methods are computationally demanding, the convergence being not obligatory given. Moreover, the positive definiteness of the VCM is rarely met so that the resulting matrices lose their physical meaning (Rao and Kleffe 1988). The model can be constrained by choosing a special structure for the VCM based for instance on a Kronecker structure (Groeneveld 1994, Shaw and Geyer 1997). Alternatively, a triangular decomposition of the covariance matrix and a maximum likelihood function as the/an objective function can be used to guarantee the positive definiteness (Xu et al. 2007).

2.4.2. Modelling the covariance between observations

A second possibility for constructing the covariance matrices without using iterative procedures is based on a stochastic modelling of the measurement errors. For this, families of positive definite covariance functions were defined (Yaglom 1987). They can be chosen based on model identification (Rasmusseum and Williams 2006) or some apriori knowledge given by physical assumptions. Another possibility is to fit predefined functions to the autocorrelation of the observations as done in Kriging (Cressie 1993, Cressie et al. 1999) or used in GNSS applications by El-Rabbany (1994), El-Rabbany and Kleusberg (2003), Howind et al. (1999). However, this last method is empirical and can show some limitations as the behaviour of the covariance function at the origin is hardly visually identifiable (Stein 1999). We here present a few commonly used covariance functions. We restrict ourselves to temporal correlations, i.e. GNSS observations being time dependent.

Matérn family

This two-parameters family (Matérn 1960) states that the correlation between $Z(t)$ and $Z(t+\tau)$ decreases as τ increases. It additionally takes into consideration that different processes may exhibit different degrees of smoothness. The covariance function reads:

$$Cov(\tau) = \phi(\alpha\tau)^{\nu} K_{\nu}(\alpha\tau) \quad (9)$$

where $\nu > 0$, $\alpha > 0$ are constant parameters. The scalar parameter $\phi > 0$ is chosen so that the variance equals 1; ν is defined as the smoothness of the time series (Sec. 2.2). The constant α is the inverse of the correlation length and indicates the rate of decay of the function with increasing time (Journel and Huijbregts 1978). The modified Bessel function of order ν (Abramowitz and Stegun 1972) is denoted by K_{ν}.

Special cases arise when the smoothness factor ν is taken to:

- ½: exponential covariance function. This function is corresponding to an AR(1) process (autoregressive process of first order). The corresponding process is not mean-square differentiable.
- 1: also called Markov process of first order. This process is often used in geodesy to analyze gravitational fields (Meier 1981, Grafarend 1976).
- ∞: squared exponential covariance function. Time series having this covariance structure would be indefinitely differentiable and thus exactly predictable at all values (Stein 1999). As it is physically implausible, this covariance function should be employed with care (Stein 1999, Handcock and Wallis 1994, see also Koivunen and Kostinski 1999).

More details on this family can be found in Guttorp and Gneiting (2005) or Kermarrec and Schön (2014) with corresponding examples. Fuentes and Smith (2001) extended this function to account for non-stationarity.

Alternatively, the powered exponential model (Diggle and Ribeiro 2007) can be used where $Cov(\tau) = \phi e^{-\left(\frac{\tau}{\theta}\right)^{\gamma}}$, $\gamma \in (0,2]$, $\theta > 0$ being a scaling parameter a.

The Cauchy family (Gneiting and Schlather 2004) has a similar parametrization with $Cov(\tau) = \phi\left(1+\left(\frac{\tau}{\theta}\right)^{\alpha}\right)^{-\beta/\alpha}$, $\theta > 0$ and $\beta > 0$ a long memory parameter describing the correlation decay as $\tau \to \infty$.

Other examples

It may be useful to add some cosine function to the covariance function to describe hole effect, i.e. small negative correlations as τ increases (Gerland et al. 2010, Chiles and Delfiner 1999). One example is the exponentially damped cosine function $Cov(\tau) = \phi e^{-\left(\eta \frac{\tau}{\theta}\right)} \cos\left(\frac{\tau}{\theta}\right),\ \eta \geq \frac{1}{\tan\frac{\pi}{2d}}$, d being the dimension of the process.

Covariance functions with compact support (Gneiting 2002), as for instance the spherical correlation function $Cov(\tau) = 1 - \frac{3}{2}\frac{\tau}{\theta} + \frac{1}{2}\left(\frac{\tau}{\theta}\right)^3$ for $\tau \leq \theta$ and $Cov(\tau) = 0$ for $\tau > \theta$, are popular in geological applications but only defined up to 3 dimensions. As the correlations become exactly 0, they are said to have computational advantages. Stein (1999) argues however that as this function is only once differentiable at $\tau = \theta$, the estimation of the parameters with likelihood methods may be problematic.

Estimation of parameters via Maximum Likelihood

The parameters of the proposed covariance function can be computed via Maximum Likelihood Estimation, provided that a first guess of the parameters is supplied (Stein 1999, Handcock and Wallis 1994). The log likelihood reads

$$L(\theta, \mathbf{Z}) = -\tfrac{1}{2}\log\left|\hat{\mathbf{W}}(\theta)\right| - \tfrac{1}{2}\mathbf{Z}^{\mathrm{T}}\mathbf{P}(\theta)\mathbf{Z} \qquad (10)$$

where θ represents a set of parameters.

$\mathbf{P}(\theta) = \hat{\mathbf{W}}^{-1}(\theta) - \hat{\mathbf{W}}^{-1}(\theta)\mathbf{X}\left(\mathbf{X}^{\mathrm{T}}\hat{\mathbf{W}}^{-1}(\theta)\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\hat{\mathbf{W}}^{-1}(\theta)$ and $\mathbf{X}$ is the design matrix for trend regression.

$L(\theta, Z)$ has to be minimized with respect to the unknown parameters of the covariance function. Possible biases due to small samples can be avoided by using Restricted Maximum Likelihood (REML) estimation as described in Zimmerman and Zimmerman (1991).

2.5. Equivalence of stochastic and functional models

Following Blewitt (1998), it can be shown that the estimation of extra parameters in the functional model (i.e. augmented functional model) is equivalent to changing the stochastic model of the observations.

If Eq. (1) is augmented by extra parameters $\mathbf{z}$, the corresponding Least-Squares problem reads $\mathbf{y} = \mathbf{Ax} + \mathbf{Bz} + \mathbf{v}$, where $\mathbf{B}$ is the design matrix describing the mathematical relationship between $\mathbf{z}$ and the observations. In case of GNSS adjustments this may be an estimation of tropospheric zenith wet delays (Hofmann-Wellenhof et al. 2001), clock parameters or ambiguities. Writing the previous equation in terms of partitioned matrices, it can be shown that $\hat{\mathbf{x}}_0 = \left(\mathbf{A}^{\mathrm{T}}\mathbf{W}_0^{-1}\mathbf{PA}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}_0^{-1}\mathbf{Py},\ \mathbf{P} = \mathbf{I} - \mathbf{B}\left(\mathbf{B}^{\mathrm{T}}\mathbf{W}_0^{-1}\mathbf{B}\right)\mathbf{B}^{\mathrm{T}}\mathbf{W}_0^{-1}$

A reduced covariance matrix with inverse $\mathbf{W}_0^{-1}\mathbf{P}$ can be therefore identified similar to the simple case of Sec. 2.2. (Eq. (1)). However, this covariance matrix is singular and a regularization has to be carried out by treating the augmented part of the model as a process noise.

The corresponding matrix $\mathbf{W}^* = \mathbf{W}_0 + \mathbf{BW}_{\mathbf{y}}\mathbf{B}^{\mathrm{T}}$ is defined where $\mathbf{W}_{\mathbf{y}}$ is the apriori covariance matrix for the parameter $\mathbf{z}$ with no apriori information on the variance of the process noise. The inverse can be expressed using the matrix inversion lemma as:

$$\mathbf{W}^{*-1} = \mathbf{W}_0^{-1} - \mathbf{W}_0^{-1}\mathbf{B}\left(\mathbf{B}^{\mathrm{T}}\mathbf{W}_0^{-1}\mathbf{B} + \mathbf{W}_{\mathbf{y}}^{-1}\right)^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{W}_0^{-1}.$$

Provided that $\mathbf{W}_{\mathrm{y}}$ is sufficiently large, the functional and stochastic approaches are equivalent. This way, the data strength is said to be preserved which may not be the case by estimating an additional parameter. When using fully populated matrices accounting for correlations between observations in Least-Squares adjustment without augmented functional model, this equivalence is implicitly used. As a consequence, exemplarily, no tropospheric parameters are additionally estimated since their impact is considered to be included in the stochastic model of the LS adjustment.

3. Stochastic models for GNSS: state of the art

As described in chapter 2, the knowledge of the stochastic model by means of the covariance matrix of the residuals is of primary importance for a trustworthy and efficient Least-Squares solution. In order to identify where the stochastic model for GNSS positioning can be improved, a review of the already proposed methods and their main results is carried out. It is deliberately chosen not to give the exact values found by the corresponding authors (for instance with respect to the standard deviations or rms) but only qualitative indications. Indeed, as results were found under different conditions, data quality, positioning strategies or baseline lengths, a direct comparison allows hardly for drawing general conclusions on the superiority of one model over the others. This holds particularly true for very specific functions. Moreover, elements such as the aposteriori variance factor or the precision which could have been used to analyse the trustworthiness of the solution, are often missing.

Since the focus of this study is on the stochastic model of GPS observations, the functional model for GPS is not further described. The corresponding derivations can be found exemplarily in Hofmann-Wellenhof et al. (2001) or Misra and Enge (2012). Additionally, Kermarrec and Schön (2016) made a concrete comparison between design matrices corresponding to different positioning strategies.

The following description of the stochastic models for GPS observations is divided into a first part related to the variance and a second one to the covariance models.

3.1. Variance models

Two main families were identified to describe the variance of GPS observations: a mapping function approach and empirical functions. The identity variance model is here not considered independently.

3.1.1. Mapping function model

Bischoff et al. (2005) proved heteroscedasticity of GPS carrier phase observations by checking the hypothesis of a non-variable variance thanks to statistical test. It was shown that the variance is strongly dependent on the elevation of the satellites, thus confirming what was already assumed in the earlier time of GPS adjustments. Indeed, if one considers the path of a GPS signal through the entire atmosphere to the receiver, assuming equal weights to all satellites seems intuitively overoptimistic. Goad (1987) and Langley (1997) showed that systematic effects caused by multipath, orbit errors or the atmosphere affect each satellite differently. By inspection of the residuals of GPS observations versus elevation, low elevation satellites were identified as more influenced by such effects than satellites having a higher elevation. Thus, the most natural model was to propose a sinus-like function for the variance, i.e.

$$\sigma_i^2 = 1/\sin^2(El_i) \qquad (11)$$

where El_i is the elevation of satellite *i*, with $El_i > El_{\min}$ and σ_i^2 the corresponding variance (Vermeer 1997, Rothacher et al. 1998). In the following, this variance model is called the ELEV model. It can be seen as a mapping or projection of the satellite path to the vertical, i.e. an obliquity factor accounting for the increased effective path lengths through the atmosphere as the elevation angle decreases. Being simple, this variance does not necessitate additional computation. It is widely used and implemented in software packages, e.g. in the Bernese

GNSS software (Dach et al. 2015). Alternatively the models $\sigma_i^2 = 1/\sin(El_i)$ or $\sigma_i^2 = 1/(El_i)^2$ were sometimes tested (da Silva et al. 2009, Wang et al. 1998).

3.1.2. Data-based models

Alternatively to the mapping function approach, empirical models were proposed based on the analysis of the residuals. Two main families can be distinguished: the exponential model and quality indicator models.

3.1.2.1. Exponential model

Trying to fit more accurately the variance versus elevation, an exponential model was in some studies preferred to the ELEV model as presented in Euler and Goad (1991), Gerdan (1995), Han (1997), Jin (1996), Wang (1999), Barnes et al. (1998). Moreover, Li et al. (2014) or Amiri-Simkooei et al. (2007, 2009, 2016) used this model as a starting point for the LS-VCE estimation. It is stated that the variance can be fitted with high confidence as:

$$\sigma_i = \left[a_1 + a_2 e^{-El_i/\theta_0} \right] \tag{12}$$

where θ_0 is a reference elevation angle valid for all satellites which has to be estimated together with a_1 and a_2 using a non-linear Least-Squares procedure. This model can be considered as a sum of a white and an elevation dependent noise. Different values of these three parameters are obtained depending on the receiver, the frequency or observation types and the configuration (long or short baselines for instance). The variance is often fitted with the Least-Squares method for the particular case of a zero baseline and the parameters found are used per extension for longer baselines (Li et al. 2016).

This variance model with $\theta_0 = 10°$ was shown in Wang et al. (1998) to give a slightly lower determinant of the ambiguity covariance matrix with respect to the ELEV model. This measure is used to estimate the apriori performance of the ambiguity resolution (Teunissen 1997). A better baseline repeatability was additionally obtained. Similar results were found by Özlüdemir (2004) for the standard deviation of the baseline solution compared with the identity model (equal weights). However, a concrete comparison between the two models with respect to the precision and the variance of unit weight is missing in these studies. Li et al. (2016) investigated the influence of the three models on the overall model test and outlier detection test and showed that the corresponding results with the exponential model for a zero-baseline scenario with single differenced observations were smaller than with the ELEV or equal weights model. The extent to which the value of the parameter θ_0 may influence the solution and its precision was not further investigated. The same comment holds true for Luo et al. (2014) who used a value of 40° derived from mean SNR-weighting values.

3.1.2.2. C/N0 or SNR0 based models

The C/N0 and SNR0 models are similarly based on a quality indicator recovered from the GPS receiver. The carrier to noise power density ratio C/N0 (Langley 1997) is defined as the ratio of the power level of the signal carrier to the noise power in a 1Hz bandwidth. The Signal to Noise Ratio (SNR) is the ratio of the signal power and noise power in a given bandwidth expressed as logarithm (Joseph 2010), SNR0 being equivalent to C/N0 (i.e. normalized to a 1Hz bandwidth). SNR0 or C/N0 measurements are e.g. depending on the antenna gain of the transmitting satellite, polarisation errors, variations in atmospheric propagation and receiver antenna gain patterns, or antenna cable (Misra and Enge 2012).

- C/N0 weighting

The first authors to use the C/N0 to weight GPS observations were Talbot (1988), Langley (1997) or Braasch and Van Dierendonck (1999). They approximated the variance of the carrier phase L1 or L2 as:

$$\sigma_i = \frac{\lambda}{2\pi}\sqrt{\frac{B}{cno_i}} \tag{13}$$

where B is the carrier tracking loop bandwidth in Hz, λ the wavelength of the carrier phase and $cno_i = 10^{CN0_i/10}$ for $CN0_i$ expressed in dB-Hz.

Collin and Langley (1997) compared the weighting scheme ELEV and a SNR-based variance model and deduced that they were consistent with each other with respect to the baseline repeatability and ambiguity resolution.

This C/N0 weighting was extended by Brunner et al. (1999) and Hartinger and Brunner (1999) to develop the sigma-ε and sigma-Δ models. The variance is expressed more accurately than in Eq. (13) as

$$\sigma_i^2 = V + C10^{-CN0/10} \tag{14}$$

where V, C are receiver/antenna/frequency depending model parameters in $\left[m^2\right], \left[m^2 Hz\right]$ respectively. A comparison of the ELEV and C/N0 based models is given e.g. in Luo (2012) where it is shown that low elevation satellites are more accurately weighted with C/N0 based improved weightings.

The sigma-Δ model (Brunner et al. 1999, Wieser and Brunner 2000) is similar but accounts for signal distortion which may decrease the signal to noise ratio. This approach is said to be more powerful for kinematic applications. However, the main problem of the C/N0 models as pointed out in Wieser and Brunner (2000) or Satirapod and Wang (2000) yields in the biased solution due to too high weights introduced for observations under diffraction. It leads to a weakening of the geometry consecutive to low weights for observations when the signal path is close to the extremity of hindrances. Using the Danish method and information from the residuals (Krarup et al. 1980), a more robust estimation was possible, which is particularly meaningful for kinematic applications.

Özlüdemir (2004) compared the standard deviations of the coordinates with the sigma-ε and the equal weighting model and highlighted the better repeatability of the solutions. However, ELEV and C/N0 weighting gave similar results. This is a common result of many studies on the stochastic model for GNSS observations when comparing elevation dependent variance functions and is mainly due to the unbiasedness of the LS estimator when the Gauss-Markov assumptions are fulfilled.

- A Fuzzy based model

Trying to overcome the weakness of the C/N0 weighting, Wieser and Brunner (2002) developed the sigma-F model using the Fuzzy strategy. The variance model is a product of an apriori variance and two variance inflations depending both on the signal quality via an adapted version of the C/N0 values comparable with the sigma-Δ model and on the predicted residuals. The input variables were chosen as the result of the outlier detection test and quantities derived from the C/N0. This iterative procedure was shown to be powerful in the case of kinematic processing when signal distortions affected only a small part of the observations. Thus, fuzzy logic is an interesting and intuitive alternative way to handle the different cases that may arise depending on the geometry and the C/N0 values. However, Wieser and Brunner (2002) concluded that such computationally demanding strategies are not preferable to other methods in usual positioning cases.

- SNR0 weighting

In the SNR0 based model, it is stated that the ratios of the signal and noise power of the modulated signal at the correlator output (i.e. SNR0) and at the receiving antenna (C/N0) are amplified by approximately the same factor (Luo 2002) and can be approximated by each other.

One of the first empirical based SNR weightings was developed by Mayer (2006). Luo (2012) improved this model by using the ratio between the minimum SNR0 value and the corresponding maximum where one additional parameter avoided singularity. The corresponding variance depends on the antenna-receiver combination and can be used for L1 or L2 observations. It reduces the downweighting of low elevation measurements and takes different factors into account such as multipath, site specific effects or meteorological conditions. A particular example with different baseline lengths showed an improvement by up to 10%-20% of Narrow Lane ambiguities success rate with respect to the ELEV model for baseline lengths of 50-200km with a cutoff angle of 3°. However, this weighting necessitates post-processing and remains computationally demanding. Therefore, Luo et al. (2014) proposed a simpler exponential-based weighting where the parameter θ_0 of Eq. (12) was empirically averaged to 40°. This simplification is comprehensible as the signal power values strongly depend on the elevation, the log function being nothing more than an exponential function.

3.1.2.3. Alternative weighting strategies / GLONASS and Beidou signals

Other proposals were made by for instance adding a $1/\sin^2(El_i)$ factor to adapt Eq. (14) in order to account for multipath in an urban canyon (Tay and Marais 2013). To incorporate scintillation effects due to the ionosphere, an alternative tracking error variance at the output of the receiver Phase Look Lock for the GPS L1 carrier phase was developed leading to a C/N0 based weighting (da Silva et al. 2009, Aquino et al. 2009).

In the case of GLONASS observations, Gaglione et al. (2011) fitted the variance of the residuals from pseudorange adjustment by an uncommon 4th order polynomial. However, the results were not employed in a Least-Squares adjustment. Using carrier phases of GPS and GLONASS single-frequency observations, Wang (2000) postfitted filtered residuals using a Kalman filter with an exponential based preset covariance matrix. An improvement of the ambiguity resolution performance was shown for a real-time kinematic example.

Li (2016) first studied the stochastic properties of Beidou residuals, highlighting strong differences between the variances of MEO or GEO satellites. Using the strategy of Parkinson and Spilker (1996), a data dependent sinus based variance model was used: $\sigma_i^2 = \left(a / \left(\sin(El_i) + b \right) \right)^2$, a, b having to be estimated.

3.1.3. Summary of the proposed variance models

Table 1 presents a summary of the proposed variance models as described in the previous section.

Name Model	Variance model	Particularity	Main references
Identity	1	Assuming homoscedasticity	OLSE
Mapping function based models			
	$1/\left(\sin(El)^2\right)$	Commonly implemented in processing software (Bernese). Strong downweighting of observations from low elevation satellites. Not defined for 0° elevation.	Parkinson and Spilker (1996) Vermeer (1997) Rothacher et al. (1998)
	$1/\sin(El)$		Wang et al. (1998) Da Silva et al. (2009)
	$1/El$		Wang et al. (1998)
Exponential	$\left(a_1 + a_2 e^{-El_i/\theta_0}\right)^2$	The exponential factor is determined empirically based on zero baseline tests. May weight low elevation satellites too strongly.	Euler and Goad (1991)
C/N0	$\left(\frac{\lambda}{2\pi}\sqrt{\frac{B}{cno_i}}\right)^2$ $\left(V + C10^{-CN0/10}\right)^2$	Sigma-ε, sigma-Δ model. antenna-receiver characteristic and site-specific effects are taken into account. Main weakness: C/N0 values have to be extracted from RINEX files, possible weakening of the geometry. For strong signals.	Langley (1997) Brunner et al. (1999) Wieser and Brunner (2000) Hartinger and Brunner (1999)
SNR0	Similar function as C/N0 models or based on min/max values	Less dependent on how well the unknown loop bandwidth is specified. Not only for signals well above the tracking threshold of the receiver (strong signals).	Mayer (1996) Luo (2012)
Fuzzy	Input variables: outlier detection and C/N0 based values	Sigma-F model for kinematic applications and punctual obstructions	Wieser and Brunner (2001)
Iono model	Depending on phase or code processing	In case of strong ionospheric scintillations, based on C/N0 value	Aquino et al. (2009)
sinus model	$\left(a/\left(\sin(El_i)+b\right)\right)^2$		Li (2016) for Beidou, Parkinson and Spilker (1996)

Tab.1 Summary of the proposed variance models for GNSS observations

3.1.4. Variance models for code observations, PPP

Besides the traditional variance models, different strategies were additionally proposed to weight code-phase or PPP phase observations to account for their specificities.

3.1.4.1. Variance of code-phase observations

Using zero-baseline experiments, the code-phase observations variances were computed in Tiberius and Kenselaar (2003) or Amiri-Simkooei et al. (2009) using LS-VCE methods. These studies mainly agree on an elevation dependency, whereas the extent to which other weightings for code observations should be used was not directly addressed. Remembering that C/N0 values were initially developed for code phase observations, this may be the most appropriate choice.

3.1.4.2. Variance of PPP phase observations

Up to now, the carrier phase for a PPP (Precise Point Positioning) positioning strategy was mostly weighted either with an identity model or an elevation dependent ELEV model (Heßelbarth and Wanninger 2008). Satirapod and Luansang (2008) used MINQUE for static data processing with PPP and showed an improvement at the submm level after 30 minutes of observations compared with the ELEV or identity model. However, up to now and due to both the computational burden and associated processing difficulties, no attempt to further improve the variance model has been made. Indeed, code and phase weightings should be modelled independently.

3.1.5. Discussion

All these studies comparing or using different weighting strategies remain dependent on the observations which may be affected by different atmospheric conditions or multipath, by the processing strategies, by the receivers used, or the length of the baselines. As a result, without bias or error analysis, it seems difficult to draw general conclusions on particular models. The main argument for using empirical weightings is that it offers a more accurate approximation of the variance for low elevation satellites which is an important criterion for estimating the Up component and tropospheric delays (Misra and Enge 2012). However, the authors focused mainly on the standard deviation of the solution or on the ambiguities resolution. Giving more weights to low elevation satellites or using different C/N0 weighting strategies may influence the aposteriori variance factor or the cofactor matrix of the estimates, i.e. the overall model test and the precision measures respectively. These quantities are usually not addressed although they represent a way to validate the trustworthiness of the solution by giving information about the efficiency of the Least-Squares solution. Thus, some empirical results are sometimes difficult to interpret. For example an improvement of the 3Drms by up to a few cm for long baselines is not always synonymous with a statistically correct solution; indeed sometimes the situation is just the opposite.

3.2. Covariance models

Besides the mathematical correlations (Hofmann-Wellenhof et al. 2001, Beutler et al. 1987) that are depending on the processing strategy (e.g. single or double differences), GPS observations are physically temporally correlated. The first contribution to this problem dates from the 90’s (Craymer et al. 1990) where the overestimation due to the use of a diagonal VCM was raised.

Two groups of correlations can be identified:

I. The temporal correlations, i.e. time dependent correlations between the observations of one satellite with itself or with another one. Although effects should be considered as spatio-temporal, purely temporal correlations are easier to access in time series analysis. They are mainly due to the path of the GPS signal through the atmosphere. In this family, we also include correlations due to multipath or due to the receiver itself, often called channel correlation (Tiberius and Kenselaar 2003).

II. The correlations between observation types, i.e. phase and code or between frequencies (Teunissen et al. 1998, Jonkman 1998, Bona 2000). Such kinds of correlations are particularly important when dealing with combinations of measurements.

Two strategies exist to account for correlations:

I. Using iterative algorithms where the correlations are taken into account, often with a predefined structure, e.g. via Kronecker products. This procedure may be computationally demanding.
II. Trying to model the temporal correlations apriori.

3.2.1. Apriori models for correlations

In this section the three main strategies to describe the correlations of GPS phase measurements are reviewed. They include the exponential model (El-Rabbany 1994), the turbulence based model (Schön and Brunner 2008) as well as the ARMA (Auto Regressive Moving Average) model.

3.2.1.1. Exponential model

This is likely the most popular model. It was derived empirically by El-Rabbany (1994) and used by Howind et al. (1999) or later Radovanovic (2002) and is often the starting point for the estimation of the covariance component in LS-VCE (Zangeneh-Nejad et al. 2015). The autocorrelation of the residuals is fitted empirically to a function corresponding to an AR(1) process (first order autoregressive model) and can be expressed as:

$$Corr(\tau) = ae^{-|\tau|/T} \tag{15}$$

where a is scaling factor, τ is the time increment and T the correlation time (i.e. the so-called 1/e point). The parameter T is empirically determined for different frequencies. El-Rabbany (1994) gave the values of 263s, 270s and 169s for L1, L2 and L3 (ionosphere-free combination) double differenced observations respectively with no dependency of the correlation time on the baseline length. He made use of modified batch Least-Squares to simplify the handling of fully populated VCM.

The main advantage of this model is its relative simplicity, i.e. an identical correlations structure is assumed for all satellites. However, the reality is slightly different as low elevation satellites may lead to higher correlation level due to e.g. the increased path length through the atmosphere (Amiri-Simkooei et al. 2009, 2013).

The second advantage of this modelling comes from the inverse of the resulting fully populated VCM which can be mathematically expressed (Rao and Toutenbourg 1999, Kermarrec and Schön 2016), potentially reducing the computational burden due to the inversion in the Least-Squares method. This possibility was for instance used by Odijk and Teunissen (2007) to investigate the influence of taking correlations into account in the ADOP (Ambiguity Dilution factor) closed formula.

However, the main weakness of the so-called exponential correlation model results from the equal variances that are given for all satellites as it does not account for heteroscedasticity.

Other studies used Eq. (15) to model correlations due to multipath by fitting the residuals after having removed the effect of noise. Exemplarily, Radovanovic (2001), Radovanovic et al. (2000), Schwieger (2007) added different covariance matrices to include correlations due to thermal noise using the C/N0 weighting or receiver noise.

3.2.1.2. Turbulence-based model

Schön and Brunner (2008) developed a model to account for correlations of GNSS phase measurements due to tropospheric refractivities in the atmosphere. Built on results from the turbulence theory, this model is physically relevant. The covariance function between the phase φ observed at stations A and B at t and $t+\tau$ reads:

$$Cov\left(\varphi_A^i(t),\ \varphi_B^j(t+\tau)\right) = \frac{2^{1/3}}{3\Gamma\left(\frac{2}{3}\right)} \frac{\kappa_0^{-2/3}}{\sin El_i^A \sin El_j^B} C_n^2 \int_0^H \int_0^H \left(\kappa_0 d\right)^{1/3} K_{1/3}\left(\kappa_0 d\right) dz_1 dz_2 \tag{16}$$

where *i* and *j* are two satellites, Γ the gamma function, d the separation distance between two GPS rays. The turbulence parameters H, κ_0, C_n^2 (tropospheric height, outer scale length and structure constant respectively) involved are described in detail in Schön and Kermarrec (2015).

This model is based on a double integration where the variance is computed as a limit. Allowing for the consideration of spatio-temporal variability, the concept of separation distance is the major finding of this model with respect to the empirical model of El-Rabbany (1999). The covariance between phase observations from different satellites can be computed. Based on the ELEV variance model, it accounts moreover for heteroscedasticity. However, some numerical instabilities may occur when evaluating the double integral. A simplification of this model was carried out and presented in Kermarrec and Schön (2014) and is briefly outlined in Sec. 4.2. It is based on the powerful Matérn family presented in chapter 2.

3.2.1.3. Modelling via ARMA processes

Luo (2012) and Luo et al. (2011) showed that filtered and studentized residuals from short term relative positioning and long term PPP can be expressed as ARMA processes where the ARMA orders may depend on factors such as atmospheric conditions or multipath. Fully populated VCM could be therefore deduced thanks to the Yule Walker equations. The similarity between the ARMA modelling and the Matérn covariance family is addressed in Kermarrec and Schön (2017a).

3.2.2. Recursive models (Delft school)

The second possibility to account for correlations in Least-Squares adjustment is to use a recursive procedure based on residuals or VCE as presented in Sec. 2.4.1. Although some simplifications have been proposed (Satirapod et al. 2002, Amiri-Simkooei et al. 2016, Xu et al. 2007), these methods remain computationally demanding and do not necessarily lead to positive definite matrices.

The MINQUE model was mostly used e.g. by Wang et al. (1998, 2002), Satirapod et al. (2002,2003) and is valid for short observation periods and baselines, i.e. the temporal correlation coefficients and the variance are assumed to remain unchanged (Satirapod et al. 2003). A simplification procedure has been proposed by Satirapod (2001) based on data segmentation. Tiberius and Kenselaar (2000) investigated additionally the correlations between different observation types thanks to this strategy.

The LS-VCE procedure was popularized by Teunissen and Amiri-Simkooei (2008), Amiri-Simkooei et al. (2009, 2013) or used by Li et al. (2011). The variances of the observations are modelled depending on the observation types or frequencies. Moreover, cross correlations are taken into account and a satellite elevation dependence based on the exponential weighting model is assumed as a starting point of the procedure. Consecutively, the VCM of the observations is expressed thanks to a Kronecker product which reads $\hat{\mathbf{W}} = \hat{\mathbf{W}}_{\mathbf{C}} \otimes \hat{\mathbf{W}}_{\mathbf{T}} \otimes \hat{\mathbf{W}}_{\mathbf{E}}$, where $\hat{\mathbf{W}}_{\mathbf{C}}, \hat{\mathbf{W}}_{\mathbf{T}}, \hat{\mathbf{W}}_{\mathbf{E}}$ are the subsequent covariance matrices for cross correlation, temporal correlations and elevation dependent effects, respectively.

The main weakness of these widely used iterative procedures (Wang et al. 1998, 2002, Li et al. 2015, Amiri-Simkooei et al. 2016) is that the temporal correlations are not accounting for non-stationarity, i.e. the matrices $\hat{\mathbf{W}}_{\mathbf{T}}$ are of Toeplitz type and are the same for all satellites. Moreover, the structure of time correlation and satellite elevation dependence is identical for code and phase observations which may not be the case in reality (Braasch and VanDierenonck 1999). The influence of the starting model and the extent to which it influences the final result was never addressed.

3.2.3. Summary of the proposed covariance functions

In Table 2, the proposed covariance functions are shortly described based on the previous classification.

Model type	Model name	Particularity	References (exemplarily)
Empirical models			
	Exponential model	Exponential covariance function with empirically determined correlation time	El-Rabbany (1994)
	ARMA	Empirically determined ARMA orders	Luo (2012)
Physically based models			
	Turbulence based model, Schön and Brunner model	Covariance function developed principally for modelling correlation due to tropospheric refractivities	Schön and Brunner (2008)
	Multipath model	Exponential covariance function to model the correlation due to multipath	Radovanovic (2001)
Iterative model			
	LS-VCE MINQUE	Models are necessary for simplification and to avoid negativeness of the VCM. No spatial temporal dependency. Computationally demanding.	Teunissen and Amiri-Simkooei (2008) Wang et al. (2002) Satirapod et al. (2002) Li et al. (2015)
Elementary error model			
		Sum of different elementary errors functions	Radovanovic (2001), Radovanovic et al. (2000), Schwieger (2007)

Tab.2 Summary of the proposed covariance models

3.3. State of the art: impact of the stochastic model on LS results

As mentioned previously, no concrete and precise values of standard deviation or rms were given in the comparison between existing models. Indeed, different data sets were used from different years with different quality and processing strategies applied making comparisons difficult. Moreover, the results of test statistics such as the overall model test which would have allowed for the identification of potential biases are rarely given. Nevertheless, some general conclusions on the impact of the stochastic model can be drawn from these previous studies which are summarized below.

- Baseline component estimation

As long as the stochastic models have an elevation-dependent weighting, the effects on positioning were shown to be similar for various stochastic models (Wang et al. 1998, Özlüdemir 2014). Consecutively, many authors compared the results provided by “their” empirical improved stochastic model with the one obtained with the identity VCM (Amiri-Simkooei et al. 2016) to obtain stronger statements. Due to the unbiasedness of the LS estimator, they should be critically considered.

Radovanovic (2001) showed that improving the stochastic model with a multipath-based model only leads to a difference at the mm-submm level for the baseline length compared with a diagonal variance model. Luo et al. (2014) found similar results for the baseline component estimation with the improved SNR based weighting.

- Overestimation of the precision

The overestimation of the precision is the main critical point addressed to a diagonal VCM only accounting for heteroscedasticity. Craymer et al. (1990) proposed to overcome this overestimation by using an apriori scale factor (see also Misra and Enge 2012). Indeed, when combining measurements of different sources (i.e. GPS and terrestrial), the wrongly overestimated weighting of GPS observations may lead to problems when forming control networks (Miller et al. 2012). Han and Rizos (1995) defined a data based scale factor for the VCM of the estimates which may be similar to the one found by inverting AR(1) VCM (Rao and Toutenburg 1999). Similarly, Jansson and Persson (2013) introduced the concept of effective number of observations to account for correlations. Using an iterative method for Beidou measurements, Li (2016) shows that the precision was improved and more realistic, i.e. it was possible to obtain a decrease of the overestimation using a more correct stochastic model.

- Effect on ambiguity resolution

O'Keefe et al. (2007) investigated the effect of temporal correlations with respect to the trustworthiness of the ambiguity resolution. Similar to the study of Odijk and Teunissen (2007) on ADOP, it was pointed out that neglecting correlations impacts the test statistics for ambiguity validation. Li and Wang (2012) made similar conclusions. The effect of an erroneous float ambiguity vector was proven by means of simulations in Joosten and Teunissen (2001), i.e. a biased ambiguity success rate. Amiri-Simkooei et al. (2016) showed that LS-VCE methods led to an improvement of the empirical success rate. However, as comparisons with the identity model (i.e. assuming homoscedasticity of the variance) were done although heteroscedasticity of GNSS observations was proven (Bischoff et al. 2005), it remains difficult from these contributions to make general conclusions with regard to other elevation dependent weightings.

Using the empirical sinus model without taking correlation into account for the particular case of Beidou observations, Li (2016) showed an improvement of the ambiguity resolution efficiency with both larger success rate and smaller false alarm. Similarly, the SNR0 weighting (Luo 2012, Luo et al. 2014) led to significant improvement in the ambiguity resolution compared with the simple ELEV model when considering a lower cutoff of the observations (3°) as well as in case of multipath. However, the focus is more often on reducing such error effects (Verhagen and Odijk 2007) than studying the influence of the stochastic model on ambiguity resolution. As a consequence, the general effects of improved stochastic models on discriminant tests used for fixing the ambiguities to integers were -to the author's knowledge- not addressed yet.

- Effect on outlier detection and overall model test

Using the exponential variance model in which parameters were determined for a zero-baseline, Li et al. (2016) compared the results given for the overall model test and outlier detection test with respect to the ELEV and identity variance model. A decrease of the false alarm rate was shown with the exponential variance model, however without studying a possible underestimation or bias of the quantity and mentioning the value of the apriori variance taken into consideration. The impact of correlations was moreover not addressed.

- Effect on tropospheric parameter estimates

Luo (2012) and Luo et al. (2014) studied the impact of the SNR0 weighting on the tropospheric parameter estimation. Differences with respect to the commonly used ELEV model in terms of standard deviations of up to a few cm were found. Moreover, a sensitivity to meteorological conditions could be shown.

3.4. Issues and objectives

In the previous section, the impact of the stochastic model on the Least-Squares solution as described in the literature was summarized, highlighting its importance to get a trustworthy Least-Squares positioning as well as an improved ambiguity resolution. However, as different observations or positioning strategies were used, an evaluation of the corresponding weightings and their performances is not straightforward.

Focusing on modelling instead of iterative procedures, some important open questions on the effects of the VCM of GNSS observations remain thus to be treated. Their answers should lead to a better understanding of the impact of improving the stochastic model by taking correlation into account.

- A need for a simple model of the correlation from GNSS (phase) observations

Besides a particular function for tropospheric correlations (Schön and Brunner 2008), only one model (El-Rabbany 1994) was suggested to account for correlations in LS adjustment. This proposal is empirical and states homoscedasticity of the variance. Moreover, it is based on fitting of the autocorrelation function of the residuals which was identified in Stein (1999) as being a sub-optimal strategy. Thus, correlations are neglected due to the lack of a general, easy to use and more realistic model, leading to a less trustworthy Least-Squares solution.

- Impact of misspecification of the correlation structure up to neglecting correlations

Omitting correlations impacts the test statistics, the ambiguity validation tests, the efficiency of the solution, as well as to some extent the estimates in particular cases (e.g. short batches of observations). However, a quantification of the errors on these Least-Squares quantities when the correlation structure is misspecified is missing. An important part concerns the biased float ambiguity solution and how it is linked to the validation procedure if correlations are present but neglected. Studying these impacts will lead to a characterization of the most adequate model to choose depending on the scenario (batch length, correlation structure, positioning scenario), i.e. answering the questions why and when neglecting correlations is adequate or not.

- Reducing the fully populated VCM into a diagonal VCM and comparison with other diagonal strategies

As an improved stochastic model is a necessary condition for reaching the minimum variance of the Least-Squares solution, the last step is to further simplify the proposed correlation model to decrease the computational burden due to matrix inversion. This can be done by compacting the fully populated VCM into a diagonal matrix. It allows to build a bridge between the more conventional variance models described previously and the proposed correlation function, deepening the understanding of how weightings work.

4. Improving the stochastic model of GNSS phase observations

In this chapter, the first two questions raised in Sec. 3.4 are answered. To this end, the methodology to analyse the impact of the stochastic model and its misspecification for both the simulations and the real case study is briefly summarized. In Sec. 4.2, the new model for correlations is presented, followed by the main results of integrating the corresponding fully populated VCM in a LS adjustment taken from the following publications:

- Simplification of a model for correlations due to tropospheric refraction:

  Kermarrec G, Schön S (2014) On the Matérn covariance family: a proposal for modelling temporal correlations based on turbulence theory. Journal of Geodesy 88(11):1061-1079

- Proposal for modelling GNSS phase correlations based on the Matérn covariance function, sensitivity analysis of the parameters, simulations and case study:

  Kermarrec G, Schön S (2017a) Apriori fully populated covariance matrices in Least-Squares adjustment – case study: GPS relative positioning. Journal of Geodesy 91(5):465-484

- Possible physical explanation for the fixing of the parameters of the proposed covariance function:

  Kermarrec G, Schön S, Kreinovich V (2017b) Possible explanation of empirical values of the Matérn smoothness parameter for the temporal covariance of GPS measurements. Applied Mathematical Science 11(35):1733-1737

- Global analysis of the effects of fully populated VCM on the LS results:

  Kermarrec G, Schön S (2017c) On modelling GPS phase correlations: a parametric model. Acta Geophysica et Geodaetica https://doi.org/10.1007/s40328-017-0209-

4.1. Methodology

In chapter 2, the influence of a misspecification of the VCM on various Least-Squares quantities was theoretically presented (Eqs. (5)-(8)) and summarized in Fig. 1. In real cases however, the true VCM is unknown and can only be estimated or generated. As a consequence, simulations are used to study the sensitivity of the LS results with respect to the stochastic model. The corresponding outcome allows to identify the needed accuracy for the determination of the model parameters of the covariance function. Additionally, a deeper understanding of the results from data analysis is gained.

4.1.1. Sensitivity analysis of the stochastic model: simulations

The methodology used for the simulations follows the principle presented in Fig. 1. The impact of the correlations on the loss of efficiency, the test statistics, the estimates including the float or integer ambiguities are thus covered. Under a particular but general enough satellite geometry, the design matrix corresponding to relative positioning is entirely known. In a first step, the VCM $\mathbf{W_0}$ of simulated observations is computed. Approximated VCM $\hat{\mathbf{W}}$ are built in a second step related to a covariance function whose parameters can be varied around the values of reference used to compute $\mathbf{W_0}$. As a consequence, a sensitivity analysis can be easily carried out allowing a quantification of the impact of the FGLSE on the proposed Least-Squares quantities.

4.1.2. Impact of misspecification of the stochastic model: real case study

4.1.2.1. Methodology

In real cases, the true VCM is unknown and estimated. Thus, strategies have to be found to identify if and when the Least-Squares solution is still efficient and trustworthy, i.e. its precision is realistic and coherent with the estimates. In this thesis, two main quantities were considered for data analysis:

- aposteriori variance factor

Thanks to the results of simulations, a parallel with the behaviour of the aposteriori variance factor and the relative efficiency of the LS estimator can be drawn when the stochastic model is misspecified. Moreover, the bias-behaviour is similar to other quantities such as for instance the one used in the outlier detection test (Teunissen 2000, Li et al. 2016).

Thus, through the publications associated with this dissertation, the aposteriori variance factor has retained attention to decide whether a chosen stochastic model is adequate or not. This factor was chosen because of its accessibility and popularity in the GPS community. Under the general term "better test statistics" found in the following sections, it is understood that the value of the aposteriori variance factor is corresponding to the expected apriori value. Indeed, improving the stochastic model by correctly taking correlations into account leads to less biased significance tests. In this thesis, the overall model test was employed, i.e. $\frac{\hat{\sigma}_{\hat{\mathbf{w}}}^{2}}{\sigma_{0\max}^{2}} > F_p\left(n-u,\infty,0\right)$ where $F_p\left(n-u,\infty,0\right)$ is a p-quantile of the central F-distribution having *n-u* and ∞ degrees of freedom. In addition, it is tested if $\frac{\hat{\sigma}_{\hat{\mathbf{w}}}^{2}}{\sigma_{0\min}^{2}} < F_t\left(n-u,\infty,0\right)$ to avoid underestimation. 95% and 5% for *p* and *t* were chosen respectively. Except in particular cases specially mentioned, normal distribution of the residuals is assumed and the overall model test can be used based on the F-distribution. The range of acceptable values of the apriori value σ_0 for relative positioning with double differences was taken between $\sigma_{0\min} = 2mm$ and $\sigma_{0\max} = 4mm$, a flexibility being allowed depending on the data quality.

- 3Drms

Processing Observed Minus Computed (OMC) observations, the 3Drms of the solution is defined as $3Drms = \sum_{i=1}^{m} \sqrt{\frac{trace\left(\left(\hat{\mathbf{x}}-\mathbf{x}_0\right)^{\mathbf{T}}\left(\hat{\mathbf{x}}-\mathbf{x}_0\right)\right)}{3}}$ where *m* is the number of batches of observations and $\mathbf{x}_0$ is the **0** vector when the coordinates are exactly known in advance. As a consequence and in ideal cases, the 3Drms is expected to be close to 0 as soon as *m* is large enough. A lower 3Drms with a given stochastic model is called in the following an "improved 3Drms solution." The impact of the VCM on the North, East or Up components is not addressed independently in this work, i.e. a global indicator was chosen to test the impact of wrongly estimated models.

Both the unbiasedness of the aposteriori variance factor detected on the basis of the overall model test and a low 3Drms are wished in real data analysis.

4.1.2.2. Positioning scenario and data analysis

Through all publications, a relative positioning scenario with double differences of observations from short and long baselines was considered. For consistency, the same EPN (European Network) stations were used, i.e. principally a 66 km long baseline between the stations KRAW and ZYWI in Poland, chosen arbitrarily. The data rate was 30s and an elevation cut-off angle of 3° was chosen. The randomly selected GPS day was DOY220 of year 2015. The starting time was chosen to be GPS-SOD 6000s and was shown not to impact the conclusions (Kermarrec and Schön 2017a, Appendix 2). All corresponding observations were not preprocessed in order to leave unmodelled effects such as troposphere, ionosphere or multipath correlating the measurements. The ambiguities were computed using the Lambda method of the Lambda toolbox of the Delft University (Verhagen and Li 2012). For the short baseline case, the stations ZIMM and ZIM2 in Switzerland were chosen as well as multipath contaminated observations collected in Hanover (lower Saxony, Germany). Additionally, 1Hz observations between the EPN stations KRAW and CTRM (medium baseline) and KRAW and KRA1 (short baseline) were computed. Although the corresponding results from these baselines were not published yet, the general conclusions on model misspecification are comparable with the one drawn from the baseline KRAW-ZYWI. Indeed, the focus of this work is on analysing the sensitivity of the LS solution to the model parameters of the proposed covariance function, i.e. on understanding the general impact of taking correlations into account. Deduced from the outcome of the simulations, the conclusions drawn from the results of a real data analysis thus remains in all cases similar with the one presented in Kermarrec and Schön (2017a) and Kermarrec and Schön (2017c). The specific values of expected improvements depend on the observations or strategies used and are here not of interest.

4.2. Modelling correlations

In this section, the new correlation function for GNSS phase observations is shortly presented, the physically derived model of Schön and Brunner (2008) for phase correlations due to turbulent tropospheric refractivity fluctuations being the starting point of the proposal. A simplification of this function allows to account more easily for correlations due to the troposphere in positioning adjustment whereas a generalization permits to model elevation-dependent correlations. The corresponding functions are described in:

- Kermarrec and Schön (2014) for the tropospheric covariance function and
- Kermarrec and Schön (2017a) for its extension.

In the following section, only the main components of the models are described. Based on this both flexible and physically relevant proposal, it is possible to study the impact of a wrongly estimated correlation structure on the Least-Squares results following the methodology presented in Sec. 4.1.

4.2.1. Simplification of the model for tropospheric fluctuations

Based on Wheelon (2001), the proposal of Schön and Brunner (2008) implies a double integration which was shown to be both computational intensive and numerically unstable, particularly for low elevation satellites or high sampling rates. As a consequence, a simplification was needed to allow an integration of the corresponding VCM in LS adjustment (Kermarrec and Schön 2014).

4.2.1.1. Modelling temporal correlation: one satellite with itself

From the concept of the turbulence theory, it is stated that correlations of phase measurements due to tropospheric refractivity fluctuations for GPS phase signals are mainly due to large and elongated eddies present in the -loosely called- free atmosphere. The turbulence at this attitude is anisotropic and the reorganization of the energy structures is less rapid. Using the von Karman spectrum adapted to account for anisotropy under the Taylor frozen hypothesis, the spectral density of the phase variations can be derived. It involves turbulence parameters such as the structure constant of the index of refractivity C_n^2, the outer scale length L_0, the tropospheric height H or the wind velocity u. Due to the Wiener-Khintchine theorem, the corresponding covariance function (Kermarrec and Schön 2014) describing the correlations between the observations of one satellite with themselves, can be obtained:

$$Cov\left(\varphi_A^i(t),\ \varphi_A^i(t+\tau)\right) = 0.7772\frac{k^2 H C_n^2 c \kappa_0^{-5/3}}{\sin\left(El_i(t)\right)\sin\left(El_i(t+\tau)\right)}\left(\frac{\kappa_0 u\tau}{a}\right)^{5/6} K_{5/6}\left(\frac{\kappa_0 u\tau}{a}\right) \qquad (17)$$

where $\kappa_0 = \frac{2\pi}{L_0}$ the wavenumber, k the electromagnetic wavenumber, a and c are scale parameters accounting for anisotropy. $El_i(t)$ is the elevation of the satellite *i* at time t. $K_{5/6}$ is the Bessel function with a smoothness parameter of $\nu = 5/6$ and a Matérn correlation time $1/\alpha$, $\alpha = \frac{\kappa_0 u}{a} \approx 0.008 s^{-1}$ in the free atmosphere.

The positive definiteness of the function is given per definition. The variance is obtained as a limit and can be scaled so that it equals 1 for satellite at 90° elevation.

In Eq. (17), a Matérn covariance function (Sec. 2.4.2) can be recognized, multiplied by a mapping function or obliquity factor. Thus spatio-temporal dependency is taken into account, i.e. each satellite is weighted differently in the covariance function depending on its elevation. Correspondingly, the VCM are not of Toeplitz type as shown in Fig.2. The structure depends on the satellite elevation, i.e. the correlation length may increase (Fig. 2 middle) or decrease (Fig. 2 right). It is higher for low elevations (Fig. 2 left).

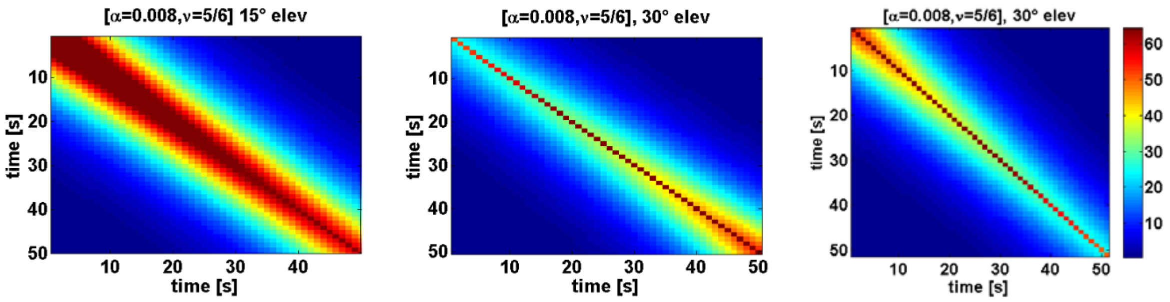


Fig.2 VCM built from the proposed model (Eq. (17)) for three satellites at different elevations (left: 15°, middle; 30° decreasing, right: 30° raising). The spatio-temporal dependency is highlighted. The batch has a length of 50 epochs (data rate 30s). The unit are [mm²]

4.2.1.2. Modelling temporal correlation: one satellite with other ones

In order to model the correlations between observations of two different satellites, eventually from different stations, the concept of the separation distance introduced by Schön and Brunner (2008) was further simplified to avoid a double integration. Based on the large eddies assumption, the distance between two rays is computed at a given height and not integrated. This height corresponds to the limit between the free atmosphere and the boundary layer, i.e. where elongated eddies are present but the structure constant of the refractivity index remains at high values (Wheelon 2001). This simplification can be compared with the definition of a ionospheric single layer model, where the electron content is considered to be concentrated on a spherical layer.

The outer scale length L_0 for the more energetic and stable large eddies that are impacting the correlations of GPS phase signals is fixed to 6000m. When the simplified separation distance d is smaller than L_0, case 1 of Eq. (18) is applied corresponding to Fig.3(left). However, although rarer at this altitude, eddies up to a length of 10km also coexist (Fig.3 middle). Thus, and in order not to set sharply the covariance function to 0 when the simplified separation distance is larger than 6000m, a physically plausible smooth transition was proposed which corresponds to case 2 in Eq. (18). Case 3 corresponds to Fig. 3(right) when tropospheric correlations between observations can be physically neglected.

$$
\begin{cases}
\text{Case 1} \\
\text{if d<6000m: } Cov\left(\varphi_A^i(t),\ \varphi_B^j(t+\tau)\right) = \ldots \\
\qquad \ldots 0.7772 \dfrac{k^2 H C_n^2 c}{\sin\left(El_i^A(t)\right)\sin\left(El_j^B(t+\tau)\right)} \left(\dfrac{2\pi}{L_0}\right)^{-5/3} \left(\dfrac{2\pi u \tau}{L_0 a}\right)^{5/6} K_{5/6}\left(\dfrac{2\pi u \tau}{L_0 a}\right) \\
\text{Case 2} \\
\text{if 6000<d<10000m: } Cov\left(\varphi_A^i(t),\ \varphi_B^j(t+\tau)\right) = \ldots \\
\qquad \ldots 0.7772 \dfrac{k^2 H C_n^2 c}{\sin\left(El_i^A(t)\right)\sin\left(El_j^B(t+\tau)\right)} \left(\dfrac{2\pi}{d_{t,H=1000m}}\right)^{-5/3} \left(\dfrac{2\pi u \tau}{d_{t,H=1000m} a}\right)^{5/6} K_{5/6}\left(\dfrac{2\pi u \tau}{d_{t,H=1000m} a}\right) \\
\text{Case 3} \\
\text{if d>10000m: } Cov\left(\varphi_A^i(t),\ \varphi_B^j(t+\tau)\right) = 0
\end{cases}
\tag{18}
$$

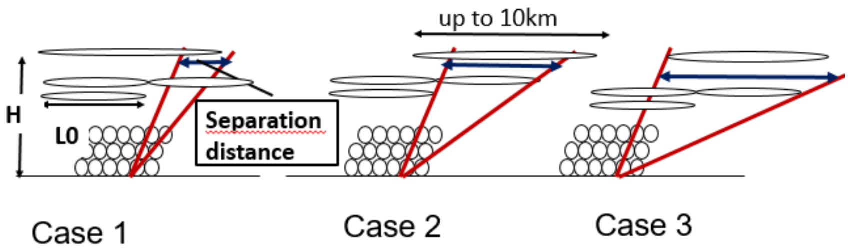


Fig.3 Simplified concept of separation distance used in the proposal to model phase correlations due to turbulent tropospheric refractivity fluctuations. The ellipses are a schematical representation of the eddies in the atmosphere. H is the tropospheric height. The different cases are corresponding to Eq. (18)

The model easily allows for the generation of fully populated covariance matrices that can be used to model the tropospheric zenith wet delay (Vennebusch et al. 2010) or as stochastic model in Least-Squares adjustment, for instance for VLBI observations (Halsig et al. 2016) allowing a better baseline repeatability and improved aposteriori variance factor. Using the derivation of Blewitt (1998) stating equivalence between an augmented functional model and a stochastic model (Sec. 2.5) and considering the influence of the troposphere as a process noise, this function can model an additional tropospheric parameter for short sessions of observations when the VCM is made sufficiently large for the equivalence to hold.

4.2.2. Generalization of the covariance function

The previous function necessitates the computation of a simplified separation distance and is only devoted to model phase correlations due to turbulent fluctuations of the index of refractivity. For this purpose, the smoothness and correlation parameters $\nu, 1/\alpha$ are fixed based on concepts from the turbulence theory. In order to account for other kinds of elevation-dependent correlation factors such as multipath, a more general function had to be proposed. The corresponding derivation is developed in Kermarrec and Schön (2017a).

4.2.2.1. Covariance function for elevation-dependent factors

Based on the analysis of the principal correlation factors (troposphere, ionosphere, multipath, eventually unmodelled phase center variations) and on the results of the previous studies presented in chapter 3, the main characteristics of this function for a wide acceptance were chosen to be:

- Elevation dependency, based on a commonly used and physically derived mapping function
- Flexible covariance function whose parameters can be estimated by Maximum Likelihood (see Sec. 2.4.2 or Handcock and Wallis 1994, Stein 1999).

An expected good acceptance for this model in the GNSS community was an important criterion in the choice of the function, i.e. it had to be historically based, easy to use, and understandable. Its flexibility should allow to carry out a sensitivity analysis in order to understand the effect of correlations on the trustworthiness of the LS solution as well as the needed accuracy in the determination of the corresponding parameters.

On purpose and for the sake of simplicity, no model selection from the observations was made (Williams and Rasmussen 2006). Moreover, an empirical fitting of LS residuals with a particular autocorrelation function was intentionally *not* investigated. First, a lack of generality would have been the consequence. Indeed, the smoothness is empirically often assumed to be ½ as in El-Rabbany (1994). The covariance function of the observations is thus not mean-squared differentiable at the origin which is a restrictive property particularly for observations from long baselines that are smoother than short baselines observations (Kermarrec et al. 2017b). Secondly, it is a non-accurate procedure for short observation sessions, being at the same time computationally demanding as a post processing is needed.

As a consequence, the covariance function of Eq. (17) was adapted to:

$$Cov\left(\varphi_A^i(t),\ \varphi_B^j(t+\tau)\right)=\frac{\rho\delta_{elev}}{\sin\left(El_i(t)\right)\sin\left(El_j(t+\tau)\right)}(\alpha\tau)^{\nu}K_{\nu}(\alpha\tau) \quad (19)$$

where El_i is the elevation of the satellite *i*. δ is a positive factor so that the value of the covariance function is scaled to have a variance of 1 for a satellite at 90° elevation. The Matérn parameters $[\alpha,\nu]$ (inverse of the correlation length and smoothness, respectively) are computed based on MLE (Sec. 2.4). In the following, the model is referred to as the FULLY model.

Eq. (19) allows to build fully populated covariance matrices to model elevation-dependent correlations between all satellites of one station as well as between two stations due to the parameter ρ. From physical considerations based on the vertical profile of microwave values for refractive-index structure constant C_n^2 (Wheelon 2010), ρ was fixed to 0.1 (Kermarrec and Schön 2017a). For the sake of simplicity as well as for numerical stability, it was assumed that correlations between stations could be neglected, the main impact in Least-Squares adjustment coming from the block diagonal VCM corresponding to correlations between observations from one satellite. The variance is obtained as a limit and is corresponding to the ELEV model (Schön and Brunner 2008)

4.2.2.2. Main characteristics of the covariance function

The main properties of the covariance model are outlined here briefly:

- *Link with AR model*. From Luo et al. (2012), it is known that the correlation structure of filtered GPS residuals can be described by means of AR or ARMA processes of different orders depending on the chosen data. Since Eq. (19) is derived from a rational spectral density, it has some similarities with the ARMA processes (Stein 1999, see also Kermarrec and Schön 2017c).
- *Spatio-temporal dependency*. For a GPS case, the covariance matrices built with Eq. (19) have no Toeplitz structure, i.e. the weighting for each satellites combination is individual. Figure 2 shows an example of the covariance matrices for satellites with different time-dependent elevations.
- *Heteroscedasticity* is taken into account due to the ELEV variance model
- *Different kinds of elevation dependent correlations.* The model summarizes all kinds of elevation dependent correlations thanks to a data-driven estimation of the parameters $[\alpha,\nu]$. It can thus be

adapted to take station- and frequency- dependent effects into account. The Matérn parameters are not fixed to a given value as for the turbulence model presented in Eq. (17) but are allowed to vary.

While the model has some advantages over other methods, its shortcomings should also be addressed:

- *The mapping function* which defines the variance model is physically derived (Wheelon 2001). However, some more complicated functions such as Niell or Marini mapping functions (Hoffmann-Wellenhof et al. 1999) that account for specific meteorological parameters may be more accurate.
- The model is well suited for *elevation dependent correlations* which can be considered as the main class of correlations of GPS phase measurements. Other types of correlation can be modelled additionally with for instance elementary covariance functions as described in Sec. 3.2.3. or without the elevation dependent weighting factor of Eq. (19).
- No additive noise (like e.g. receiver noise) is explicitly taken in consideration by adding a specific diagonal covariance matrix. However, due to the use of the Matérn covariance family and the non-orthogonality of the parameters (Gelfand et al. 2010), it is implicitly allowed to model noise as long as it is not dominant in the total variance (Kermarrec and Schön 2017a). In the case study and for the sake of numerical stability (Tikhonov 1995), a scaled identity VCM was added with a factor 0.1 for the identity VCM and 0.9 for the fully populated VCM.
- In a first approximation, the model is not developed for the modelling of code-phase correlations as needed for Precise Point Positioning (PPP) applications.

4.2.3. Estimating the Matérn parameters in real cases

While in simulations the parameters $[\alpha, \nu]$ can be varied freely, a methodology has to be proposed to fix the Matérn parameters in real cases. It is assumed that the estimates are deterministic (Caspary and Wichmann 1994) and thus the Observed Minus Computed (OMC) observations are here used to compute the model parameters. For cases where the true coordinates are unknown and trends remain in the observations, the same procedure as proposed can be followed based on detrended observations or residuals.

Due to the satellite-dependent weighting, only one parameter combination (smoothness and correlation length) is estimated for a given data set. This strategy represents a simplification in order to avoid computational burden and is based on the results of the sensitivity analysis of the parameters (Kermarrec and Schön 2017a). Depending on the processing strategies, three different cases were identified for fixing the Matérn parameters. The terms “long” or “short” sessions correspond to the batch length into which the observations time span is divided, independent of the data rate. Long sessions are batches longer than 100 epochs. This categorization is rough and subject to adaptation depending eventually on the results of test statistics. The same holds true for the term “short baseline”.

- Long batches - Short baseline (<15 km)

As Eq. (19) accounts for a geometrical satellite individual weighting, the computational burden induced by estimating the parameters for each satellite can be reduced to the estimation of one set of parameters $[\alpha, \nu]$ for a given satellite with a mean elevation of 40-60° observed during 1-2 hours. The correlation structure is thus depending on the observations and e.g. meteorological conditions at each station. Although from the results of simulations not mandatory, a higher accuracy in the parameter determination can be obtained by estimating satellite-dependent parameters. Both procedures allow for the computation of fully populated VCM at the two stations and can be further simplified by constraining the smoothness to 1 following the approximated structure of the tropospheric VCM (Eq. (17), smoothness 5/6) as mostly atmospheric effects lead to elevation-dependent correlations (Kermarrec et al. 2017). This way, only the correlation parameter has to be estimated. This simplification follows the results of Luo et al. (2012) who shows that AR processes of low orders accurately approximate the correlation structure of the stationary residuals. Moreover, taking a smoothness parameter larger than 1.3 leads to numerical instability. Fixing the smoothness, the property of the non-orthogonality of the Matérn function is used (Diggle and Ribeiro 2007, Gerland et al. 2010). An empirical validity of 4 hours for the Matérn parameters set is proposed in order to account for changes of the geometry or atmospheric conditions at the stations (Kermarrec and Schön 2017a). It should be noted that for an accurate determination with MLE, the batch of observations chosen should be long enough (i.e. minimum 1 hour).

- Long batches - Medium and Long baselines (>15km)

For medium-long baselines (>15 km), the correlation structure of the double differences may not be reflected in the OMC observations of each station. Thus, the MLE values are not computed from the observations themselves but directly from the double differences. As previously, the smoothness can be fixed to 1. Eventually, if the overall model test fails, α can be decreased or increased by steps of 0.005 so that $\frac{\hat{\sigma}_{\hat{\mathbf{W}}}^2}{\sigma_0^2} < F_p\left(n-u,\infty,0\right)$.

- Short batches

In this scenario, the effect of taking correlations into account becomes intuitively stronger as the correlation length and batch size have close values. In such cases, no tropospheric parameters are usually estimated in order not to weaken the data strength. Thus, when the covariance matrix is strongly fully populated, the equivalence stochastic-augmented functional model presented in Sec. 2.5 can be applied. The value of the smoothness can be fixed slightly higher than 1 following the physical explanations proposed in Kermarrec et al. (2017). It is close to the 5/6 value of the tropospheric model but allows a mean-square differentiability at the origin of the corresponding covariance function, this property being physically desirable considering GNSS observations as results of intensity/voltage measurements. The correlation length parameter can be taken empirically to 0.01s^{-1} in a first approximation (Kermarrec and Schön 2014) and varied by steps of 0.005 depending for instance on the results of the overall model test. For numerical stability it is advantageous that the covariance function reaches the zero-value inside the batch, i.e. α has to be approximated to fit this criterion (Stein 1999). Incorrect stochastic models and in particular, too high correlation lengths are detected by the overall model test, conservative parameters which lead to sparser VCM being always preferable. For very short sessions (<50 epochs), the F-distribution may not hold anymore due to the violation of the normal distribution of the residuals. Alternatively, a student distribution could be used (Schön et al. 2018).

4.2.4. Impact of the covariance function on whitening the observations: an example

The fully populated matrices with Matérn parameters estimated as previously described were shown to be able to whiten the double differenced GPS observations better than the corresponding diagonal matrices, particularly in the low frequency domain. A whitened time series $\mathbf{y}_{\mathbf{white}}$ is defined as $\mathbf{y}_{\mathbf{white}} = \hat{\mathbf{W}}^{-1/2}\mathbf{y}$ where $\mathbf{y}$ is the original time series and $\hat{\mathbf{W}}^{-1/2}$ is the inverse of the square root of $\hat{\mathbf{W}}$. A typical challenging scenario is presented in Fig. 4, where the double differenced time series correspond to the medium baseline (66 km length ZYWI-KRAW following Sec. 4.1.3). Using the FULLY matrix with $[\alpha,\nu]=[0.01,1]$, i.e. accounting for atmospheric-like correlations, the whitened observations are shown to correspond more to zero-mean white noise compared with the time series whitened both with the ELEV matrix (magenta line) or the FULLY matrix with $[\alpha,\nu]=[0.01,1/2]$ (green line), i.e. the exponential covariance function. This holds particularly true for observations related to low elevation satellites (PRN19,7).

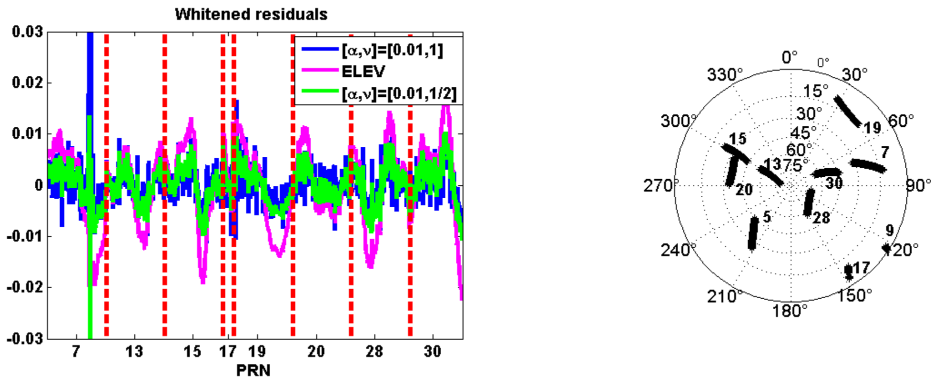


Fig.4 right: Whitening potential of a FULLY matrix (blue and green lines) with respect to a diagonal ELEV model (red line). The original time double differenced time series corresponds to a medium baseline of 66 km, the observations were not preprocessed and sorted per satellites (30s data rate, 100 epochs), right: corresponding skyplot

The analysis of the amplitude spectrum from a Fourier decomposition of the whitened double differences PRN5-30 allows a more accurate comparison between the models. The satellites were observed during 100 epochs (i.e. approximately 1 hour of observations). The corresponding results found by varying the structure of $\hat{\mathbf{W}}$ (correlation length and the smoothness) are presented in Fig.5. Other examples can be found in Kermarrec and Schön (2017c).

Due to the elevation dependency of our model, the whitened double differenced time series are automatically studentized (Luo 2012). No further filtering of the data was applied, the study of the whitening potential of the proposed double differenced VCM being the focus of this short analysis, i.e. for instance the impact of a fully populated VCM on the presence of low frequencies coming from unmodelled effects. The exact values of the amplitudes are here not of interest, as soon as they remain comparable for each VCM which is the case thanks to the studentization. Using FULLY matrices, it can be seen that the amplitudes of the low frequency part of the spectrum are decreased whereas the high frequencies are stronger with respect to the ELEV model. Generally, the smoothness parameter acts on the high frequency part of the spectrum (green line with $\nu = \frac{1}{2}$ versus blue line with $\nu = 1$) whereas the correlation length has an effect on the low frequency part of the spectrum. This behaviour is directly related to the rational spectral density of a Matérn function (Stein 1999). However, it can be noted that the change of the correlation length from 0.007 to 0.01 s^{-1} has only a small influence at the amplitude level. It is furthermore confirmed that the exponential correlation model with $\nu = \frac{1}{2}$ is suboptimal to whiten the observations, particularly at high frequencies but to some extent also at low frequencies (Kermarrec and Schön, 2017c).

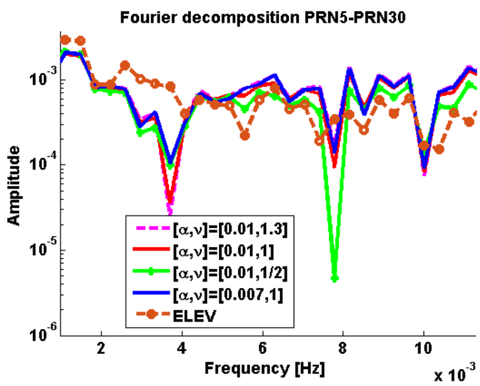


Fig.5 Fourier decomposition of different whitened observations versus log-frequency (Hz), Amplitude as log scale. Whitened L1 double differences by varying the correlation length and the smoothness parameter of the VCM (100 epochs a 30s) Double differences PRN5-30

This case study is an example. However, other baselines with different lengths and satellites were computed without changing the previous conclusions on the spectrum. It is clearly not to be expected from an apriori covariance matrix that the observations will be perfectly whitened. Moreover, some factors cannot be "repaired" or taken into account in the stochastic model which remains anyway an approximation of the true but unknown VCM.

4.3. Simulations: sensitivity analysis of the parameters

From Eq. (19), a VCM with a desired correlation structure can be built. In order to assess the effect of misspecifying the stochastic model on the Least-Squares results, simulations are necessary. This way, the structure of the true covariance matrix is exactly known and observations with the corresponding behaviour simulated. Following Sec. 4.2. and Fig. 1, the effects of varying the Matérn parameters can be quantified and compared e.g. with the results given with the corresponding ELEV diagonal model. The main results are presented in Kermarrec and Schön (2017a) and Kermarrec and Schön (2017c).

The main focus is on the differences when correlations are taken into account (FULLY model), misspecified or neglected (ELEV model), i.e. the same diagonal variance model is used. This is a more realistic comparison than usually found in the literature.

Dealing with phase observations that are inherently ambiguous, two cases were identified to be treated:

1. Biases of the aposteriori variance factor and loss of efficiency due to the stochastic model when the ambiguities are fixed in advance and impact at the estimates level.
2. Biases of discriminant validation tests, when the ambiguity vector is estimated together with the other parameters (i.e. the position) and eventually led to its float value.

4.3.1. Case 1: ambiguities fixed in advance

Extended results are presented in Kermarrec and Schön (2017a). It includes the detailed effects of adding noise matrices, changing the batch lengths or the known correlation structure of the true VCM. Moreover, the impact of neglecting the correlations between observations from different satellites on the Least-Squares solution is addressed. For this study, the reference VCM was chosen to have a correlation structure corresponding to a Matérn parameter set $[\alpha,\nu]=[0.01,1]$ which is a physically plausible correlation structure for double differenced observations from medium baselines. A relative positioning was adopted although no main change occurs by using other strategies. The reasons for these findings are developed in the next chapter based on the structure of the design matrix or in Kermarrec and Schön (2017a, Appendix 1). The VCM was sorted per satellites, i.e. a batch-based scenario was adopted. Following the methodology of Sec. 4.1., only the main results are summarized below:

- Aposteriori variance factor

Using the result of chapter 2, the error when misspecifying the stochastic model is expressed as a trace of a matrix product (Eq. (5)). It may have a minimum which is not corresponding to the correct solution. In the ideal case of simulations, a small bias of $0.1m^2$ for the chosen structure of reference was found when the diagonal ELEV model was wrongly used. Conservative parameters (i.e. low correlation length and low smoothness) were preferable to wrongly estimated higher correlation lengths than the reference that were associated with an exponential increase of the bias of the variance of unit weight. These results were consistent with Rao (1967). It is worth noticing that different sets of Matérn parameters yield the unbiased variance of unit weight solution which is explained by the non-orthogonality of the Matérn parameters.

- Relative efficiency / Mean Squared Error

Studying the relative efficiency of an estimator (Eq. (8)) is a powerful tool to assess the extent to which the approximated estimated VCM impact the precision of the solution. It was shown that correlation lengths lower than the reference, particularly for higher smoothnesses (+0.3) than the original one, were responsible for a positive increase of the relative efficiency, i.e. the outcome follows the shape of the Frobenius norm of the matrix difference (true minus estimated). As for the bias of the aposteriori variance factor, higher correlation lengths than the true one led to a rapid increase of the relative efficiency. Taking wrongly a smoothness corresponding to the exponential correlation model ($\nu = 1/2$) made the results being closer to the one given with the diagonal model ELEV, i.e. overoptimistic. However, the mean squared difference of the estimates obtained with FULLY and ELEV, for the assumed reference correlation structure moved at the submm level, which is negligible.

- Apriori covariance matrix of the estimates

Considering the example of the covariance matrix corresponding to an AR(1) process which inverse is obtained with a closed formula, the effects of fully populated VCM on the apriori covariance matrix of the estimates are presented in Appendix 2 of Kermarrec and Schön (2017a) or graphically in Kermarrec and Schön (2017c). Whereas the use of a fully populated VCM built with Eq. (19) only slightly changes the orientation of the corresponding error ellipsoid by a few degrees maximum with respect to the ELEV model, its volume is strongly affected. The overestimation of the precision with the diagonal model is corrected due to the fully populated VCM which acts as a scaling factor. This mathematical effect comes from the inverse of the VCM, its eigenvalues exhibiting higher values depending on the chosen correlation structure. Correspondingly, the scaling strategies presented in the literature and described in Sec. 3.3 can be better understood as a way to artificially increase the determinant of the cofactor matrix of the estimates. However, this procedure does not allow to get a statistically trustworthy solution which is only possible by using realistic fully populated VCM in the LS adjustment.

4.3.2. Case 2: ambiguities estimated as part of the solution

Through the contributions associated with this thesis, the Lambda method (Teunissen 1995) was adopted to fix the float ambiguity vector to its integer values. Within this fixing strategy, the VCM of the observations is introduced in the cofactor matrix of the ambiguities. The focus being here on the impact of the stochastic model on the error of the float solution as well as on different discriminant tests such as the ratio test (Euler and Schaffrin 1991), other methods would have led to the same conclusions.

- Discriminant tests, distance fixed-float ambiguity vector

Following the methodology for the simulations explained in Sec. 4.1.1., the Matérn parameters of the estimated VCM were varied around the reference value. The true fixed ambiguity vector was entirely known. As a consequence, the impact on different validation tests such as the ratio test or R ratio, as well as the distance between the float and the fixed solution in the same metric could be studied. Since the discriminant tests are based on the observations, no general formulation of the biases under the wrong stochastic model can be developed and Monte Carlo simulations were used.

From Fig. 6 right, a minimum of the ratio test is obtained when using the correct VCM, similar to the shape of the relative efficiency (Sec. 4.3.1.). The same conclusion hold true for the Euclidian distance between the float and the true integer ambiguity vector. Thus, neglecting correlations when they exist can be linked with a slight decrease of the probability of fixing the ambiguity vector to integer. However, the same fixed ambiguity vector is obtained with fully populated or diagonal models, except in case of overestimation of the correlation length, for which the overall model test also fails. This highlights the bias of the corresponding test statistics and the lack of trustworthiness of the LS solution. The non-uniqueness of the "nearly" best solution is furthermore highlighted, a given variability of the Matérn parameters being allowed without impacting strongly the results.

An error in the distance between the float and the integer ambiguity vector with the ELEV model can be further identified (Fig. 6 left). This error increases as the smoothness or the correlation parameter decreases, i.e. neglecting or overestimating the correlations. Thus compared with Sec. 4.3.1., it is expected that the more accurate float ambiguity solution plays a central role when the ambiguities cannot be fixed with enough confidence and are let floating.

Although these simulations may seem to depend on the design matrix and thus be not reproducible, it is worth mentioning that only the behaviour of the quantities versus the Matérn set are of interest in this study. Thus, the corresponding shape of the presented curves does not change with other satellite constellations or batch lengths (Kermarrec and Schön 2017a).

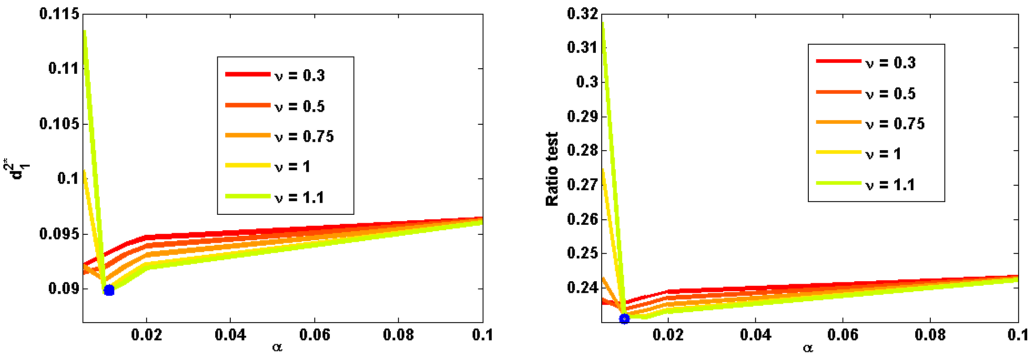


Fig.6 Impact of the estimated covariance matrices on the ambiguity validation. Left: the distance float-fixed $d_1^{2*} = \left\| \mathbf{x}_{\text{fix}} - \hat{\mathbf{x}}_{\text{float}} \right\|_{ID}$ between the correct integer vector $\mathbf{x}_{\text{fix}}$ and the estimated float ambiguity vector $\hat{\mathbf{x}}_{\text{float}}$ (Euclidian distance). Right: Ratio test. The Matérn parameters are varied around the known structure of $[\alpha, \nu]_0 = [0.01, 1]$ corresponding to the blue point (true solution). For this Monte Carlo simulation, a standard GPS constellation with 8 satellites was taken as well as batches with 100 epochs a 30s (Fig.3)

4.3.3. Summary of the simulations

Following the previous results, a two-step procedure is proposed in Fig. 7 to prevent the highlighted loss of efficiency of the LS estimator and biases of test statistics such as the overall model test or ambiguity discriminant tests, when a wrongly estimated stochastic model is used.

In a first step, the float ambiguity vector is estimated as part of the LS solution. Based on the result of the Ratio test (a threshold of 0.5 is usually considered, Wei and Schwarz 1995), it is either let floating or fixed to an integer. In a second step, the overall model test is applied following the methodology of Sec. 4.1.2.1. to compare the aposteriori variance factor under the estimated VCM with a realistic and plausible apriori value. The chosen value depends on the processing strategy (e.g. relative or single positioning). The Matérn parameters may have to be adapted depending on the outcome of the overall model test, i.e. the correlation parameter α decreased by step of 0.005 if the test fails. Using this procedure, a model misspecification can be detected with a higher confidence. At the same time, a statistically trustworthy least-squares solution is expected with a minimum loss of efficiency allowing at the same time more ambiguities to be fixed to an integer due to more accurate test statistics.

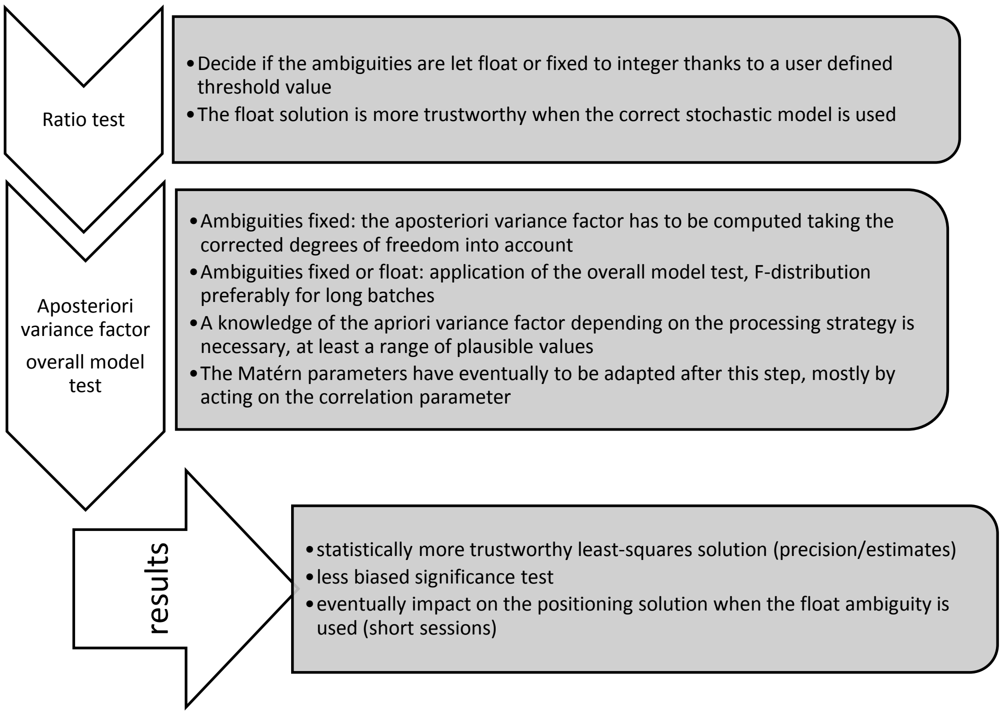


Fig.7 Summary of the results from simulations using a fully populated VCM: a proposal to prevent erroneous solutions when the correlations are unknown and estimated.

4.4. Real data analysis

The previous conclusions of the simulations (Sec. 4.3) on the impact of the stochastic model on the Least-Squares adjustment were confirmed for real case studies (Sec. 4.1.2.). Following the same methodology, a sensitivity analysis was carried out by varying the Matérn parameters in a physically plausible range. The corresponding results can be found in Kermarrec and Schön (2017c) for the case 1 (ambiguity fixed) and Kermarrec and Schön (2017e) for the case 2, i.e. estimating the float ambiguity vector as estimates together with the position. In this last contribution, the Matérn parameter set was computed based on the strategy described in Sec. 4.2.3. for the short batches case.

4.4.1. Case 1: ambiguity fixed

Similarly to Sec. 4.3.1., the ambiguities were firstly considered to be resolved in advance, i.e. fixed to integer. From the simulations, modelling elevation-dependent correlations in case of fixed ambiguities was shown to lead to more trustworthy precision and aposteriori variance factor. The mean-squared errors of the estimates difference were at the submm level for the ideal simulated case corresponding to zero-mean observations having an exactly known correlation structure.

Kermarrec and Schön (2017a,c) confirmed the results of simulations for real data analysis at the estimates level as only mm-submm differences were obtained between the results given under the FULLY or ELEV model for a baseline length of 66 km (KRAW-ZYWI, see Sec. 4.1.3.). This result, presented in Tab.3 by varying the correlation structure of the estimated FULLY VCM, is expected, the Least-Squares estimator being unbiased (Sec. 2.2.). However, using the overall model test, the ELEV solution was shown to be statistically incorrect, i.e. the value $E\left(\hat{\sigma}_{Welev}\right)$ was up to a few mm higher than the values found under the FULLY model (9.7 mm). For long sessions, σ_0 could be adjusted to $E\left(\hat{\sigma}_{Welev}\right)$, corresponding to the mean of $\hat{\sigma}_{Welev}$ over all batches (Kermarrec and Schön 2017c). Using this strategy, similar results as with FULLY VCM were reached, i.e. up to a few sub-mm differences. However, because a plausible apriori of 4mm for relative positioning σ_0 was only used for the FULLY case, this proposed "bias corrective procedure" remains statistically incorrect. From the simulations, high aposteriori variance factors are linked with a higher loss of efficiency and thus a biased and incorrect significance.

Table 3 highlights further the already mentioned non-uniqueness of the best solution, i.e. different values of $\left[\alpha,\nu\right]$ gives the same 3Drms. Moreover, the effect exaggerating the correlation length by decreasing α for the same smoothness of 1 (e.g. $\left[\alpha,\nu\right]=[0.001,1]$) degrades the solution by up to 20 mm, less batches being used to compute the solution due to the higher $\hat{\sigma}_{Wfully}$. Such cases were identified in the simulation as corresponding to a high loss of efficiency of the LS estimator. The exponential model corresponding to $\left[\alpha,\nu\right]=[0.01,\frac{1}{2}]$ is similarly suboptimal. The ID model disregards the heteroscedasticity of the observations and is detected as a model misspecification.

	$[\alpha,\nu]_0 = [0.012, 1.1]$	$[\alpha,\nu] = [0.005, 1]$	$[\alpha,\nu] = [0.001, 1]$	
$E(3Drms)$ [mm]	***55.9***	63.4	75.0	
$E(\hat{\sigma}_{\hat{w}})$ [mm]	***3.4***	3.9	4.5	
	$[\alpha,\nu] = [0.01, \frac{1}{2}]$	$[\alpha,\nu] = [0.005, \frac{1}{2}]$	ELEV	ID
$E(3Drms)$ [mm]	56.1	55.9	*57.0* (adapted σ_0)	*82.7* (adapted σ_0)
$E(\hat{\sigma}_{\hat{w}})$ [mm]	3.8	3.4	*9.7*	*22.8*

Tab.3 $E(\hat{\sigma}_{\hat{w}})$ and the mean of the 3Drms computed with estimated VCM. The FULLY model by varying $[\alpha,\nu]$ is compared with the results given with the ELEV and ID models, i.e. neglecting correlations. From Kermarrec and Schön (2017c).

Observations from short baselines or from the ionosphere-free linear combination (L3) led to similar conclusions. As many effects cancel out in such cases and the observations are ideal (zero-mean, Gaussian distribution), the impact of taking correlation into account follows even more closely the results of the simulations. At the estimates level, sub-mm differences are obtained when comparing different models. However, test statistics still remain less biased, particularly for short batches of observations (<1 hour). Thus, risky potential underestimation of the aposteriori variance factor can be avoided by taking correlation into account. This can impact positively the 3Drms due to a more realistic rejection of batches if the overall model test fails, what can be seen as a “snowball effect”. Multipath contaminated observations of short sessions can thus profit from improved stochastic model. However, a realistic apriori variance factor has to be considered.

In Kermarrec and Schön (2017a,c), the impact of estimating or not a tropospheric parameter was not addressed, the batch length chosen being less than 1 hour, i.e. 100 epochs a 30s. For longer batches from long baselines, the use of a FULLY model does not avoid the estimation of an additional tropospheric parameter as the condition for the equivalence stochastic-functional are not fulfilled (see chapter 5 for more details).

4.4.2. Case 2: the float solution

In this thesis, it is investigated in which cases and at which level taking correlations into account impacts the LS solution. Thus it is not aimed to fix the ambiguities to integers by all means. Following the results of simulations, the impact of the less erroneous float ambiguity when taking correlations into account was identified as an important quantity to test in real cases. Therefore, the global approach was adopted where parameters and ambiguities are estimated together in a Least-Squares adjustment. The ambiguities are fixed if the validation test (ratio test) is below a threshold value of 0.5, i.e. the computationally demanding Fixed Failure Rate Ambiguity Validation Method (Wang and Feng 2013) was not adopted. In that case and in the simulations, the improvement of the stochastic model was shown to lead to a more correct float ambiguity vector and a slightly lower ratio test value. In a real case scenario, this result turns out to be the main improvement that can be expected from taking correlations into account at the estimates level.

The baseline ZYWI-KRAW (Sec. 4.4.1.) was retained for this study. The following results can be extended for other observations, such as those from short baselines with strong multipath or higher data rate (1 Hz). Three batch lengths were selected to show the influence of the stochastic model using the global approach. In order to average over different geometries and to study the float solution with more details, the total observation time span was subdivided into:

a) 20 batches with 200 epochs (session length 100 minutes)
b) 40 batches with 100 epochs (session length 50 minutes)
c) 80 batches with 50 epochs (session length 25 minutes)

For short batches of observations, the fixing strategy of the Matérn parameters as presented in Sec. 4.2.3. could be adopted without additional ML estimation. Indeed, this case corresponds to an estimation of a tropospheric parameter by the stochastic model. No overall model test was applied due to the computation of short sessions for which the student distribution should be preferred (Schön et al. 2017).

The results are shortly presented in Tab. 4. As the ratio test values are influenced by the chosen stochastic model, the FULLY model allowed slightly more batches to be fixed (5-10% in our case study). As a consequence, estimated 3Drms differences at the cm level for session length of up to 1 hour between the ELEV and FULLY models were reached. Thus, both the great potential of a more correct float solution and less biased test statistics is highlighted for very short sessions. These results were confirmed with 1 Hz observations and are thus independent of the data rate or the baseline length as soon as the fixing of the ambiguities to integer is challenging and the float ambiguity vector is used.

For longer sessions of observations (i.e. from one hour), the 3Drms differences between the ELEV and FULLY models are less important. This has to be linked with the short correlation length compared with the batch length as well as the fact that ambiguities are fixed to integer with a higher confidence (Sec. 4.4.1.).

The mean value of $\hat{\sigma}_{\hat{w}}$ presented in Tab. 4 is below 4 mm for the three cases with the FULLY model and remains closer to the apriori values for double differences. The less biased $\hat{\sigma}_{\hat{w}}$ with respect to the ELEV case highlights the previously mentioned higher confidence that can be put in the LS solution, i.e. a lower loss of efficiency. The float solution and its associated precision are thus more trustworthy (Sec. 4.3.).

Ratio test: threshold 0.5	FULLY	ELEV
3Drms (mean) [mm]		
a)	28.5	29.1
b)	82.3	84.2
c)	184.0	216.4
Aposteriori variance factor [mm] $\hat{\sigma}_{\hat{w}}$		
a)	4.1	9.9
b)	3.8	8.0
c)	3.1	4.8

Tab.4 Comparison of the impact of the batch length on the 3Drms for 3 particular scenarios: a) 2 hours, case b) 1 hour and case c) 30 minutes. A threshold of 0.5 for the ratio test is applied. The aposteriori variance factor is given for the two models: FULLY and ELEV. The apriori variance factor is assumed between 2 and 4 mm (double differences).

4.5. Concluding remarks

In this chapter, an innovative way to take correlations into account was proposed. Built with a Matérn covariance function weighted by an elevation-dependent factor, heteroscedasticity and spatio-temporal dependency of the observations are thus taken into account. Correlations between observations from one satellite or different satellites can be easily computed thanks to an additional factor that can be fixed apriori based on physical considerations. This proposal avoids thus the shortcoming of the previous empirical model based on a LS-fitting of the autocorrelation of the residuals with an exponential function. Derived from a model to account for correlations due to turbulent variations of the refractivity index in the atmosphere, the physical component of the extended function was highlighted. Two Matérn parameters called the smoothness and the correlation length of the observations can be either fixed or estimated, depending on the processing strategy and the baseline length. The most optimal whitening of double differenced observations was obtained for a smoothness parameter of at least 1 (Kermarrec and Schön 2017b). The correlation parameter impacts the low frequency part of the Fourier spectrum of the observations and has thus a stronger impact for short sessions and suboptimal data.

This model allows to build realistic fully populated VCM that can be used in a LS-adjustment. Simulations have shown that neglecting correlations when the simulated observations are correlated led to a higher loss of efficiency of the LS estimator with a biased aposteriori variance factor. As a consequence, even if the solution remains unbiased by using different estimated VCM, a more trustworthy solution can be expected, i.e. an estimated precision that corresponds to the scatter of the coordinates (Kermarrec and Schön 2014). This overcomes thus the overoptimistic precision obtained with diagonal VCM. Due to the less biased validation test used for ambiguity resolution, a reliable fixing of the ambiguity vector to an integer is reached. Combined with a float solution with a smaller error, the 3Drms of the coordinates could be improved by up to 30 cm for a case study with 30 minutes-batches of observations.

5. Simplifying the stochastic model: why and how?

In this chapter, the two last questions raised in Sec. 3.4. are answered. In a first section, the effect of correlations in the Least-Squares adjustment is explained using an intuitive approach. Then, the expected impact on the Least-Squares solution when correlations are taken into account is summarized, introducing the concept of the “hidden parameter” for short batches of observations, giving an additional justification of the results from Sec. 4.4.2.. In a last section, the equivalent diagonal model is presented. This chapter follows the results of the publications:

- Equivalent diagonal model:

  Kermarrec G, Schön S (2016) Taking correlation in GPS Least-Squares adjustments into account with a diagonal covariance matrix. Journal of Geodesy 90(9):793-805

- Diagonal correlation model:

  Kermarrec G, Schön S (2017d) Taking correlations into account: a diagonal correlation model. GPS solution 21(4):1895-1906

- The hidden parameter:

  Kermarrec G, Schön S (2017e) Fully populated VCM or the hidden parameter. Journal of Geodetic Science 7(1):151-161

5.1. About the impact of correlations

In this section, an intuitive interpretation of the impact of correlations on the Least-Squares solution is proposed. It aims to understand why correlations -when present- should not be neglected. The low effect of the underlying geometry or positioning strategy is shortly developed.

5.1.1. Inverse of the covariance matrix

Independently of the stochastic model used, the Least-Squares estimates remain unbiased. Thus, no strong differences between the ELEV and FULLY models are expected in an ideal positioning scenario as presented in chapter 4, i.e. mainly when the residuals are zero-mean. In this section, the structure of the inverse of the fully populated VCM is used to explain “visually” the effect of correlations on the solution and its precision. It aims to help understanding the results found for short batches of observations.

For a didactical explanation, a fully populated VCM with a structure corresponding to a simple AR(1) process is here considered. As illustrated in Kermarrec and Schön (2017a) for the case with fixed ambiguities, the three eigenvalues of the cofactor matrix of the estimates are, up to a “correlation factor” which depends on the chosen correlation parameter α, similar with or without correlations, the eigenvectors being the same. This factor is the reason why the precision is not overestimated under the FULLY model (Kermarrec and Schön 2017c) and is related to the structure of the inverse of the fully populated VCM.

Analysing the slope matrix (Eq. (2)) more closely in that particular case, it can be observed that the two first and last values of one line of the slope matrix for one particular satellite exhibit different values for the fully populated model than for the diagonal model (Fig. 8 left). In the middle of the batch, ELEV and FULLY leads nearly to the same slope (Fig. 8 right), i.e. the mean of the values of $\mathbf{K}$ for one satellite should be similar due to the unbiasedness of the LS estimator. For less simple correlation structures, corresponding for instance to a Matérn model with higher smoothness (i.e. when the covariance function is mean square differentiable as defined in

Sec. 2.1. with $\nu > 1$), the number of values oscillating at the beginning and end of one batch for one satellite around the constant slope will be slightly more important by remaining small with respect to the number of epochs. Luati and Proietti (2011) showed e.g. that this number increases following the order of the AR process of consideration. The closed formula of the inverse of the AR(1) gives e.g. an explanation for this behaviour (Sec. 5.4.1., see Kermarrec and Schön 2016). For long batches of observations, this shape has however only a small impact on the Least-Squares estimates as long as the beginning and ending observations are not outliers (i.e. unbiasedness of the LS estimator under the assumptions of Sec. 2.2). For a batch of 100 epochs per satellite, more than 98% of the observations are combined identically with the FULLY and ELEV model.

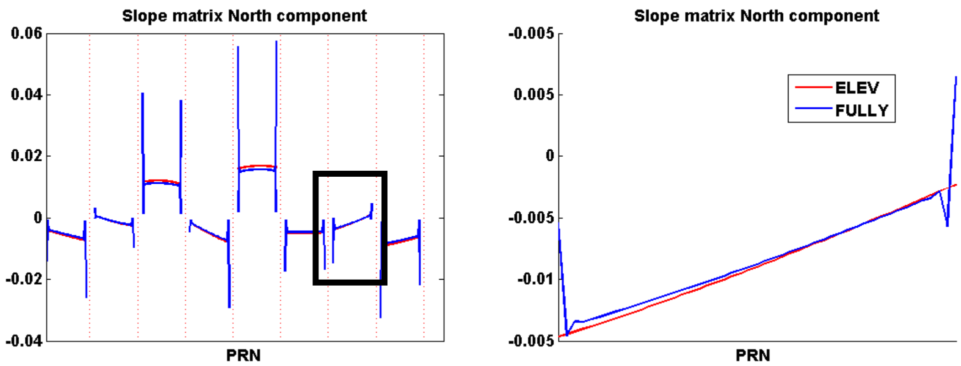


Fig.8 Left: First line of the slope matrix for the FULLY and ELEV VCM sorted per PRN corresponding to the North component. The first and last epoch exhibit stronger values (AR(1) case). Right: zoom corresponding to a particular PRN (black box in Fig.4 left)

The previous conclusions on the slope matrix are similar for $\mathbf{K}^{1/2} = \left(\mathbf{A}^{\mathrm{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\hat{\mathbf{W}}^{-1/2}$. This matrix being multiplied by $\hat{\mathbf{W}}^{-1/2}\mathbf{y}$, i.e. the whitened observations vector (Sec. 4.2.4.), the impact of the fully populated VCM is more important depending on how powerful the whitening is. As for short batches of observations, the ideal condition of the LS estimator are more often violated (zero-mean or normal distribution), the impact of taking correlation into account on the position increases, the whitening effect of the FULLY VCM having a stronger effect on the spectrum by decreasing the low frequencies content of the whitened vector of observations with respect to diagonal VCM.

5.1.2. Similarities with BLUP and impact of the positioning scenario

In Kermarrec and Schön (2017a) Appendix 2, an interpretation of the small dependencies on the underlying geometry (i.e. the design matrix) was proposed. Considering the positioning as an interpolation, results from the BLUP (Best Linear Unbiased Predictor) can be applied, where the relative efficiency can be expressed as an integral of estimated and true spectral density. The corresponding figure by varying the Matérn parameters shows similarities with the one given for a GPS positioning scenario (Kermarrec and Schön 2017a), allowing to motivate the low impact of changing the design matrix on the results.

Throughout our work and for the sake of homogeneity, the relative positioning scenario with double differences was chosen. The previous comment on the impact of the design matrix clarifies that changing the positioning scenario, e.g. PPP or single positioning, will not strongly change the conclusions of this work.

5.2. On neglecting correlations or not

In the previous chapter, two scenarios were identified and treated: ambiguities fixed in advance or estimated together with the position in a global strategy. Whereas in the first case the main improvements with respect to the stochastic model were mostly related to the test statistics and by “domino effect” could eventually impact the averaged solution, in the second case, a stronger effect was highlighted for both quantities, particularly for short batches of observations. Sec. 5.1.1. proposed a first interpretations of this result based on the slope matrix. Thus, neglecting correlations does influence the Least-Squares solution and should be avoided for more trustworthy results, test statistics and precision, i.e. a globally statistically correct solution. In this section, it is aimed to specify in which cases and for which purposes correlations should be considered or not in the LS adjustment.

5.2.1. Long sessions: ideal case

Placing ourselves in the ideal case of observations corresponding for instance to short baselines without multipath (zero-mean, normal distribution of the residuals, high amount of high frequencies in the observations), the ambiguities are easily fixed to integers. As a consequence, the effect of taking correlations into account is small, i.e. the errors, biases and the loss of efficiency were shown in simulations to remain similar to the results with a diagonal VCM having the same ELEV variance. Using diagonal VCM is moreover a better alternative than a wrongly estimated correlation structure, i.e. particularly with too low correlation length when no overall model test is used. For long sessions of observations (> 1,5 hours), as long as the variance of unit weight remains statistically correct and a plausible apriori value is used, correlations could be neglected. Nevertheless, in order to increase the trustworthiness of the solution, it remains always possible to estimate the Matérn parameters by MLE following Sec. 4.2.3. However, as our proposal is based on an ELEV variance model which may be suboptimal when the geometry and data quality are enough averaged, an exponential variance function as proposed by Luo et al. (2014) is a better alternative in that case. Section 5.4. highlights how such variance models can be considered as improved "hidden" correlation models.

5.2.2. Short sessions: the hidden elevation-dependent parameter

The impact of correlations increases with the correlation length or the amount of power at low frequencies, particularly for short batches and independently of the data rate, as the shape of the slope matrix (Fig. 8) highlighted. In such cases, the normal distribution of the errors as well as the zero-mean assumption is often questionable. Thus, using a more accurate stochastic model by means of fully populated matrices that whiten the residuals is an answer to these violations.

If the global approach is used and float ambiguities are estimated together with the position without being fixed, the stronger impact of the FULLY model at the estimates level with respect to the ELEV model was highlighted. Taking correlation into account should be thus definitively preferred to a diagonal model. Moreover, in such cases, it could be shown that the Matérn parameters can be taken to fixed values ($[\alpha, \nu] = [0.01, 1]$, Sec. 4.2.3.) to make use of the equivalence between the improved stochastic model and the augmented functional model (Kermarrec and Schön 2017e). Indeed, the VCM computed with the Eq. (19) "models" for short batches of observations a non-estimable tropospheric parameter, which if estimated would have weakened the data strength and influenced negatively the Up parameter. Thus, in that particular case, a specific smoothness and correlation length can be taken into account for a trustworthy solution. However, and particularly for longer baselines, the correlation length still can be moved around the values used to model correlations due to the troposphere depending on the results of the overall model test. It is proposed to act mostly on the correlation length by steps of 0.005 by letting the smoothness slightly higher than 1 for physical reasons (Kermarrec et al. 2017b). In order to avoid the negative impact of a wrong specification of the apriori correlation structure, the aposteriori variance factor should be controlled in order not to overestimate the correlation length. The knowledge of the apriori variance of unit weight value to which the aposteriori value has to be compared is thus necessary. Ranges of plausible values are however available (Sec. 4.1.2.).

On micronumerosity…..

Micronumerosity is defined as the impact of very small batches on the results of regression. This topic is often not addressed because no rule of thumb can be found to assess when the length of the batches may be problematic with respect to the decrease of precision (Goldenberg 1991, chapter 23). As the sample size decreases, the impact of a correct stochastic model and its ability to eventually correct or whiten errors that are not exactly zero-mean becomes important. Fully populated matrices have this capacity, i.e. the low frequencies amplitudes of whitened double differenced observations are smaller with respect to diagonal models and thus drifts are reduced. This is the “hidden” parameter concept. Due to the use of the float solution, particularly for long baselines, the impact of correlations becomes automatically stronger. For small batches, student distribution instead of the F-distribution should be preferred for the overall model test (Williams et al. 2013). This change was not investigated in this work and remains an open question.

5.2.3. Which stochastic model for which effect: a summary

Figure 8 proposes a summary of which stochastic model can be used for which expected effect. Thus, particularly for longer baselines and short observation sessions, the computation of fully populated VCM is of key importance for better precision, test statistics and estimates, i.e. a realistic and trustworthy solution.

The covariance model developed in this thesis is a good answer to this challenging task by being both physically plausible and simple. Its main limitation was identified for long batches where improved variance models are a better alternative. However, it remains difficult or unwanted to compute fully populated VCM for some applications. As a consequence, a proposal was made to reduce such matrices into their diagonal equivalence. The results are presented in Kermarrec and Schön (2016) and Kermarrec and Schön (2017d). They allow to draw a parallel between an empirical exponential variance model and the proposed correlation model. Corresponding developments are presented in the next section but already included in Fig. 9 for the sake of readability.

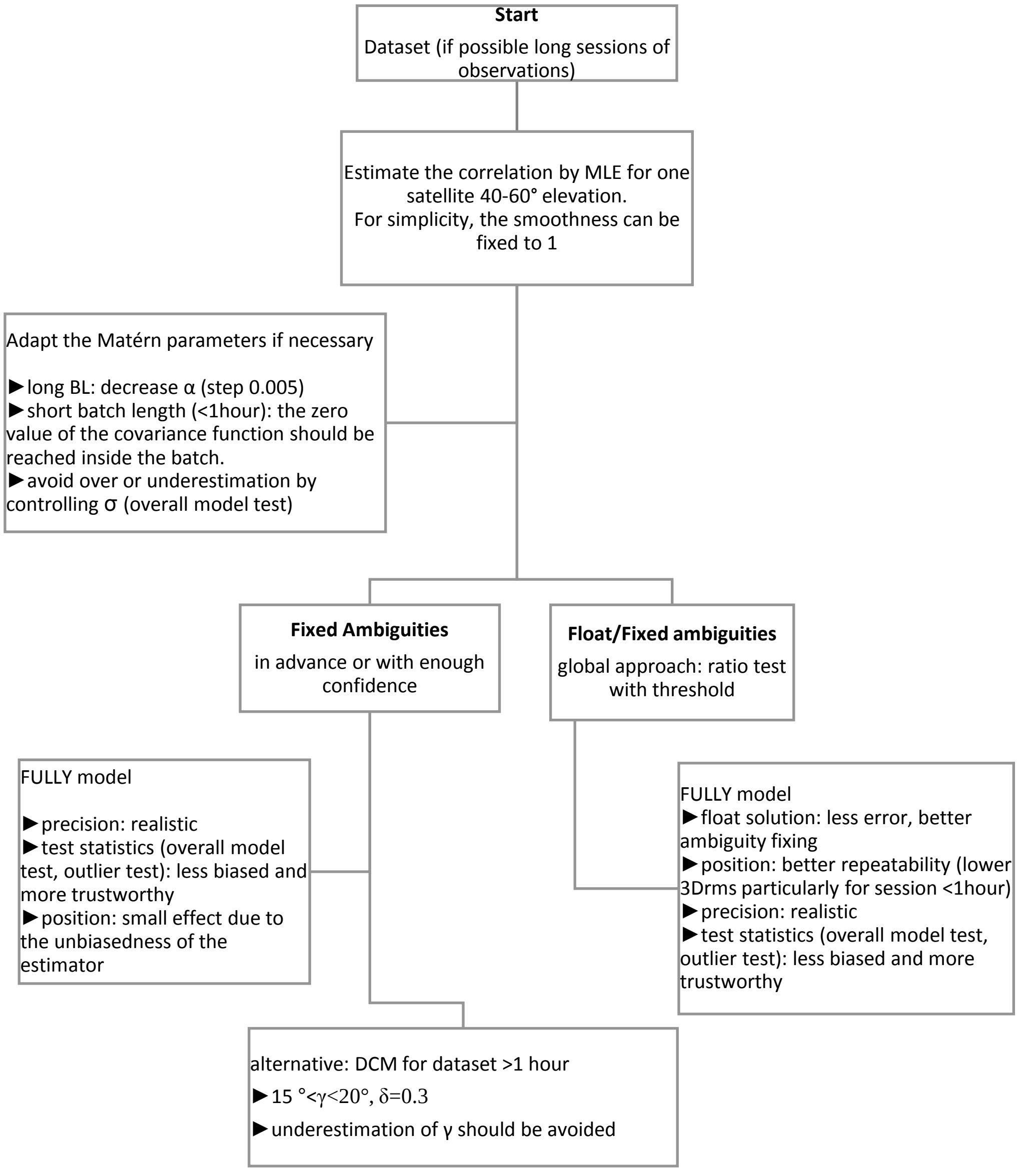


Fig.9 Impact of the stochastic model: a summary

5.3. The equivalent diagonal model

Taking correlations into account in LS-adjustments was identified as useful for a trustworthy solution and should not be neglected. However, fully populated matrices are less easy to invert and handle than diagonal VCM. In this section, an equivalent diagonal model, further extended in a diagonal correlation model are presented. They allow to account for correlations in a simplified way, linking thus variance and covariance functions presented in Sec. 3.1. and 3.2. together.

5.3.1. Mathematical concept

The following results are based on the work of Luati and Proietti (2011) for special cases of polynomial regression where an equivalence between diagonally weighted Least-Squares (DWLS) and the generalized Least-Squares (GLS) estimator was proven. The theorem on which the equivalence is based states that DWLS and GLS estimators are equivalent if and only if the $\left(n \times u\right)$ $\mathbf{A}$ matrix $\left(n > u\right)$ can be decomposed as

$$\mathbf{A} = \mathbf{V}^{*}\mathbf{M} \tag{20}$$

where the *u* columns of $\mathbf{V}^{*}$ are eigenvectors of $\mathbf{W}_{FULLY}\mathbf{W}_{EQUI}^{-1}$ and $\mathbf{M}$ is a non-singular matrix. $\mathbf{W}_{FULLY}, \mathbf{W}_{EQUI}$ are the fully populated and equivalent matrices respectively. In the following the model corresponding to the equivalent diagonal $\mathbf{W}_{EQUI}$ is called EQUI.

For the mean estimator case, the necessary and sufficient condition for the equivalence between GLS and DWLS can be simplified so that each element of the diagonal matrix $\mathbf{W}_{EQUI}^{-1}$, is the sum of the row elements of the inverse of the fully populated covariance matrix $\mathbf{W}_{FULLY}^{-1}$ (Luati and Proietti 2011). This result was extended empirically for GNSS positioning and the corresponding procedure to compute the equivalent matrix is resumed in Fig. 10 following Kermarrec and Schön (2016).

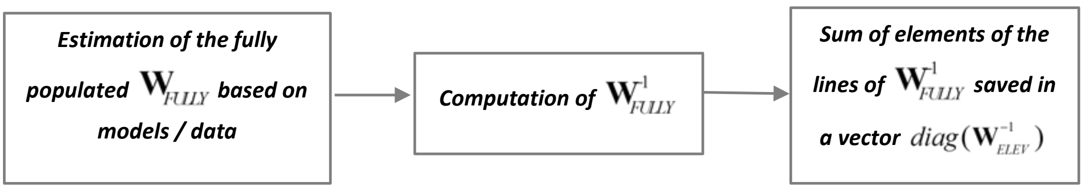


Fig.10 Computation of the equivalent diagonal matrix

The equivalence holds true for the estimates of the parameters and the apriori cofactor matrices of the estimates. It was extended to GPS positioning strategies due to the slowly varying GPS constellation, i.e. the values of the columns of the design matrix sorted per satellite vary nearly linearly.

5.3.2. How to take advantage of the equivalent diagonal model

The two main weaknesses of this proposal are the need to invert a fully populated VCM, as well as the underestimated aposteriori variance factor when the equivalent diagonal model is used in a Least-Squares adjustment. In this section, we briefly suggest solutions for dealing with this inconvenience.

- Inverse of the VCM

In Kermarrec and Schön (2016), it was proposed to invert the VCM outside the software package. A second possibility makes use of the non-orthogonality of the Matérn parameters of the covariance function (Eq. (19)), in forcing the MLE to correspond to an AR(1) process. In that case, the inverse of the VCM is explicitly known and no inversion has to be carried out. It therefore becomes an attractive alternative to fully populated VCM for undifferenced positioning, particularly when used in a Kalman Filter. However, the price to pay is the suboptimal sharp decrease of the covariance function near the origin and thus results closer to the ELEV model. Moreover, when making use of the equivalence between the stochastic model and an augmented functional model, this approximation is strong and should be avoided.

- Aposteriori variance factor

The second weakness of this proposal consists in the systematic underestimation of the aposteriori variance factor. As it was shown to be a key parameter in deciding whether or not a given stochastic model is optimal, this question has to be addressed.

Two solutions to this issue can be proposed: the first one is simple and considers that if correlations are present, Eq. (19) is an adequate model. It can thus be adopted with high confidence. Eventually and in order to prevent from strong biases that occur if the correlation length is misspecified, taking systematically a slightly larger α value than the MLE estimation or the atmospheric correlation length (α between 0.008 and 0.01s^{-1}) prevents from resulting biases and eventually a wrong fixing of the ambiguities to an integer. Another strategy is to adapt the apriori variance factor to $\sigma_0 = E\left(\hat{\sigma}_{\hat{\mathbf{W}}\mathbf{equi}}\right)$, i.e. the mean of the aposteriori variance factor over all batches following the proposal of Sec. 4.4.1. for the ELEV model. This way, solutions computed with wrongly estimated correlation parameters for a given batch can be excluded.

An empirical correction of the aposteriori variance factor is nearly impossible to obtain. Indeed, this could only be done by means of Monte Carlo simulations and necessitates that all combinations of smoothnesses and correlation lengths are taken into account for each geometry. This is computationally demanding and far from easy. As all these solutions remain unsatisfactory, a “diagonal correlation model” was proposed as a better alternative to face this challenge.

5.4. An alternative to the equivalent diagonal model

5.4.1. The Diagonal Correlation Model (DCM)

Detailed explanations and derivations of this model are developed in Kermarrec and Schön (2017d). In order to derive the alternative “diagonal correlation model”, called DCM, the AR(1) model for which the inverse of the VCM is known was used for didactic reasons. In that particular case, and following the results for the equivalent diagonal model, the FULLY VCM for one satellite can be reduced to a diagonal matrix with elements:

- First and last epoch at which satellite *i* is present: $\frac{1}{\sin\left(El_i\left(t\right)\right)^2}\left(\frac{1-\rho_{AR}{}^2}{1-\rho_{AR}}\right) = \frac{1}{\sin\left(El_i\left(t\right)\right)^2}\left(1+\rho_{AR}\right)$
- All other diagonal values: $\frac{1}{\sin\left(El_i\left(t\right)\right)^2}\left(\frac{1-\rho_{AR}{}^2}{\left(1-\rho_{AR}\right)^2}\right) = \frac{1}{\sin\left(El_i\left(t\right)\right)^2}\left(\frac{1+\rho_{AR}}{1-\rho_{AR}}\right)$ (21)

assuming slow variations of the satellite elevation, i.e. allowing a factorization of the elevation-dependent factor.

ρ_{AR} is the correlation factor as defined in Kermarrec and Schön (2016). Thus it can be seen that taking correlation into account is similar to giving a variance higher than 1 for satellites at 90° elevation. This is the reason why the aposteriori variance factor using the equivalent model is not comparable with the one found with commonly used functions, i.e. it is always underestimated. Trying to find an alternative variance model that would have approximately the same effects as the equivalent model without this weakness, the exponential variance function - called DCM for Diagonal Correlation Model - was found to be an interesting alternative. This model is defined as:

$$\sigma_{DCM}^{2}(t)=\sigma_{0}^{2}\left[(1-\delta_{DCM})+\delta_{DCM}\frac{\exp^{-El_i(t)/\gamma}}{\exp^{-90/\gamma}}\right]. \quad (22)$$

where γ [°] is a positive factor called exponential factor. The cofactor is scaled to 1 for a satellite at 90 ° elevation, corresponding to the ELEV model, El_i is the elevation of the satellite i at time *t*. $\delta_{DCM}\in[0,1]$ is a factor that allows to take non elevation-dependent noise into account.

This model provides an interesting flexibility through the exponential factor which can be adapted freely. Its main property leads to the fact that it is always possible to find for a fix ρ_{AR} a value of γ so that the mean of both functions DCM and EQUI over all elevations is similar, i.e. the DCM can be considered as a "hidden" correlation model. δ_{DCM} acts as adding a scaled identity matrix (i.e. a white noise VCM) to the equivalent matrix.

In Fig. 11, the three models DCM, EQUI and ELEV are compared for a particular batch with 8 satellites observed during 100 epochs a 30s corresponding to the case of Fig. 2. It can be seen that the blue line corresponding to the DCM model with an exponential factor of 15° is similar to the red one, which represents an AR(1) equivalent diagonal model with a correlation factor of 0.8. A value of $\delta_{DCM}=0.3$ was used.

Comments

This proposal is not a usual variance model and should not be directly compared with the models proposed in chapter 2. Although tempting, a direct comparison of the obtained variance plotted versus elevation with the commonly used variance models is not meaningful. The DCM should be described and considered as a diagonal correlation model. Thus, it can be used in the presence or not of correlations by adapting the parameters when fully populated matrices cannot be computed. Exemplarily, increasing the value of γ is similar to neglecting more and more correlations.

The DCM model is moreover not trying to fit the EQUI model exactly. Much more, it proposes a way to correct its main weakness and is thus an alternative without claiming to replace it. Its main advantage comes from the possibility to use test statistics such as the overall model test which is not valid for the EQUI model. Moreover, the variability of the exponential factor allows to model different correlation lengths.

The weakness of the proposal is identified for shorter observation sessions (<30 minutes) where the FULLY model that whitened observations remains a better alternative than the DCM. It is thus preferable to use the DCM as soon as the geometry is enough averaged due to the weighting of high elevation satellites that may be otherwise exaggerated (Kermarrec and Schön 2017d).

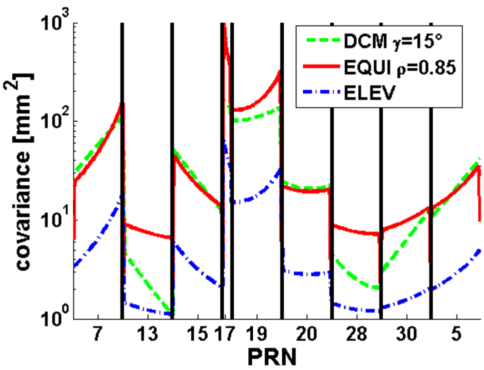


Fig.11 Exponential variance model versus equivalent model. Comparison between EQUI, DCM and ELEV models

5.4.2. Sensitivity of the LS solution to the parameters of the DCM

In order to assess how the exponential factor influences the LS results, simulations similar to those explained in Sec. 4.3. were carried out. The results, presented in Kermarrec and Schön (2017d) highlighted that values of the exponential factor smaller than 12° lead to a strong underestimation of the aposteriori variance factor, i.e. a negative bias compared with the correct value. Taking larger values of γ than the reference value into account were shown to be less risky in terms of loss of efficiency, leading moreover to a positive bias of the aposteriori variance factor. Thus, the corresponding batches of observations can be easily excluded from the solution using the overall model test.

These results were confirmed by a case study with unpreprocessed observations following the methodology of Sec. 4.1.3. However, as for the FULLY model, a good knowledge of the apriori variance factor is necessary to avoid an underestimation which may lead to a wrongly improved 3Drms and high loss of efficiency.

The DCM model was shown to provide more realistic test statistics and estimates with respect to the ELEV model. Moreover, comparable results with the FULLY model for the 3Drms and aposteriori variance factor were empirically found for $\gamma = 15°$ which correspond to a correlation length close to the "atmospheric-like correlation length", i.e. $\rho_{AR} = 0.85$ with an AR(1) model. Thus, the potential of using the DCM to take correlations into account was confirmed. Moreover, choosing values of γ close to 30-40° was identified as corresponding to neglecting correlations, i.e. 3Drms differences compared with the ELEV model of maximum 2-3mm for a 66 km baseline and submm level for short baselines. These results held true for long sessions of observations (>1 hour) where averaging of both the geometry and the data quality happen and are corresponding to Luo (2012).

5.4.3. A proposal to fix the parameters of the DCM

A simple proposal to fix the exponential factor could be derived in Kermarrec and Schön (2017d) based on the results of the simulations and the case study. An explanation in which cases the DCM should be replaced by the correlation model was further proposed. The results are summarized in Fig.12. In a first step, the correlation coefficient ρ_{AR} of the double differences is estimated for a given satellite at mean elevation (40-60°) using a simple routine. A noise factor is either empirically fixed or estimated from the observations. The corresponding γ can be derived by plotting the mean of the equivalent variance over all elevations versus ρ_{AR} and the mean of the DCM versus γ for a given noise factor. For the case $\gamma < \gamma_{\min} = 12°$ which was identified as critical in terms of loss of efficiency in case of misspecification, it is advised to fix ρ_{AR} to 0.85 to account for atmospheric-like correlations, i.e. a correlation length of 600 s (Kermarrec and Schön 2017a). Independent of the baseline length, the DCM should be used with care for short batches when the geometry is not enough averaged. In such cases, a high sensitivity of the 3Drms to variations of γ was identified and the FULLY model should be preferred. Although from the results of simulations not mandatory to have a trustworthy LS solution, it always remains possible to complicate the estimation of the exponential parameter by making it satellite dependent.

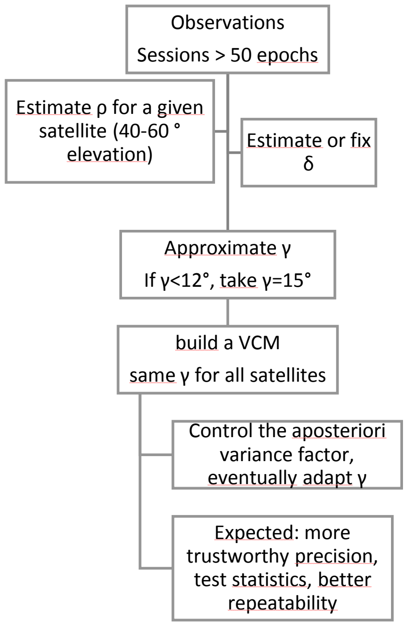


Fig.12 Estimation of the parameters of the DCM model: summary (relative positioning scenario)

5.5. Concluding remarks

Neglecting correlations leads to an overoptimistic precision of the LS solution, i.e. a less efficient estimator. If for long sessions of observations no main differences are expected at the parameters level because of the unbiasedness of the LS estimator, for shorter sessions and longer baselines, correlations should not be disregarded. Indeed, besides the fact that more ambiguities can be fixed to an integer with a higher confidence, the float solution contains less error. In order to simplify the computational burden induced with fully populated VCM, an equivalent diagonal model was proposed that summarized the information contained in the FULLY matrices. The shortcomings of the simplification were overcome by approximating the equivalent variance with an exponential function, allowing thus test statistics based on the aposteriori variance factor to be used. A simple procedure to fix the exponential factor showed how the proposed correlation model and the widely used empirical variance models based on an exponential function can be linked together, i.e. the exponential variance function being able to take correlations into consideration by varying accordingly its parameters, principally the exponential factor.

6. Conclusions and outlook

An accurate knowledge of the stochastic model of GNSS observations is of key importance for a trustworthy Least-Squares solution. Any misspecification leads to an incorrect estimation of the precision, and possibly to wrong decisions due to biased test statistics such as e.g. the validation test in the ambiguity validation procedure, the overall model test, or the outlier detection test.

In order to obtain an improved stochastic model of GNSS phase measurements, three main directions were followed in previous studies: (i) better description of the elevation-dependent variance model, i.e. heteroscedasticity, (ii) modelling or empirical estimation of a temporal covariance function to describe the correlations between observations, (iii) iterative procedure for variance-covariance estimation. Such computational demanding strategies were not investigated in this thesis, the emphasis being on stochastic modelling.

The accurate description of the variance has gained a strong interest in the literature as the corresponding VCM of the observations are diagonal and thus easy to invert and handle. Data-based models such as SNR, fuzzy or exponential models were proposed besides simple models based on a sinus mapping function which disregard data quality indicators or station dependent effects. Correlations between measurements are usually not taken into account. On the one hand the existing models - empirical and improvable - only describe approximately the truth of the correlation structure of the observations. On the other hand, the effect of neglecting correlations when observations are correlated remains unstudied. In order to overcome this weakness and adopting a GNSS user perspective, the first objective of this thesis was to develop a physically derived description of the correlation structure of GNSS phase observations with as less computational burden as possible. The underlying intention was to refine in an understandable way the stochastic model by taking correlations into account to lead to a trustworthy Least-Squares solution.

1. *A correlation model for GNSS phase observations*

To this aim, a simplified version of a function describing the correlations due to turbulent tropospheric refractivity fluctuations based on turbulence theory was developed. This new model includes both the heteroscedasticity (i.e. elevation dependency) of the variance as well as a Matérn covariance function. Fully populated covariance matrices describing the observation dependencies can be computed and integrated in the Least-Squares positioning adjustment. Two parameters were shown to play a major role to whiten the observations: the smoothness related to the mean-square differentiability of the covariance function at the origin and the correlation length. Estimated by MLE, these two Matérn parameters are observation-dependent which permits a great variability in the description of the correlation structure. For short batches, it was shown that an empirical fixing is possible to make use of the equivalence between the augmented functional model and the augmented stochastic model and implicitly takes a non-estimated elevation dependent (for instance tropospheric) parameter into account in the LS adjustment.

2. *Impact of correlations on the LS adjustment: sensitivity analysis of the correlation model*

The second objective of this work was to quantify the effect of misspecification of the correlation structure on the trustworthiness of the LS solution. Because of the flexibility of the proposed correlation function, a sensitivity analysis on its parameters could be carried out. The methodology was based on the analysis of the bias of the aposteriori variance factor which could be linked to the loss of efficiency of the estimator by means of simulations. It was shown that conservative parameters corresponding to sparse variance-covariance matrix were preferable than an overestimated smoothness and correlation length in case of uncertainties in the estimation of the model parameters. The correlation model proved its superiority over diagonal variance models when used concretely for real data analysis in a relative positioning scenario. Both a more realistic precision and less biased test statistics, including ambiguity discriminant tests, were obtained. The main improvements were achieved for short batches of observations, when the ambiguities could not be fixed due to poor validation tests. Such cases correspond to the estimation of a “hidden tropospheric parameter”. This concept could be concretely applied.

3. *Simplification of the correlation model*

The third and last objective of this thesis was to propose a way to take correlations into account without additional computation burden. Thus a method for reducing fully populated variance covariance matrices into their diagonal form for GNSS positioning cases was developed. An unexpected side effect of this proposal was the development of an alternative diagonal correlation model to counterbalance the systematic underestimation of the aposteriori variance factor with the equivalent model. Based on an exponential function with freely varied parameters, this new diagonal correlation model has allowed linking commonly used variance models with the correlation function.

The results of this thesis were based on relative positioning scenarios. However, it is not expected that they are strongly influenced or totally different for other positioning strategies. This statement remains to be tested in practice; particularly the impact of taking correlation into account in Precise Point Positioning (PPP) demands further testing. The effect of the float ambiguity solution on the convergence time could be an interesting follow-on research topic, although the use of an ionospheric-free linear combination may decrease the impact of fully populated VCM. The equivalent diagonal model proposed in this thesis as a way to take correlations into account easily could be concretely applied in a Kalman Filter to model the observation noise and thus find applications beyond the scope of GNSS correlation modelling. The extent to which our model can be also used for code or code/phase correlations remains furthermore an open question. For cross correlations e.g., another weighting of the Matérn covariance function would be worth studying.

The model developed in this work allows for modelling elevation-dependent correlations of GNSS phase observations. By means of simulations and case studies, it was shown to give trustworthy positioning results using Least-Squares adjustment, allowing even to model the effect of a non-estimable tropospheric parameter. The proposed description is neither the most general nor the unique one. The topic of stochastic modelling is a lively area of research and subject to many interpretations.

Bibliography

Abramowitz M, Stegun IA (1972) Handbook of mathematical functions. Dover Publications, Incorporate

Aitken AC (1935) On Least-Squares and Linear Combinations of Observations. Proceedings of the Royal Society of Edinburgh 55:42–48

Amiri-Simkooei AR (2007) Least-Squares variance component estimation: Theory and applications. PhD thesis, Delft University of Technology, Delft

Amiri-Simkooei AR, Teunissen PJG, Tiberius C (2009) Application of Least-Squares variance component estimation to GPS observables. Journal of Surveying Engineering 135(4):149-160

Amiri-Simkooei AR, Zangeneh-Nejad F, Asgari J (2013) Least-Squares variance component estimation applied to GPS geometry based observation model. Journal of Surveying Engineering 139(4):176-187

Amiri-Simkooei AR, Jazaeri S, Zangeneh-Nejad F, Asgari J (2016) Role of stochastic model on GPS integer ambiguity resolution success rate. GPS Solutions 20(1):51-61

Aquino M, Monico JFG, Dodson AH, Marques HA et al. (2009) Improving the GNSS positioning model in the presence of ionospheric scintillation. Journal of Geodesy 83(10):953-966

Barnes BJ, Ackroyd N. Cross PA (1998) Stochastic Modelling for Very High Precision Real-Time Kinematic GPS in an Engineering Environment. Proceedings of FIG XXI International Conference, 21-25 July, Brighton, U.K., Commission 6, 61-76

Berk RA (2004) Regression analysis: a constructive critique. Thousand Oaks, CA:SAGE

Blewitt G (1998) GPS Data Processing Methodology: From Theory to Applications. In: Teunissen PJG and Kleusberg A (Eds.) GPS for Geodesy (pp 231-270). 2nd ed. Springer-Verlag Berlin Heidelberg New York

Bischoff W, Heck B, Howind J, Teusch A (2005) A procedure for testing the assumption of homoscedasticity in Least-Squares residuals: a case study of GPS carrier-phase observations. Journal of Geodesy 78(7-8):397–404

Bona P (2000) Precision, cross correlation, and time correlation of GPS phase and code observations. GPS Solutions 4(2):3-13

Braasch MS, Van Dierendock AJ (1999) GPS receiver architectures and measurements. Proceedings of the IEEE 87(1)48-64

Brunner FK, Hartinger H, Troyer L (1999) GPS signal diffraction modelling: the stochastic Sigma-Δ model. Journal of Geodesy 73(5):259-267

Caspary, W.F. (1987). Concepts of Network and Deformation Analysis. The University of New South Wales, School of Surveying, Monograph No. 21, Kensington (Australia)

Caspary WF, Wichmann K (1994) Lineare Modelle: algebraische Grundlagen und statistische Anwendungen. R. Oldenburg Verlag, München

Chatfield C (1989) The analysis of time series: An introduction. 4th ed. Chapman & Hall, London, U.K.

Chiles J-P., Delfiner P (1999) Geostatistics. Modelling Spatial Uncertainty. Wiley New York

Cohen J, Cohen P, West SG, Aiken L S (2003). Applied multiple regression/correlation analysis for the behavioral sciences. 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates.

Craymer MR, Wells DF, Vanicek P, Davlin RL (1990) Specifications for urban GPS surveys. Surveying and Land Information Science 50(4):251-259

Cressie N (1993) Statistics for spatial data. Wiley, New York

Cressie N, Huang C-H (1999) Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions. Journal of the American Statistical Association 94(448):1330-1340

Dach R, Lutz S, Walser P, Fridez P (2015) Bernese GNSS Software Version 5.2. User manual, Astronomical Institute, University of Bern, Bern Open Publishing. DOI: 10.7892/boris.72297, ISBN: 978-3-906813-05-9.

da Silva HA, Camargo P, Joa˜o, Monico JFG, Aquino M, Marques HA, De Franceschi G, Dodson A (2010) Stochastic modelling considering ionospheric scintillation effects on GNSS relative and point positioning. Advances in Space Research 45(9):1113–1121

Diggle PJ, Ribeiro PJ (2007) Model Based Geostatistics. Springer Science+Business Media, New York

El-Rabbany AE-S (1994) The effect of Physical Correlations on the Ambiguity Resolution and Accuracy Estimation in GPS Differential Positioning. Ph.D. thesis, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada

El-Rabbany A, Kleusberg A (2003) Effect of temporal physical correlation on accuracy estimation in GPS relative positioning. Journal of Surveying Engineering 129(1):28-32

Euler HJ, Goad CC (1991) On optimal filtering of GPS dual frequency observations without using orbit information. Bulletin Geodesique 65(2): 130-143

Fuentes M, Smith RL (2001) A new class of nonstationary spatial models. Tech report, North Carolina state University. Institute of Statistics Mimeo Series 2534, Raleigh

Gaglione S, Angrisano A, Pugliano G et al. (2011) A stochastic sigma model for GLONASS satellite pseudorange. Applied Geomatics 3(1):49–57

Gaspari G and Cohn SE (1999) Construction of correlation functions in two and three dimensions Quarterly. Journal of the Royal Meteorological Society 125(554 Part B):723–757

Gelfand AE, Diggle P, Fuentes M, Guttorp P (2010) Handbook of Spatial Statistics. Boca Raton: Chapman & Hall/CRL

Gerdan GP (1995) A Comparison of Four Methods of Weighting Double-Difference Pseudorange Measurements. Trans Tasman Surveyor, Canberra, Australia, 1(1):60-66

Gneiting T, Schlather M (2004) Stochastic models that separate fractal dimension and the Hurst effect. SIAM Review 46(2):269-282

Goad CC (1987) Precise positioning with the GPS, in S. Turner (Ed.): Applied Geodesy, Lecture Notes in Earth Sciences, Springer-Verlag, 12:17-30

Goldenberg AS (1991) A course in econometrics. Havard University Press

Grafarend EW (1974) Optimization of geodetic networks. Bolletino di Geodesia a Science Affini 33(4):351-406

Grafarend EW (1976) Geodetic applications of stochastic processes. Physics of The Earth and Planetary Interiors 12(2):151-179

Grafarend EW, Awange J (2012) Applications of Linear and Nonlinear Models, Springer Verlag, Berlin-Heidelberg-New York

Greene W H (2003). Econometric Analysis, 5th ed. Upper Saddle River, Prentice Hall, NJ

Guttorp P, Gneiting T (2005) On the Whittle-Matérn correlation family. NRCSE Technical Report Series n°80

Groeneveld E (1994) A reparameterization to improve numerical optimization in multivariate REML (co)variance component estimation. Genetics Selection Evolution 26(6):537-545

Hahn M, Van Mierlo J (1987) Abhängigkeit der Ausgleichungsergebnisse von der Genauigkeitsänderung einer Beobachtung. Zeitschrift für Vermessungswesen 3:105-115

Halsig S, Artz T, Leek J, Nothnagel A (2014) VLBI analyses using covariance information from turbulence models, IVS 2014 General Meeting Proceedings, "VGOS: The New VLBI Network", D Behrend, KD Baver, and KL Armstrong (eds.), Science Press (Beijing), 272-276, ISBN 978-7-03-042974-2

Han S (1997) Quality Control Issues Relating to Instantaneous Ambiguity Resolution for Real-Time GPS Kinematic Positioning. Journal of Geodesy 71(7):351-361

Han S, Rizos C (1995) Standardization of the variance covariance matrix for GPS rapid static positioning. Geomatics Research Australasia 62:37-54

Handcock MS, Wallis JR (1994) An Approach to Statistical Spatial-Temporal Modelling of Meteorological Fields. Journal of the American Statistical Association 89(426):368-378

Hartinger H, Brunner FK (1999) Variances of GPS phase observations: The SIGMA- model. GPS Solutions 2(4):35–43

Heck B (1981) Der Einfluß einzelner Beobachtungen auf das Ergebnis einer Ausgleichung und die Suche nach Ausreißern in den Beobachtungen. Allgemeine Vermessung Nachrichten 88:17-34

Heßelbarth A, Wanninger L (2008) Short-term stability of GNSS satellite clocks and its effect on Precise Point Positioning. Proceedings of the ION 2008 Savannah, GA, September 2008, pp. 1855-1863.

Hofmann-Wellenhof B, Lichtenegger H, Collins J (2001) GPS Theory and Practice. Fifth revised edition. Springer Wien New York

Howind J, Kutterer H, Heck B (1999) Impact of temporal correlations on GPS-derived relative point positions. Journal of Geodesy 73(5):246-258

Jansson P, Persson CG (2013) The effect of correlation on uncertainty estimates – with GPS examples. Journal of Geodetic Science 3(2):111–120

Jin X (1996) Theory of Carrier Adjusted DGPS Positioning Approach and Some Experimental Results. Ph.D. thesis, Delft University Press, Delft The Netherlands

Jonkman NF (1998) Integer GPS-ambiguity estimation without the receiver-satellite geometry. Publications of the Delft Geodetic Computing Centre, LGR Series, 18(95)

Joosten P, Teunissen PJG (2001) On the error sensitivity of the GPS ambiguity success rate. Proceedings of KIS2001, June 5-8 2001, pp. 409-414. Banff, Canada University of Calgary

Joseph A (2010) What is the difference between SNR and CN0. Inside GNSS November December 2010 p20-25

Journel AG, Huifbregts CJ (1978) Mining Geostatistics. Academic Press, New York

Kermarrec G, Schön S (2014) On the Matérn covariance family: a proposal for modelling temporal correlations based on turbulence theory. Journal of Geodesy 88(11):1061-1079

Kermarrec G, Schön S (2016) Taking correlation into account with a diagonal covariance matrix. Journal of Geodesy 90(9):793-805

Kermarrec G, Schön S (2017a) Apriori fully populated covariance matrices in Least-Squares adjustment – case study: GPS relative positioning. Journal of Geodesy 91(5):465-484

Kermarrec G, Schön S, Kreinovich V (2017b) Possible explanation of empirical values of the Matérn smoothness parameter for the temporal covariance of GPS measurements. Applied Mathematical Science 11(35):1733-1737

Kermarrec G, Schön S (2017c) On modelling GPS phase correlations: a parametric model. Acta Geophysica et Geodaetica https://doi.org/10.1007/s40328-017-0209-5

Kermarrec G, Schön S (2017d) Taking correlations into account: a diagonal correlation model. GPS solution 21(4):1895-1906

Kermarrec G, Schön S (2017e) Fully populated VCM or the hidden parameter. Journal of Geodetic Science 7(1):151-161

Koch KR (1986) Maximum likelihood estimate of variance components. Bulletin Géodésique 60(4):329–338

Koch KR (1999) Parameter Estimation and Hypothesis Testing in Linear Models, Springer, Berlin

Koivunen AC, Kostinski AB (1999) The feasibility of Data Whitening to Improve Performance of Weather Radar. American Meteorological Society 38(6):741-749

Krarup T, Juhl J, Kubik K (1980) Götterdämmerimg over Least-Squares adjustment. Proceedings of the 14th Congress of the International Society of Photogrammetry, Hamburg, Vol B3:369-378

Kutterer H (1999) On the sensitivity of the results of Least-Squares adjustments concerning the stochastic model. Journal of Geodesy 73(7):350-361

Langley R (1997) GPS receiver system noise. GPS World, 8:40-45

Li B, Shen Y, Lou L. (2011) Efficient Estimation of Variance and Covariance Components: A Case Study for GPS Stochastic Model Evaluation. IEEE Transactions on Geoscience and Remote Sensing 49(1):203-210

Li B, Verhagen S, Teunissen PJG (2014) Robustness of GNSS integer ambiguity resolution in the presence of atmospheric biases. GPS solution 18(2):283–296

Li B, Lou L, Shen Y (2016) GNSS Elevation-Dependent Stochastic Modelling and Its Impacts on the Statistic Testing. Journal of Surveying Engineering 142(2):04015012

Li B (2016) Stochastic modelling of triple-frequency BeiDou signals: estimation, assessment and impact analysis. Journal of Geodesy 90(7)593:610

Linkwitz K., 1961. Über den Einfluss verschiedener Gewichtsannahmen auf das Ausgleichungsergebnis bei bedingten Beobachtungen. Zeitschrift für Vermessung, 86, 179–186, 239–245

Luati A, Proietti T (2011) On the equivalence of the weighted Least-Squares and the generalised Least-Squares estimators, with applications to kernel smoothing. Annals of the Institute of Statistical Mathematics 63(4):673-686

Luo X, Mayer M, Heck B (2011) Verification of ARMA identification for modelling temporal correlations of GNSS observations using the ARMASA toolbox. Studia Geophysica et Geodeatica 55(3):537-556

Luo X, Mayer M, Heck B (2012) Analysing Time Series of GNSS Residuals by Means of ARIMA Processes. In: Sneeuw N., Novák P., Crespi M., Sansò F. (eds) VII Hotine-Marussi Symposium on Mathematical Geodesy. International Association of Geodesy Symposia, vol 137. Springer, Berlin, Heidelberg

Luo X, Mayer M, Heck B, Awange JL (2014) A realistic and easy-to-implement weighting model for GNSS phase observations**.** IEEE Transactions on Geoscience and Remote Sensing 52(10):6110-6118

Loeve M (1963) Probability theory, 3rd ed. Van Nostrand, Princeto

Matérn B (1960) Spatial variation-Stochastic models and their application to some problems in forest surveys and other sampling investigation. Medd. Statens SkogsforskningsInstitute 49(5)

Mayer M (2006) Modellbildung für die Auswertung von GPS-Messungen im Bereich der Antaktischen Halbinsel. Deutsche Geodätische Kommission, NoC597, Verlag der Bayerischen Akademie der Wissenschaften, Munich

Meier S (1981) Planar Geodetic Covariance Functions. Reviews of geophysics and space physics 19(4):673-686

Miller C, O’Keefe K, Gao Y (2012) Time Correlation in GNSS positioning over short baselines. Journal of Surveying Engineering 138(1):17-24

Misra P, Enge P (2012) Global Positioning System. Ganga-Jamuna Press; Revised Second Edition edition

Odijk D, Teunissen PJD (2007) Sensitivity of ADOP to changes in the single-baseline GNSS model. Artificial satellites 42(2)

O'Keefe K, Petovello M, Lachapelle G, Cannon ME (2007) Assessing probability of correct ambiguity resolution in the presence of time-correlated errors. Navigation 53(4)269-282

Özlüdemir MT (2004) The stochastic model of GPS observations. Turkisch Journal of Engineering and Environmental Sciences 28:223-231

Parkinson BW, Spilker JJ (1996) Global positioning system: theory and applications, progress in astronautics and aerodynamics, American Institute of Astronautics, Washington DC

Radovanovic RS (2001) Variance-Covariance Modelling of Carrier Phase Errors for Rigorous Adjustment of Local Area Networks. IAG 2001 Scientific Assembly. Budapest, Hungary, September 2-7, 2001

Radovanovic R, Teskey W, El-Sheimy N (2000) Development of a Precision GPS Deformation Monitoring System. Vol.3: Correlation of Noise and Multipath. Technical Report. University of Calgary

Rao C (1967) Least-Squares theory using an estimated dispersion matrix and its application to measurement of signals. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1:355--372, University of California Press, Berkeley, California

Rao CR (1971) Estimation of variance and covariance components – MINQUE theory. Journal of Multivariate Analysis 1(3):257-275

Rao, CR, Kleffe J (1988). Estimation of Variance Components and Applications. North-Holland, Amsterdam, New York.

Rao CR, Toutenburg H (1999) Linear Models, Least-Squares and Alternatives. 2 ed., Springer-New York

Rasmussen C.E., Williams C. (2006), Gaussian Processes for Machine Learning. the MIT Press, 2006 Massachusetts Institute of Technology

Rothacher M, Springer TA, Schaer S and Beutler G (1998) Processing strategies for regional GPS Networks. Proceedings of the IAG General Assembly 1997 (e. Brunner FK), Rio de Janeiro Brazil 3-9 September 1997, IAG Symposia, Springer Verlag, Berlin Heldelberg New York, 1998

Satirapod C (2001) Improving the accuracy of static GPS positioning with a new stochastic modelling procedure. Proc., 14th International Technical Meeting of the Satellite Division of the U.S. Institute of Navigation, Salt Lake City, Utah

Satirapod C*,* Wang J *(*2000*)* Comparing the quality indicators of GPS carrier phase observations. Geomatics Research Australasia 73:75-92

Satirapod C, Wang J, Rizos C (2002) A simplified MINQUE procedure for the estimation of variance-covariance components of GPS observables. Survey Review 36(286):582-590

Satirapod C, Wang, J, Rizos C (2003) Comparing different GPS data processing techniques for modelling residual systematic errors. Journal of Surveying Engineering 129(4):129–135

Satirapod C, Luansang M (2008) Comparing stochastic models used in GPS precise point positioning technique. Survey Review 40(308):188-194

Schaffrin B (1981) Best Invariant Covariance Component Estimator and its Application to the Generalized Multivariate Adjustment of Heterogeneous Deformation Observations. Bulletin Geodesique, 55(1):73-85.

Schaffrin B (1983) Varianz-Kovarianz-Komponenten-Schätzung bei der Ausgleichung heterogener Wiederholungsmessungen. Deutsche Geodätische Kommission Reihe C, Heft 282, München 1983

Schuh W-D, Krasbutter I, Kargoll B (2014): Korrelierte Messung - was nun? In: Neuner, H. (Hrsg.): Zeitabhängige Messgrößen - Ihre Daten haben (Mehr-)Wert, Deutsche Vermessungswesen-Schriftenreihe 74. Wissner Verlag Augsburg

Schön S, Brunner FK (2008) Atmospheric turbulence theory applied to GPS carrier-phase data. Journal of Geodesy (1):47-57

Schön S, Kermarrec G (2015) Turbulence theory. In Handbook of Geomathematics. Freeden W Zuhair Nashed M, Sonar T (Eds) Springer-Verlag Berlin Heidelberg. pp1297-1348

Schön S, Kermarrec G, Kargoll B, Neumann I, Kosheleva O, Kreinovich V (to appear 2018) Why Student Distributions? Why Matérn's Covariance Model? A Symmetry-Based Explanation. To appear in Proceedings of the Vietnam Econometric Conference, Hanoi, Vietnam, January 16-18, 2018.

Schwieger V (2007) Determination of Synthetic Covariance Matrices – an Application to GPS Monitoring Measurements. Proceedings on 15th European Signal Processing Conference, Poznan, Polen, 03-07.09.2007

Shaw FH, Geyer CJ (1997) Estimation and testing in constrained covariance component models. Biometrika 84(1):95-102

Stein ML (1999) Interpolation of spatial data. Some theory for kriging. Springer New York

Strand ON (1974) Coefficient errors caused by using the wrong covariance matrix in the general linear model. The Annals of Statistics 2(5):935-949

Talbot N (1988) Optimal weighting of GPS carrier phase observations based on the signal-to-noise ratio. Proceedings of the International Symposia on Global Positioning Systems. Queensland, Australia, 17-19 October, ppV.4.1-4.17

Tay S, Marais J Weighting models for GPS Pseudorange observations for land transportation in urban canyons. 6th European Workshop on GNSS Signals and Signal Processing, Dec 2013, Germany. 4p, 2013

Taylor H, Karlin S (1998) An introduction to stochastic modelling. Academic Press, 3rd edition, https://books.google.de/books?id=UtPgVrVthF8C

Teunissen PJG (1995) The Least-Squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. Journal of Geodesy 70(1-2):65-82

Teunissen PJG (1997) A canonical theory for short GPS baselines. Part IV: Precision versus reliability. Journal of Geodesy 71(9):513-525

Teunissen PJG (2000) Testing theory and introduction, Series on Mathematical Geodesy and Positioning. Delft University Press, 2000

Teunissen PJG, Amiri-Simkooei AR (2008) Least-Squares variance component estimation. Journal of Geodesy 82(2):65–82

Teunissen PJG, Jonkman NF, Tiberius CCJM (1998) Weighting GPS Dual Frequency Observations: Bearing the Cross of Cross Correlation. GPS Solutions 2(2):28-37

Teunissen PJG, Kleusberg A (1998) GPS for Geodesy. 2nd ed. Springer-Verlag Berlin Heidelberg New York

Tiberius C, Kenselaar F (2003) Variance Component Estimation and Precise GPS Positioning: Case Study. Journal of Surveying Engineering 129(1):11-18

Tikhonov AN, Goncharsky AV, Stepanov VV, Yagola AG (1995) Numerical Methods for the Solution of Ill-Posed Problems, Springer Netherlands

Vennebusch M, Schön S, Weinbach U (2010) Temporal and spatial stochastic behavior of high-frequency slant tropospheric delays from simulations and real GPS data. Advances in Space Research 47(10):1681–1690

Verhagen S, Li B (2012) Challenges in ambiguity resolution: Biases, weak models, and dimensional curse Proceedings of 6th ESA Workshop on Satellite Navigation Technologies, NAVITEC 2012 December 2012

Verhagen S, Odijk D (2007) Trustworthy Multi-carrier Ambiguity Resolution in the Presence of Multipath. ION GNSS 20th International Technical Meeting of the Satellite Division, 25-28 September 2007, Fort Worth, TX

Vermeer M (1997) The precision of geodetic GPS and one way of improving it. Journal of Geodesy 71(4):240-245

Wang J, Stewart MP, Tsakiri M (1998) Stochastic modelling for GPS static baseline data processing. Journal of Surveying Engineering 124(4):171-181

Wang, J (1999) Modelling and Quality Control for Precise GPS and GLONASS Satellite Positioning, Ph.D. thesis, School of Spatial Sciences, Curtin University of Technology, Perth, Australia, 171pp.

Wang J (2000) Stochastic modelling for RTK GPS/Glonass positioning. Journal of US Institute of Navigation 46(4)297-305

Wang J, Satirapod C, Rizos C (2002) Stochastic assessment of GPS carrier phase measurements for precise static relative positioning. Journal of Geodesy 76(2)95:104

Wang J, Lee HK, Musa T and Rizos C (2005) Online stochastic modelling for network based GPS real-time kinematic positioning. Journal of Global Positioning System 4(1-2):113-119

Watson GS (1955) Serial Correlation in Regression Analysis. Biometrika 42(3-4):327-341

Weisberg S (2005) Applied Linear Regression. Wiley, 3rd edition

Wheelon AD (2001) Electromagnetic Scintillation part I Geometrical optics. Cambridge University Press, Cambridge

White H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica 48(4):817–838.

Wieser A, Brunner FK (2000) An extended weight model for GPS phase observations. Earth Planet Space 52(10):777-782

Wieser A, Brunner FK (2002) SIGMA-*F*: Variances of GPS Observations Determined by a Fuzzy System. In: Ádám J., Schwarz KP. (eds) Vistas for Geodesy in the New Millennium. International Association of Geodesy Symposia, 125:365-370 Springer, Berlin, Heidelberg

Williams MN, Gomez Grajales CA, Kurkiewicz D (2013) Assumptions of Multiple Regression: Correcting Two Missconceptions. Practical Assessment. Research and Evaluation 18(11)

Wolf H (1961) Der Einfluss von Gewichtsänderungen auf die Ausgleichungsergebnisse. Zeitschrift für Vermessungswesen, 86:361-362

Xu P (1991) Least-Squares collocation with incorrect prior information. Zeitschrift für Vermessungswesen 116:266-273

Xu P, Liu Y, Shen Y, Fukuda Y (2007) Estimability analysis of variance and covariance components. Journal of Geodesy 81(9):593-602

Xu P (2013) The effect of incorrect weights on estimating the variance of unit weigth. Studia Geophysica et Geodetica 57(3):339-352

Yaglom AM (1987) Correlation theory of stationary and related random functions I, Springer Verlag New York, Berlin, Heidelberg

Zimmerman DL, Zimmerman MB (1991) A comparison of spatial semivariogram estimators and corresponding ordinary kriging predictors. Technometrics 33(1):77-91

Acknowledgments

My first thank goes to my thesis advisor Prof. Dr.-Ing. Dr.h.c. Bernhard Heck of the Geodetic Institute at Karlsruhe Institute of Technology (KIT). This work would not have been possible without his support as well as interesting discussions on the topic.

I would also like to acknowledge Prof. Dr. Lambert Wanninger of the Geodetic Institute of the Technical University of Dresden as the co-supervisor of this thesis. I am gratefully indebted for his very valuable comments on my researches as well as fruitful and inspired discussions which helped me conclude this dissertation.
I am thankful to Prof. Dr. Stefan Hinz as Decan of the Faculty and member of the doctoral committee for having accepted me as external doctoral candidate at KIT. All other members of the doctoral committee are also warmly thanked for their interest on this work.

At the Institute für Erdmessung in Hanover, special thanks go to Prof. Dr Jürgen Müller for his support and Dr. Tobias Kersten for his technical help.

I am particularly grateful to Prof. Dr. Fritz Brunner who inspired my work on correlations through all the last years. His warm welcome and support in Graz stays unforgettable in our hearts.

Finally, I express my very profound gratitude to my husband for providing me with unfailing support and continuous encouragement thanks to scientific discussions, reading and improving the articles I wrote. This accomplishment would not have been possible without his experience on GNSS.

Appendix

Kermarrec G, Schön S (2014) On the Matérn covariance family: a proposal for modelling temporal correlations based on turbulence theory. Journal of Geodesy 88(11):1061-1079

Kermarrec G, Schön S (2016) Taking correlation in GPS Least-Squares adjustments into account with a diagonal covariance matrix. Journal of Geodesy 90(9):793-805

Kermarrec G, Schön S (2017) Apriori fully populated covariance matrices in Least-Squares adjustment – case study: GPS relative positioning. Journal of Geodesy 91(5):465-484

Kermarrec G, Schön S, Kreinovich V (2017) Possible explanation of empirical values of the Matérn smoothness parameter for the temporal covariance of GPS measurements. Applied Mathematical Science 11(35):1733-1737

Kermarrec G, Schön S (2017) On modelling GPS phase correlations: a parametric model. Acta Geophysica et Geodaetica https://doi.org/10.1007/s40328-017-0209-5

Kermarrec G, Schön S (2017) Taking correlations into account: a diagonal correlation model. GPS solution 21(4):1895-1906

Kermarrec G, Schön S (2017) Fully populated VCM or the hidden parameter. Journal of Geodetic Science 7(1):151-161

J Geod (2014) 88:1061–1079
DOI 10.1007/s00190-014-0743-7

ORIGINAL ARTICLE

On the Mátern covariance family: a proposal for modeling temporal correlations based on turbulence theory

Gaël Kermarrec · Steffen Schön

Received: 7 January 2014 / Accepted: 10 June 2014 / Published online: 2 July 2014

Abstract Current variance models for GPS carrier phases that take correlation due to tropospheric turbulence into account are mathematically difficult to handle due to numerical integrations. In this paper, a new model for temporal correlations of GPS phase measurements based on turbulence theory is proposed that overcomes this issue. Moreover, we show that the obtained model belongs to the Mátern covariance family with a smoothness of 5/6 as well as a correlation time between 125–175 s. For this purpose, the concept of separation distance between two lines-of-sight introduced by Schön and Brunner (J Geod 1:47–57, 2008a) is extended. The approximations made are highlighted as well as the turbulence parameters that should be taken into account in our modeling. Subsequently, fully populated covariance matrices are easily computed and integrated in the weighted least-squares model. Batch solutions of coordinates are derived to show the impact of fully populated covariance matrices on the least-squares adjustments as well as to study the influence of the smoothness and correlation time. Results for a specially designed network with weak multipath are presented by means of the coordinate scatter and the a posteriori coordinate precision. It is shown that the known overestimation of the coordinate precision is significantly reduced and the coordinate scatter slightly improved in the sub-millimeter level compared to solutions obtained with diagonal, elevation-dependent covariance matrices. Even if the variations are small, turbulence-based values for the smoothness and correlation time yield best results for the coordinate scatter.

Keywords GPS · Physical correlations · Temporal correlations · Turbulence theory · Mátern covariance family

1 Introduction

The use of diagonal covariance matrices in which temporal correlations between GPS phase measurements are ignored leads to unreliable positioning results and an overestimation of the precision (El-Rabbany 1994; Wang et al. 2002; Satirapod et al. 2003). Being mainly computed with elevation-dependent models (Euler and Goad 1991), C/N0 models (Hartinger and Brunner 1999; Brunner et al. 1999; Wieser and Brunner 2000), or SNR models (Luo et al. 2011), diagonal covariance matrices are however easy to handle in weighted least-squares models since no computational issue due to matrix inversions occurs (Howind et al. 1999; Beutler et al. 1987). Leading to fully populated variance–covariance matrices (VCM), temporal correlations can depend on receiver type, occur between channels and observation types (Borre and Tiberius 2000) or come from multipath (Radovanovic 2001). However, the main temporal correlations are known to come from the GNSS signal propagation through the atmosphere, considered as a random medium.

Up to now, few propositions have been done to specify time-dependent correlations. El-Rabbany (1994) proposed an empirical modeling based on the study of autocorrelation functions of phase residuals which leads to a simple exponential function with an empirically determined correlation time. Howind et al. (1999) used the results of El-Rabbany to build covariance matrices. They principally showed that the coor-

G. Kermarrec (✉) · S. Schön
Institut für Erdmessung (IfE), Leibniz Universität Hannover, Schneiderberg 50, 30167 Hannover, Germany
e-mail: gael.kermarrec@web.de

S. Schön
e-mail: schoen@ife.uni-hannover.de

dinate estimates are not much improved; nevertheless, the a posteriori accuracy of the least-squares solution is much more realistic. Wang et al. (2002) and Satirapod et al. (2003) used a recursive whitening procedure based on residuals. They highlighted that the determination of ambiguities was strongly improved. In the last few years, some authors have also made used of ARMA processes at the least-squares residuals' level to study temporal correlations of GPS measurements (Luo et al. 2012; Wang et al. 2002).

Based on Kolmogorov turbulence theory and the concept of eddies, Schön and Brunner (2008a) developed the SIGMA-C model. This model, as the Treuhaft and Lanyi (1987) one, is involving a double integration, making its concrete use time consuming. Thus, until yet, there is a lack of a simple, physically derived and easy to handle model for temporal GPS phase correlations.

Thanks to the equations of electromagnetic propagation and geometrical approximations as well as the Kolmogorov turbulence theory, we develop a new covariance model, based on the flexible Mátern covariance family. Using the results of Schön and Brunner (2008a), it is possible to estimate covariances between phase observations for all relevant cases, i.e.

- a given satellite observed at one station with itself,
- one satellite at one station with another one at the same station,
- one satellite at a given station with another one at another station.

Thanks to the specially designed "Seewinkel Network" (Schön and Brunner 2008b), first promising results, both for the quadratic deviation of the computed batch coordinates as well as for the a posteriori variance of the unknowns validate the feasibility and the utility of taking temporal correlations into account with the Mátern covariance family.

The paper is organized as follow: in a first part, a new way to model GPS phase temporal correlations based on models using turbulence (Schön and Brunner 2008a) will be presented. The second part is devoted to the comparison with other models as well as the study of the model parameter dependencies. In a last part, a concrete case based on the Seewinkel Network, designed to study the effect of the tropospheric fluctuations on GPS phase signals (Schön and Brunner 2008b) is presented. The flexibility of the Mátern family will be highlighted as well as the possibility to improve in a manageable way the reliability and standard deviation of least-squares coordinate estimations by taking temporal correlations into account. Finally, an appendix gives necessary mathematical background on the Mátern covariance functions.

2 Physical background: atmospheric transmission and turbulence theory

2.1 Basic concepts—overview

The atmosphere can be considered as a medium varying randomly in time and space (Ishimaru 1997, chapter 17). Thus, GPS satellite signals that propagate through the atmosphere have to be described by statistical methods. The troposphere (a layer between approximatively 0–10 km of altitude) and the ionosphere (85–500 km of altitude) are the most important parts of the atmosphere to be considered for GPS transmission (Wheelon 2001, chapter 2). Frequency-dependent ionospheric effects due to the ionization by solar radiation can mainly be canceled via ionosphere-free linear combination. However, the troposphere being a non-dispersive medium, whether double differencing neither linear combination of observations eliminate its refractive effects on GPS phase measurements.

Turbulence, particularly in the boundary layer between 0–2 km height, is a fundamental phenomena in the troposphere and remains an actual research field. As examples, we cite here the wavelet approach by Khujadze et al. (2013), Farge (1992) or the large Eddy simulation (2008). From the point of view of GNSS signals, turbulence causes variations of the refractive index, that act on phase measurements, causing tropospheric slant delay fluctuations.

The turbulent troposphere can be modelized as a superposition of eddies or "swirl of motion" of different sizes, from the millimeter to the kilometer level depending on the altitude (Stull 2009; Wheelon 2001; Coulman and Vernin 1991); eddies and the surrounding atmosphere having different refractive indices. Thus, with this model in mind, we can motivate sources of correlations for GNSS observations using a geometric optics model: rays that are closer (temporally or spatially) encounter nearly identical eddies and are correlated together. Figure 1 proposes a schematic representation of the troposphere where eddies of different size and energy coexist, from small and isotropic in the boundary layer, the so-called 3D turbulence, to elongated, anisotropic in the loosely called free troposphere (Stull 2009).

In the boundary layer characterized by a high Reynolds number, strong turbulence occurs due to the influence of the Earth and the water vapor content; changes in the refractive index are rapid and eddies are small and isotropic (Coulman and Vernin 1991; Hunt and Morrison 2000). The energy cascade (Kolmogorov 1941) models the transfer and associated breakdown of eddies at a constant rate. Above the boundary layer, in the loosely called free atmosphere ($H > 1{,}000$ m) the turbulence is more 2D-like and the validity of the energy cascade which represents eddies breaking from large to small scales is questionable (Gage 1979).

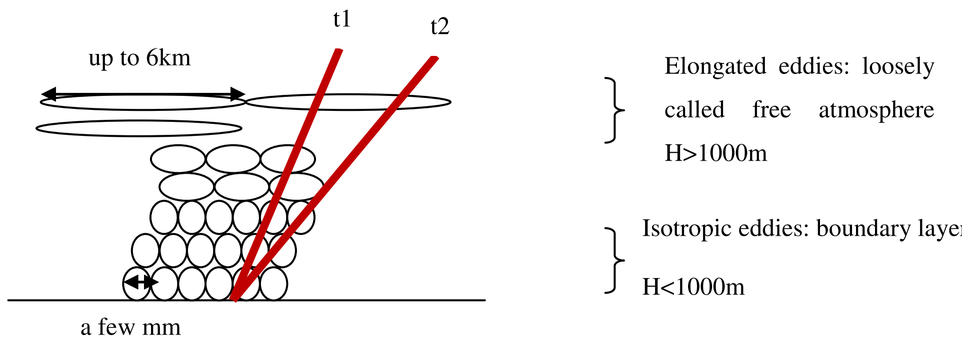


Fig. 1 Schematic representation of the troposphere seen from a satellite signal (adapted from Wheelon (2001, p83) based on measurements of the outer scale)

An important parameter for tropospheric turbulences is the structure constant C_n^2, which is a kind of measure of the intensity of turbulence; see for example Nilsson and Haas (Nilsson and Haas 2010) for a description of how this parameter can be evaluated. A typical profile for microwave C_n^2 in the troposphere (Wheelon 2001, p68) shows a decrease with height from approximately 10^{-14} $\mathrm{m}^{-2/3}$ at 1,000 m altitude to 10^{-17} $\mathrm{m}^{-2/3}$ at 7,000 m. However, the values are remaining in a range of 5.10^{-14}–10^{-15} $\mathrm{m}^{-2/3}$ between 1,000–3,000 m. Thus and following Wheelon (2001, chapter 2), for a GPS ray that propagates through the whole atmosphere, the free troposphere from 1,000 m up to 3,000 m will play a much more important role than the boundary layer below 1,000 m in creating correlations between phase GPS measurements. The intensity of turbulence is large enough and at the same time the reorganization is slower than at a lower altitude making the medium more stable.

This intuitive result can also be explained by considering the weak fluctuation mathematical approximation (Ishimaru 1997, chapter 17). The dielectric constant ε of the troposphere depends on the position $\mathbf{r}$ and time t and is expressed in a first-order approximation: $\varepsilon(\mathbf{r}, t) = n^2(\mathbf{r}, t)$, where n is the refractive index. Under this approximation, the covariance function for phase at a plane $x = L$ can be expressed by means of a filter function (Ishimaru 1997, p352):

$$C_\varphi(L, r) = 2\pi^2 k^2 L \int_0^\infty \kappa \mathrm{d}\kappa\, J_0(\kappa r)\, f_\phi(\kappa)\, \Phi_n(\kappa), \qquad (1)$$

with J_0 the ordinary Bessel function of 0th order, $\Phi_n(\kappa)$ the 3D power spectrum of refractivity fluctuations for an homogenous medium independent of the location, $\kappa = \frac{2\pi}{L}$ the wavenumber, L the scale length and finally k the electromagnetic wavenumber. The function

$$f_\varphi(\kappa) = 1 + \frac{\sin(\kappa^2 L/k)}{\kappa^2 L/k} \qquad (2)$$

can be considered as a filter function of the spectrum (Fig. 2).

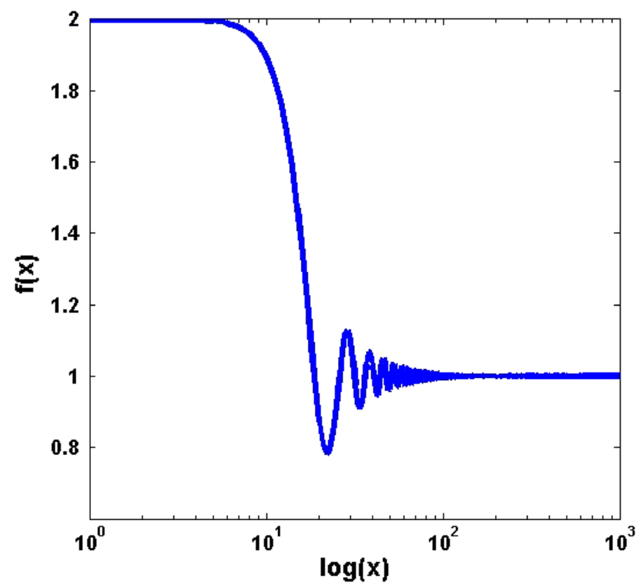


Fig. 2 Filter function $f_\varphi(x)$ versus $\log(x)$. For $x < 1$, the filter function is having nearly constant values close to 2

The region $\kappa < \frac{\sqrt{2\pi}}{\sqrt{\lambda L}}$ is strongly emphasized, meaning that larger eddies are more affecting the phase measurements than smaller ones. Since for GNSS signals $\lambda \approx 20$ cm, we will have $\frac{L_0^2}{\lambda} \approx 10^9 >> H$, where L_0 is the correlation length (in the free troposphere, i.e. the outer scale length $L_0 \approx 6{,}000$ m for horizontal direction, $L_0 \approx 70$ m for vertical direction) and H is approximately the height of the troposphere (5,000–10,000 m).

The covariance function for phase measurement can be further simplified to:

$$C(L, r) = 4\pi^2 k^2 L \int_0^\infty \kappa \mathrm{d}\kappa\, J_0(\kappa r)\, \Phi_n(\kappa). \qquad (3)$$

However, a further integration is not possible without knowledge of the power spectrum of the atmospheric fluctuations, which is a weighting function of the wavenumber κ.

From dimensional analysis, Kolmogorov (1941) has shown that the energy spectrum of turbulence should follow a power law. The related velocity power spectrum as well as the power spectrum of passive scalars such as temperature or refractive index (Monin and Yaglom 1975) can be written as:

$$\Phi_n(\kappa) = \frac{0.033\, C_n^2}{(\kappa_x^2 + \kappa_y^2 + \kappa_z^2)^{11/6}}, \tag{4}$$

where C_n^2 is the structure constant. C_n^2 differs for optical and microwave measurements; microwave being more influenced by water vapor content and optical frequencies by temperature fluctuations. The vector of 3D wavenumbers is $\kappa = [\kappa_x\ \kappa_y\ \kappa_z]^{\mathrm{T}}$. This model is valid in the inertial range for homogeneous and isotropic turbulence, where $\frac{2\pi}{L_0} \le \kappa \le \frac{2\pi}{l_0}$ with L_0, l_0 being the outer and inner scale length of turbulence that bounds the inertial range, respectively.

However, this model yields infinite values for some quantities such as the mean square fluctuations of the refractive index $\langle(\delta n)^2\rangle$. As a consequence, the empirical Von Karman Model is often preferred:

$$\Phi_n(\boldsymbol{\kappa}) = \frac{0.033\, C_n^2}{(\kappa_x^2 + \kappa_y^2 + \kappa_z^2 + \kappa_0^2)^{11/6}}, \quad \kappa_0 = \frac{2\pi}{L_0}. \tag{5}$$

Please refer to Voitsekhovich (1995) or Wheelon (2001, chapter 2) for a presentation of further models such as the Greenwood model or the exponential model.

Although only developed for the inertial range and isotropic turbulence, the Von Karman power law model has shown to be valid beyond these limits, particularly for 2D turbulence (Wheelon 2001; Kraichnan 1974). We will therefore make use of it.

2.2 Anisotropy, inhomogeneity

Isotropy and homogeneity are the main assumptions of the Kolmogorov model. However, GNSS phase measurements are especially affected by the propagation through the free atmosphere where eddies are mainly elongated. Inhomogeneity as well as anisotropy have to be taken into account in the power spectrum model.

Inhomogeneity

The troposphere can be considered in a first approximation as a locally inhomogeneous field with smoothly varying mean characteristics. Inhomogeneity can be expressed by a product of a slowly varying function which describes the spectral distribution of turbulent fluctuations for the whole medium $\Phi_{n,0}(\kappa)$ and a term with faster variations $C_n^2\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2}\right)$ describing the intensity of the fluctuations for the refractive index in a given region of the medium where $\mathbf{r}_1$ and $\mathbf{r}_2$ denote two different position vectors (Tatarskii 1971, p36). Thus, for different regions separated beyond the outer scale length, the power spectrum reads:

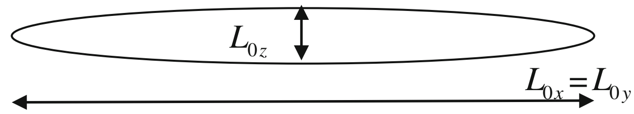


Fig. 3 Elongation of eddy

$$\Phi_n\left(\boldsymbol{\kappa}, \frac{\mathbf{r}_1+\mathbf{r}_2}{2}\right) = C_n^2\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2}\right)\Phi_{n,0}(\boldsymbol{\kappa}). \tag{6}$$

Anisotropy

As seen in Fig.1, horizontal elongated eddies of the free atmosphere are impacting the phase measurements. Following Wheelon (2001), the wavenumber spectrum can be expressed by:

$$\Phi_n(\boldsymbol{\kappa}) = abc\Phi_n\left(\sqrt{a^2\kappa_x^2 + b^2\kappa_y^2 + c^2\kappa_z^2}\right), \tag{7}$$

where a,b,c are stretching factors describing the elongation of the eddies in the three dimensions, i.e. $L_{0x} = aL_0$, $L_{0y} = bL_0$, $L_{0z} = cL_0$. $L_{0x} = L_{0y}$ are the horizontal elongations and L_{0z} the vertical one.

Figure 3 is a schematic representation of an horizontal elongated eddy corresponding to the layered turbulence above the planetary boundary layer, in the loosely called free troposphere where the effects of the Earth's surface friction on the air motion are becoming negligible. The outer scale length's value is between 6,000 and 10,000 m (please refer to Wheelon 2001, chapter 2 for results of experiments of the Global Atmospheric Sampling Program). The vertical elongation is 100 times smaller and between 10 and 70 m, the higher values being measured at an altitude around 1,000 m (five years campaign New Mexico, Wheelon 2001 p83).

2.3 Taylor's frozen hypothesis

To access to the temporal covariance of phase measurements $C(t) = \langle\varphi(t), \varphi(t+\tau)\rangle$, τ being a time increment, the Taylor's frozen hypothesis (Taylor 1938) is assumed. It postulates that the phase covariance between two instants separated by τ is identical to the spatial covariance of phase measurement at two stations separated by a vector $\mathbf{r} = \mathbf{u}\tau$, where $\mathbf{u}$ is the mean wind vector. Thus, the atmosphere is said to be "frozen" during the measurement, eddies are only moved by the mean wind $\mathbf{u}$. A value of $\|\mathbf{u}\| = 8\,\mathrm{ms}^{-1}$ seems relevant (Stull 2009).

This approximation performs better as the mean wind speed increases. In the case of GPS phase covariance at synoptic scales, the geostrophic wind which blows parallel to the isobar can be chosen (Wheelon 2001, chapter 6). Thus, the covariance (Eq. 3) can be written as $C(\tau) = 4\pi^2k^2L\int_0^\infty \kappa\mathrm{d}\kappa\, J_0(\kappa\tau u)\,\Phi_n(\kappa)$ where it is assumed that the wind vector does not change with time and position and the satellite geometry varies slowly with time.

3 Formulation of the new covariance model

3.1 Formulation of the temporal covariance via spectral density

Following Wheelon (2001) and using the previous approximations as well as the von Karman spectrum for the refractivity fluctuations expressed in stretched coordinates, the spectrum of phase measurements $W_{\varphi_i}(\omega)$ can be obtained by integrating along the lines-of-sight:

$$W_{\varphi_i}(\omega) = 2.192H\frac{k^2C_n^2ca^{-5/3}u^{5/3}}{\sin^2(El_i)}\frac{1}{\left[\omega^2+\left(\frac{\kappa_0 u}{a}\right)^2\right]^{4/3}}, \quad (8)$$

where El_i is the elevation of the satellite i, H the tropospheric deep or height, $a = b$ and $c$the horizontal and vertical stretched parameters for the outer scale length, ω is the angular frequency and u the wind velocity.

A rational spectral density can be recognized. The considered process being 1D (only a time dependency), the previous formula can be reformulated, introducing the dimensionality $D = 1$:

$$W_{\varphi}(\omega) = 2.192H\frac{k^2C_n^2ca^{-5/3}u^{5/3}}{\sin^2(\mathrm{El})}\frac{1}{[\omega^2+\alpha^2]^{5/6+1/2}}, \quad (9)$$

where $\alpha = \frac{\kappa_0 u}{a}$ and $\nu = \frac{5}{6}\ \left(\nu + \frac{D}{2} = \frac{4}{3}, D = 1\right)$. Using the equivalence of Appendix A (Rasmussen and Williams 2006) and the Wiener–Khinchin theorem, the covariance is a so-called Mátern covariance function which reads:

$$C_{\varphi_i}(t, t+\tau) = 0.7772\frac{k^2HC_n^2c\kappa_0^{-5/3}}{\sin(\mathrm{El}_i(t))\sin(\mathrm{El}_i(t+\tau))} \times \left(\frac{\kappa_0 u\tau}{a}\right)^{5/6}K_{5/6}\left(\frac{\kappa_0 u\tau}{a}\right), \quad (10)$$

with a smoothness parameter of $\nu = 5/6$ and a Mátern correlation time $1/\alpha$, $\alpha = \frac{\kappa_0 u}{a}$. Equation (10) is a closed formula and thus free of integrals. Moreover, the identification as Mátern covariance opens up new interpretations (cf. Appendix A). Due to the use of the von Karman power spectrum, the continuity at the origin is not given. A formulation of the variance can be found in Wheelon (2001, p164) or using the limit of the Bessel function (Abramowitz and Segun 1972):

$$\lim_{\tau->0} K_{5/6}(\tau) = \frac{1}{2}\Gamma\left(\frac{5}{6}\right)\left(\frac{2}{\tau}\right)^{\frac{5}{6}}\left(1-\tau^{\frac{5}{3}}\frac{\Gamma\left(\frac{1}{6}\right)}{\Gamma\left(\frac{11}{6}\right)}\ldots\right)$$

yielding

$$C_{\varphi_i}(t, t) = 0.782\frac{k^2HC_n^2c\kappa_0^{-5/3}}{\sin^2(El_i(t))}. \quad (11)$$

Approximations

Until now, following approximations were made to express the covariance structure of GPS phase signals propagating through the turbulent free troposphere:

- No dependency of the structure constant with height is taken into account. The constant value of $5.10^{-14}\ \mathrm{m}^{-2/3}$ was taken. Following Treuhaft and Lany (1987), for simplification and homogeneity, a constant value should be enough as proposed by Schön and Brunner (2008a), Wheelon (2001). If more accuracy is needed, profiles of the structure constant for microwave could be used as well as a layered model (Kleijer 2004; Gradinarsky 2002).
- No dependency of the outer scale length with height is assumed. It is very difficult to access the structure of the horizontal outer scale length. Range of values between 6,000 and 10,000 m has been experimentally determined in the free atmosphere, cf. Wheelon (2001).
- Following Wheelon (2001, chapter 4), only the turbulence in the free atmosphere and not in the boundary layer is taken into account. However, by changing the outer scale length to $L_0 = 100-600$ m and the value of H and the structure constant C_n^2 accordingly, our model can be extended to the boundary layer. In this case, the Mátern correlation time is more than 10 times smaller than for the free atmosphere case.
- The wind speed is taken constant which is a good approximation above 1,000 m under normal meteorological condition; cf. Wheelon (Wheelon 2001, chapter 6) for wind fluctuations.
- Taylor's Frozen Hypothesis is assumed. This approximation should be valid as long as the GPS lines-of-sight are not too far from each other. Thus, the model for computing the phase covariance between two different satellites should be carefully used if the distance between two rays at approximately $H = 1{,}000-2{,}000\,m$ is larger than 10 km (the maximal outer scale length in the atmosphere, Wheelon 2001) since the turbulence at such scales is not a priori having a Kolmogorov behavior.
- This model assumes a flat atmosphere using a $\frac{1}{\sin(\mathrm{El})}$ mapping which is a good approximation for high elevation angles. However, for lower elevation angles (below 10°), more elaborated tropospheric mapping functions should be used for more accuracy (Böhm and Schuh 2013).

3.2 Dependencies and parameter sensitivity

3.2.1 General remarks

Several tropospheric parameters are involved in the proposed model. Thus, in a next step, the physical dependencies are studied, leading to a proposal for modeling the phase correlations between GPS signals of two different satellites.

Four parameters coming from the turbulence theory have a scaling effect on the variance and covariance: the tropospheric height H, the structure constant C_n^2, the vertical elongation parameter $c$of the stretched coordinates, and the

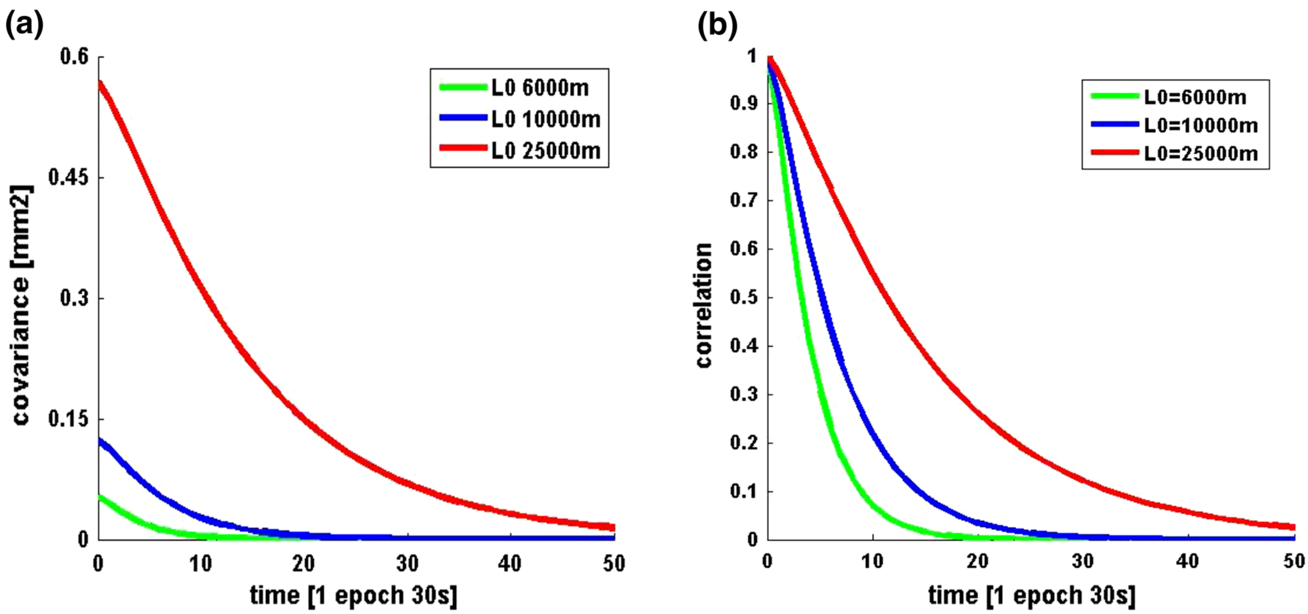


Fig. 4 Covariance (**a**) and correlation function (**b**) by varying the outer scale length from 6,000 to 25,000 m. A satellite at low elevation (15°) was taken for the simulation as well as $c = 0.01$, $C_n^2 = 5.10^{-14}\ \mathrm{m}^{-2/3}$, $u = 8\ \mathrm{ms}^{-1}$

outer scale length $\kappa_0 = \frac{2\pi}{L_0}$. In addition, both the wind velocity and the outer scale length are also acting on the Mátern correlation time (CT) via $\alpha = \frac{\kappa_0 u}{a}[\mathrm{s}^{-1}]$, which we reduce to $\alpha = \kappa_0 u$ by taking $a = 1$ and an horizontal elongation of $L_0 = 6{,}000\,\mathrm{m}$. Thus, since $a/c = 1/c > 100$ can be assumed in the free troposphere, c will be set to 0.01.

If a geostropic wind value between 8–10 $\mathrm{ms}^{-1}$is taken as well as an outer scale length of 6,000–10,000 m, the typical range of values of α should be $0.005-0.01\ \mathrm{s}^{-1}$, i.e. the correlation time as defined in El-Rabbany (1994) is between 100–200 s.

In the following, the impact of different values of the parameter (outer scale length and wind velocity) on the covariance is exemplary studied.

3.2.2 Changing the outer scale length L_0

In Fig. 4, the outer scale length parameter was changed from 25 km (very elongated eddies, synoptic scales) to 6 km, which should be considered as a reference value (Wheelon 2001) for the GPS temporal phase covariance. For a better comparison, the same value of the structure constant was used for the three cases.

A large value of the outer scale length (red 25 km) leads to a longer Mátern correlation time (CT) and larger values of the covariance. A standard value of 6 km results in a shorter Mátern CT as well as smaller covariance. Acting on both the correlation time and the variance, the outer scale length determines the behavior of the spectrum at low frequencies.

3.2.3 Changing the wind velocity

In Fig. 5, we changed the wind velocity from 4 to 10 ms^{-1}. The variance does not depend on this parameter, thus only the Mátern correlation time $1/\alpha[\mathrm{s}]$ will change. For large values of the wind speed, the Mátern CT is shorter than for small values (here 4 ms^{-1}) for which the spectral energy is shifted at low frequencies. Values between 8 and 10 ms^{-1} should be physically most relevant (blue and green line) corresponding to the approximate value of the geostropic wind (Stull 2009). Moreover, working under Taylor's hypothesis, higher values of the wind speed are preferable.

3.3 Extended formulation of the covariance—case two satellites—one or two stations

We propose to extend the "one satellite-one station" covariance model to the case when the covariance between two phase measurements of different satellites i and j at different stations A and B is needed. The case "two satellites, one station" is given by taking A = B in the following formula. Our development is based on the observations that the correlations times of GPS phase measurements are typically 3–7 min (Schön and Kutterer 2006). During that time, the satellite geometry is only little varying and a decomposition into a temporary fixed satellite geometry at a time t and temporal variations seems adequate. Consequently, both temporal and spatial correlations are taken into account.

At a given height $H = 1{,}000\,m$ (tropospheric height), and for one epoch when satellite i and j are present, the paths of

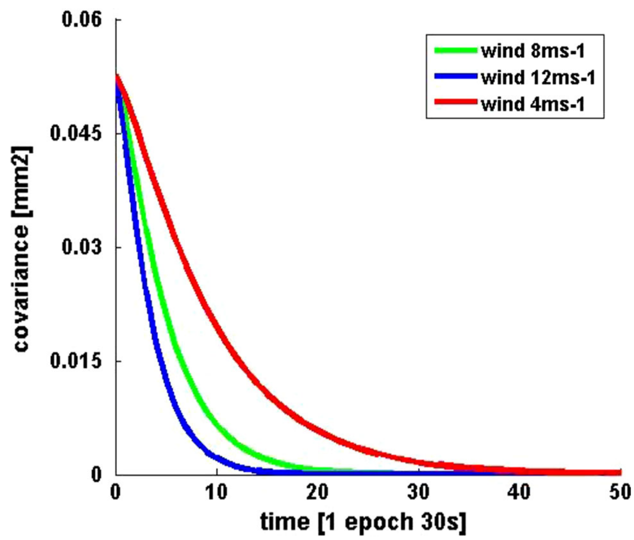


Fig. 5 Covariance functions by varying the wind velocity from 4 to 12 ms^{-1}. $c = 0.01$, $C_n^2 = 5.10^{-14}$ m$^{-2/3}$, $L_0 = 6{,}000$ m and a satellite at low elevation (15°) were taken for the simulation

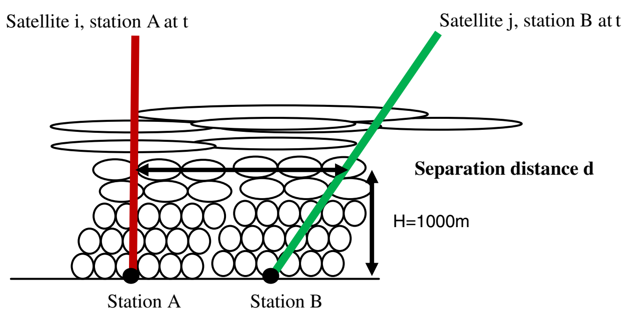


Fig. 6 Simplified version of the concept of separation distance at t (Schön and Brunner 2008a)

satellite i and j are separated by a distance called the "separation distance" $d_{t,H=1{,}000\,\mathrm{m}}$ which depends on the geometry (elevation, azimuth) of the two satellites at a given time (Schön and Brunner 2008a). The following figure (Fig. 6) illustrates in a simple way the concept of separation distance. More details can be found in Schön and Brunner (2008a).

No influence of the wind vector with time on the separation distance (Schön and Brunner 2007) is here taken into account since the separation distance is computed at one epoch. A special case occurs for separation distance larger than the maximum outer scale length of the troposphere (approximately 10 km, Wheelon 2001, p82). Indeed, neither the Kolmogorov law nor the Taylor's frozen hypothesis should be applied. For this particular case, correlations due to tropospheric turbulences can be neglected since the rays are considered too far away from each other. Thus, the procedure can be summarized as follow:
First step: Compute the separation distance $d_{t,H=1{,}000\text{ m}}$
Second step: If $d_{t,H=1{,}000\,\mathrm{m}} < 6{,}000$ m, set $L_0 = 6{,}000$ m.

If $6{,}000\text{ m} < d_{t,H=1{,}000\text{ m}} < 10{,}000$ m, the outer scale length is taken to $L_0 = d_{t,H=1{,}000\text{ m}}$. The vertical elongation is 100 times smaller than the horizontal one (Wheelon 2001, chapter 2). Thus, for high values of the separation distance, the correlations between microwave phase measurements will come from longer eddies, that appear at higher altitudes in the free troposphere. The structure constant should be taken accordingly smaller than for the case "one satellite". Following the previous mentioned structure constant profile (Wheelon 2001, chapter 2) we propose to take approximately the structure constant by a factor 10 lower as for the previous case $C_n^2 = 5.10^{-15}\,\mathrm{m}^{-2/3}$.

As a consequence, the covariance between two phase measurements of two different satellites (i, j) is given by:

$$\begin{cases} \text{if} \quad d < 6{,}000\text{ m} : \langle \varphi_A^i(t), \varphi_B^j(t+\tau)\rangle \\ \quad = \ldots 0.7772 \frac{k^2 H C_n^2 c}{\sin(\mathrm{El}_i^A(t))\sin(\mathrm{El}_j^B(t+\tau))}\left(\frac{2\pi}{L_0}\right)^{-5/3} \\ \quad \left(\frac{2\pi u\tau}{L_0 a}\right)^{5/6} K_{5/6}\left(\frac{2\pi u\tau}{L_0 a}\right) \\ \text{if} \quad 6{,}000 < d < 10{,}000\text{m} : \langle \varphi_A^i(t), \varphi_B^j(t+\tau)\rangle \\ \quad = \ldots 0.7772 \frac{k^2 H C_n^2 c}{\sin(\mathrm{El}_i^A(t))\sin(\mathrm{El}_j^B(t+\tau))}\left(\frac{2\pi}{d_{t,H=1{,}000\,\mathrm{m}}}\right)^{-5/3} \\ \quad \left(\frac{2\pi u\tau}{d_{t,H=1{,}000\,\mathrm{m}} a}\right)^{5/6} K_{5/6}\left(\frac{2\pi u\tau}{d_{t,H=1{,}000\,\mathrm{m}} a}\right) \\ \text{if} \quad d > 10{,}000\text{ m} : \langle \varphi_A^i(t), \varphi_B^j(t+\tau)\rangle = 0 \end{cases} \tag{12}$$

As previously, the wind speed u is taken to its geostropic value. For longer observation spans, different time t may be used.

As the separation distance is replacing the outer scale length in this new model, the behavior of the covariance versus time is the same as in Fig. 4. Typical values of the separation distance at $H = 1{,}000$ m depends on the satellite geometry (azimuth and elevation) and are between a few hundred meters to 10,000 m or more.

The rational spectral density of phase measurements allowed us to propose a new and simple model for the computation of GPS phase covariance based on the Mátern covariance family. Inhomogeneity as well as anisotropy and nonstationarity were taken into account and reflect the physical effects of the atmosphere on GPS signals. Moreover, by allowing a great flexibility through the smoothness and the correlation time parameters which can be changed if needed and estimated via likelihood estimation (Stein 1999; Handcock and Wallis 1994), such covariance functions are quite promising for modeling temporal correlations, not only tropospheric correlations as proposed in the paper but also multipath or receiver-related internal correlations.

Amplitudes fluctuations, which are affected by small eddies, do not allow the geometrical–optical approximations valid for microwave signals and diffraction theory has to be used (Ishimaru 1997, chapter 17). As a consequence, the proposed model is not adequate for modeling the tropospheric temporal correlation of amplitude measurements.

3.4 Comparison with other covariance models

Other models for GPS temporal correlations have been proposed in the past such as the exponential model, the Treuhaft and Lanyi model (1987) or the Schön and Brunner (2008a) model. In this section, the differences and similarities of these models are shortly described.

3.4.1 Exponential model

Proposed by El-Rabbany (1994), the exponential model was concretely used exemplary by Howind et al. (1999) and Radovanovic (2001) to describe temporal correlations. A direct link to tropospheric correlations is not explicitly given. However, it is a special case of the Mátern family with a smoothness parameter of $1/2$ (see Appendix A for more details). Consequently, this model is close to our proposal with $\nu = 5/6 \approx 0.833$. Moreover, the correlation time as defined by Howind et al. (1999) was taken constant between 100–300 s, empirically chosen by data analysis. In our proposal, the Mátern correlation time, although not directly corresponding to the correlation time, varies approximately in the same range. Moreover, it depends on atmospheric conditions (wind, outer scale length) allowing the computation of the covariance for phase measurements of two different satellites.

3.4.2 ARIMA processes

In geodesy (see exemplary Grafarend 1976), the so-called first autoregressive Markov model is often used to model correlations. It corresponds to a smoothness parameter of 1, which is also close to the $\nu = 5/6 \approx 0.833$ given by turbulence theory. Jansson and Persson (2013) fitted a covariance function with this smoothness factor of 1 to GPS observations.

Continuous ARIMA processes have a rational spectral density (Rasmussen and Williams 2006). Luo et al. (2012) made use of ARIMA(p, q) procedure to decorrelate least-squares residuals. They showed that some values of p and q perform better than others depending on the stations (for instance on the influence of multipath). Such models are however empirical and the corresponding covariance functions are mathematically more complicated to express than the Mátern one (see exemplary Jones and Vecchia 1993).

3.4.3 Treuhaft and Lany model

Treuhaft and Lany (1987) (TL) proposed a turbulence-based VCM for tropospheric delays in VLBI. This model is used by Nilsson and Haas (2010), Pany et al. (2011) or Romero-Wolf et al. (2012) for simulations or real data analysis of VLBI measurements.

The covariance between two tropospheric slant delays t_1, t_2 reads:

$$C_\varphi(t_1, t_2) = \frac{1}{\sin(\mathrm{El}_1)\sin(\mathrm{El}_2)}\left(H^2\sigma^2 - \frac{1}{2}\int_0^H\int_0^H \mathrm{d}z\mathrm{d}z' \times D_n\left(\frac{\left|\mathbf{s}_1(z) - \mathbf{s}_2(z') - \langle\mathbf{u}\rangle\Delta t\right|}{\sin(\mathrm{El})}\right)\right), \tag{13}$$

where $\langle\mathbf{u}\rangle$ is the mean wind velocity, $\mathbf{s}_1$ and $\mathbf{s}_2$ denote the line-of-sight vectors, σ^2 the variance of the wet refractivity fluctuations which is assumed independent of $\mathbf{s}$. It can be expressed by $\sigma^2 = \frac{1}{2}D_n(r)$ as $r \to \infty$, assuming that the troposphere parameters are completely uncorrelated. Built on a modified version of the Kolmogorov structure function for the refractive index, $D_n(r) = C_n^2\frac{r^{2/3}}{1+\left(\frac{r}{L}\right)^{2/3}}$, this covariance model leads to a double integral which must be solved numerically. Furthermore, Eq. (13) is based on the relation $C_\varphi(t, t+\tau) = \frac{1}{2}(D_\varphi(\infty) - D_\varphi(\tau))$ valid only for a stationary process, where D_φ is the phase structure function computed by directly integrating the modified version of the Kolmogorov structure function of the refractive index along the lines-of-sight. The parameter L, chosen in physically reasonable range, was taken to 3,000 km, a value at which empirical structure functions, for VLBI data saturate for large values of r.

In the TL proposal, neither anisotropy nor inhomogeneity was taken into account. Moreover, this model is based on a double integral which can only be solved numerically, necessitating a large computational effort.

3.4.4 Schön and Brunner covariance model

Schön and Brunner (2008a) developed a covariance model based on the time-dependent integrated separation distance d:

$$\begin{aligned}\langle\varphi_A^i(t),\ \varphi_B^j(t+\tau)\rangle &= \frac{12}{5}\frac{0.033}{\Gamma\left(\frac{5}{6}\right)}\frac{\sqrt{\pi^3}\kappa_0^{-2/3}2^{-1/3}}{\sin\mathrm{El}_i^A\sin\mathrm{El}_j^B}C_n^2\int_0^H\int_0^H(\kappa_0^d)^{1/3}K_{1/3}(\kappa_0^d)\mathrm{d}z_1\mathrm{d}z_2\\ &= \frac{2^{1/3}}{3\Gamma\left(\frac{2}{3}\right)}\frac{\kappa_0^{-2/3}}{\sin\mathrm{El}_i^A\sin\mathrm{El}_j^B}C_n^2\int_0^H\int_0^H(\kappa_0^d)^{1/3}K_{1/3}(\kappa_0^d)\mathrm{d}z_1\mathrm{d}z_2\end{aligned} \tag{14}$$

The parameters H, κ_0, c are the same as previously described. Not only the wind velocity but also the wind orientation (wind vector) is involved in the computation of the separation distance d, and the model can therefore be considered as a 2D one. A double integrated Mátern kernel with $\nu = \frac{1}{3}$ is obtained. However, for the case "one satellite" where the geometries are slowly varying with time, it will lead as in our previous formulation to a spectral density with a power law dependence of $\frac{4}{3}$ since $\nu + \frac{D}{2} = \frac{4}{3}$ with $\nu = \frac{1}{3}$, $D = 2$, D being the dimensionality (Appendix A). Thus the two formulations (Eqs. 14, 10) are equivalent for the GPS phase covariance for one satellite.

Schön and Brunner (2008a) proposed a direct integration for the variance which reads:

$$\langle\varphi^2\rangle = \frac{12}{5}\frac{0.033}{\Gamma\left(\frac{5}{6}\right)}\frac{\sqrt{\pi^3}\kappa_0^{-2/3}2^{-1/3}}{(\sin \mathrm{El})^2}C_n^2 H^2 \times \left\{\frac{\pi 2^{1/3}}{\sqrt{3}\,\Gamma\left(\frac{2}{3}\right)}{}_2F_3\left(\left[\frac{1}{2},1\right],\left[\frac{2}{3},\frac{3}{2},2\right],\frac{z^2}{4}\right) - \frac{27}{80}2^{2/3}\Gamma\left(\frac{2}{3}\right)z^{2/3}\,{}_1F_2\left(\left[\frac{5}{6}\right],\left[\frac{11}{6},\frac{7}{3}\right],\frac{z^2}{4}\right)\right\}, \quad (15)$$

where F denotes the hypergeometric function (Abramowitz and Segun 1972). The dimensionless argument z is given by $z = \frac{p\kappa_0 H}{\sin \mathrm{El}}$, where the factor p describes the impact of anisotropy on the variance. For small values of c, the variance should be replaced using the small argument approach (Eq. 11)., since the hypergeometric function (Eq. 15) takes rapidly high values, leading to computational issues by taking the difference of the two hypergeometric functions.

Schön and Brunner (2008a) computed the separation distance for a more general case by allowing a time dependency through the wind vector. As a consequence, this model is more general than our proposal. However, no dependency of the outer scale length with the separation distance or of the structure constant with the outer scale length is proposed. Moreover, for some satellite geometry, the maximum of covariance is not at the first epoch but "delayed". Such a behavior is physically difficult to understand when considering the rapid reorganization of the troposphere. The corresponding covariance matrices may be not positive definite anymore. Moreover, considering anisotropy for the turbulence parameters of interest ($c = 0, 01$, $L_0 = 6{,}000$ m) some numerical instabilities due to the double integral occur.

Thus our simplification for the case two satellites-one or two stations, allowing both a rapid computation and a physical interpretation, should be preferred.

4 Case study

In the following part, the influence of our model on the coordinate estimates in least-squares adjustments as well as the influence of Mátern parameters (smoothness and Mátern correlation time) will be studied. We aim to validate the physically derived smoothness factor and Mátern correlation time presented in Sect. 3.1. Fully populated covariance matrices are computed with the previous formulas and implemented in a weighted least-squares model. After a short presentation of the methodology used, the results for the repeatability as well as for the quadratic deviation of the batch coordinates will be discussed.

4.1 Least-squares solution

We use the Seewinkel Network (Schön and Brunner 2008b) specially designed to study the temporal correlations due to turbulence on GPS measurements. It consists of six exactly aligned stations P0, P1, P2, P4, P8, and P16 with separation of approximately 1, 2, 4, 8, and 16 km, respectively. It was measured on April, 15th 2003 during 8h (5:45–13:45 GPS time) using identical equipment, a 1Hz data rate, and a cutoff-angle of 3°. Multipath is weak, thus the correlations are assumed to come principally from tropospheric fluctuations.

The coordinates of the first station P0 were held fixed and double differences were formed. Ambiguities were precomputed. The North, East and Up (N, E, U) components of P1 and P8, respectively, were estimated for the baseline: P0P1 (1,000 m) as well as the longer baseline P0P8 (8,000 m).

Since the observations can be assumed to be uncorrelated after 600s (Schön and Kutterer 2006), the whole observation period of 8 h is split into nonoverlapping batches of 600s. To analyze the impact of the stochastic model on the coordinate repeatability, two sampling rates are used: overall 45 batches à 20 epochs, 1 epoch = 30 s and 45 batches a 600 epochs, 1epoch = 1 s were computed in a weighted least-squares adjustment. No tropospheric parameters were estimated and the CODE reprocessing orbits and clocks were used (Dach et al. 2009). An apriori standard deviation of 1 mm was assumed for the L1 carrier phase measurements.

For each batch b, the coordinates are computed:

$$\hat{\mathbf{x}}_b = [\,N_b\ E_b\ U_b\,]^{\mathrm{T}} = (\mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{A})^{-\mathbf{1}}\mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{y},\ \text{with } \mathbf{P} = \mathbf{Q}_{\mathrm{DD}}^{-1}, \quad (16)$$

$\mathbf{Q}_{DD}$ is the cofactor matrix of the double difference observations, thus $\mathbf{Q_{DD}} = \mathbf{M^T Q M}$, $\mathbf{M}$ is the matrix of the mathematical correlation (Beutler et al. 1987), $\mathbf{Q}$ the positive-definite cofactor matrix of the undifferenced observations. $\mathbf{A}$ is the design matrix of each batch, $\mathbf{y}$ the vector of double differences.

The diagonal elements of the matrix $\widehat{\boldsymbol{\Sigma}}_{\mathbf{apost},b} = \widehat{\sigma}_0^2(\mathbf{A^T P A})^{-1}$, $\widehat{\sigma}_0^2 = \frac{\mathbf{v^T Q_{DD}^{-1} v}}{n_b - 3}$ (n_b number of double differences for the batch b) are $\widehat{\sigma}_b^2 = [\,\widehat{\sigma}_{N,b}^2\ \widehat{\sigma}_{E,b}^2\ \widehat{\sigma}_{U,b}^2\,]^{\mathrm{T}}$. They represent the a posteriori variances of the unknowns. $\mathbf{v}$ is the residual vector of the least-square solution for the batch b.

The mean over all batches is computed leading to:

$$\hat{\sigma}_{i,\mathbf{apost}} = \frac{1}{m}\sum_{b=1}^{m}\hat{\sigma}_{i,b}, \tag{17}$$

where $i = \{N, E, U\}$ and m is the number of batches.

The quadratic deviation of the coordinates in [mm] reads

$$\delta_i = \sqrt{\frac{1}{m}\sum\nolimits_{b=1}^{m}\Delta\hat{x}_{i,b}^2}, i = \{N, E, U\}, \tag{18}$$

$\Delta\hat{x}_{i,b}$ is the parameter deviation for the batch b of the estimated coordinates $\hat{N}, \hat{E}, \hat{U}$ and the reference values N_0, E_0, U_0 obtained from static positioning over the whole 8-h observation window.

For all coordinate components, their quadratic deviation δ_i is compared to the a posteriori variance $\hat{\sigma}_{i,\mathbf{apost}}$. The ratio $R_i = \frac{\hat{\sigma}_{i,\mathbf{apost}}}{\delta_i}$ is formed. It should be as close as possible to 1, meaning that no overestimation occurs (Rao and Toutenburg 1999). Both $\hat{\sigma}_{i,\mathbf{apost}}$ and δ_i are independent of the a priori variance factor (Kutterer 1999). Thus the cofactor matrices and not the covariance matrices will be used.

4.2 Methodology

For different cofactor matrices, the three parameters $\delta_i, \hat{\sigma}_{i,\mathbf{apost}}, R_i$ are determined. We compare our model to the results given by other smoothness and correlation time factors as well as to the typically used elevation-dependent weighting. In the following, we will call:

- EPS model, the elevation depending model with $\mathbf{Q} = \mathbf{Q}_\varepsilon$

$$\mathbf{Q}_\varepsilon(i,i) = \text{diag}\left(\frac{1}{\sin^2(\text{El}_i)}\right) \text{ with } \mathbf{Q}_\varepsilon(i,i) = 1 \text{ when } El_i = 90^\circ \tag{19}$$

- CORR model. In this case, the global cofactor matrix before mathematical correlations is given by

$$\mathbf{Q} = (1-\beta)\,\mathbf{Q}_{\mathbf{temp}} + \beta\mathbf{Q}_\varepsilon, \tag{20}$$

  where $0 \leq \beta \leq 1$, called noise factor, is a positive parameter depending on the observations noise. Following Jansson and Persson (2013), β is defined as the ratio $\beta = \frac{\text{nugget}}{\text{sill}}$ of the structure function of the double difference observations. We found a mean value of $\beta = 0.3$ for double differences with low elevation satellites for the baseline length P0P1 and $\beta = 0.05$ for P0P8, meaning that relatively more high-frequency noise remains in the time series.

  The temporal cofactor matrices $\mathbf{Q}_{\mathbf{temp}}$ are computed thanks to the Mátern covariance functions with a given smoothness factor and Mátern correlation time and scaled as Eq. 19. To check the influence of the variation of the Mátern parameters on the least-squares solutions, they were varied from 1/6 to 4/3 for the smoothness ν and $[10^{-3}–10^{-2}]\text{s}^{-1}$ for α (inverse of the Mátern CT), respectively. The case $\nu = 1/2$ corresponds to the exponential case, while $\nu = 1$ is the AR(1) model (Appendix A). First values of the cofactor matrices were computed using the limit of the Bessel function (Abramowitz and Segun 1972) as shown in Eq. 11.

  It follows that the two cofactor matrices $\mathbf{Q}$ and $\mathbf{Q}_\varepsilon$ have the same diagonal elements. Thus, the temporal cofactor matrices are not "underweighted" as long as the observation noise is not close to 1.

Note Other noise effects such as thermal noise (Radovanovic 2001; Schön and Brunner 2008b), are modeled in a first approximation by adding an elevation-dependent matrix to the Mátern temporal cofactor matrix $\mathbf{Q}_{\mathbf{temp}}$. A positive quantity added to the diagonal elements of the fully populated covariance matrix represents measurement errors or the so-called "nugget" effect in kriging (Cressie 1993). Moreover, this matrix yields a stabilization of the covariance matrices (Tikhonov et al. 1995). Although exemplary, Williams et al. (2004) modeled the covariance of GPS observations by adding a white noise covariance matrix (power spectrum of 0 corresponding to the identity covariance matrix) and a flicker noise matrix (e.g. power spectrum with a power law of −1), we chose not to add an identity matrix but to model the observations noise as elevation-dependent.

As we do not intend to compare the influence of different covariance models but only the influence of temporal correlations that gives an optimal quadratic deviation, the case where $\mathbf{Q} = \mathbf{I}$ which leads to poorer coordinate scatter and ratio than elevation-dependent models was not analyzed here.

4.3 Building the fully populated covariance matrices

Without loss of generality, we assume a satellite-by-satellite ordering scheme of the undifferenced GPS carrier phase observations. As a consequence, for one station A, the covariance matrix is built as follow:

$$\mathbf{C}_A^{i,j} = \begin{bmatrix} \sigma_{i1}^{j1} & \sigma_{i1}^{j2} & \sigma_{i1}^{j3} & \cdots & \sigma_{i1}^{ju} \\ \sigma_{i2}^{j1} & \sigma_{i2}^{j2} & \sigma_{i2}^{j3} & & \sigma_{i2}^{ju} \\ & \cdots & \cdots & & \\ \sigma_{ik}^{j1} & & & & \sigma_{ik}^{ju} \end{bmatrix}$$

where σ_{ik}^{in} is the covariance between the satellite i at epoch k and satellite i at epoch n. The size of the covariance matrices for two different satellites is depending on the size of the

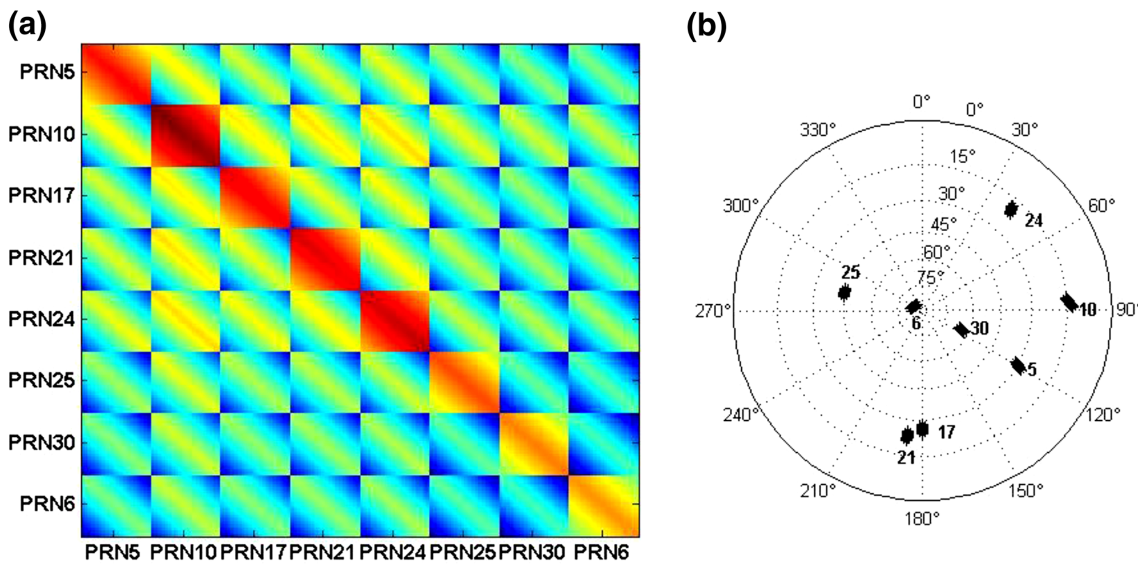


Fig. 7 **a** Example of the structure of the covariance matrix $\mathbf{C}_A$ for one station (log covariance values were plotted for more readability) and **b** the corresponding skyplot

elevation vectors: here for satellite i: k epochs and for satellite j: u epochs. It follows that the covariance matrix for the first station is given by:

$$\mathbf{C}_A = \begin{bmatrix} \mathbf{C}_\mathbf{A}^{\mathbf{1,1}} & \mathbf{C}_\mathbf{A}^{\mathbf{1,2}} & \mathbf{C}_\mathbf{A}^{\mathbf{1,3}} & \dots & \mathbf{C}_\mathbf{A}^{\mathbf{1,n}} \\ & \mathbf{C}_\mathbf{A}^{\mathbf{2,2}} & \mathbf{C}_\mathbf{A}^{\mathbf{2,3}} & & \mathbf{C}_\mathbf{A}^{\mathbf{2,n}} \\ & \dots & \mathbf{C}_\mathbf{A}^{\mathbf{3,3}} & & \\ & & & & \mathbf{C}_\mathbf{A}^{\mathbf{n,n}} \end{bmatrix}$$

$\mathbf{C}_A$ is a symmetric positive-definite covariance matrix. The diagonal blocks $\mathbf{C}_A^{\mathbf{i},j}$ of $\mathbf{C}_A$ are also symmetric but not exactly Toeplitz due to the small non-stationarity of the model. For two stations, the global covariance matrix reads:

$$\Sigma_{\mathbf{temp}} = \begin{bmatrix} \mathbf{C}_A & \mathbf{C}_{A,B} \\ \mathbf{C}_{A,B} & \mathbf{C}_B \end{bmatrix} \text{ and } \mathbf{Q}_{\mathbf{temp}} = \gamma\, \Sigma_{\mathbf{temp}},$$

$$\gamma \text{ being computed so that } \mathbf{Q}_{\mathbf{temp}}(\mathrm{El} = 90°) = 1 \text{ (cofactor matrix).}$$

Figure 7a shows an example of the structure of such a covariance matrix and Fig. 7b the corresponding skyplot. The smoothness parameter were taken to $\nu = 5/6$ and the inverse of the Mátern correlation time is $\alpha = 0.006\,\mathrm{s}^{-1}$.

Figure 7a highlights that the fully populated covariance matrices have a block diagonal structure; each sub matrices showing globally a "Toeplitz" like form. This property could be used in future to accelerate the inversion algorithm (Meurant 1992). The computation of such fully populated covariance matrices—one per batch—took only a few second for the whole observation time and is faster than the Schön and Brunner or Treuhaft and Lany model as no integration must be performed.

4.4 Results

In the following plots, black stars correspond to results with $\mathbf{Q}_\varepsilon$ called the EPS model, whereas blue or green circles are obtained with the fully populated cofactor matrices $\mathbf{Q}$ called CORR.

4.4.1 Baseline P0P1 of 1,000 m

Varying the smoothness factor

Figure 8 shows the ratio R_i (b) as well as the quadratic deviation (a) of the batch coordinates when changing the value of the smoothness parameter ν from 1/6 to 3/2 by keeping α constant, $\alpha = 0.006\,\mathrm{s}^{-1}$. This value corresponds to a physical relevant Mátern correlation time by taking $L_0 = 6000\mathrm{m}$ and $u = 8{-}10\,\mathrm{ms}^{-1}$ (see also Schön and Brunner 2008a). We plotted the results for two different values of the noise factor: $\beta = 0.3$ (blue circle) and $\beta = 0$ (green circle, no observation noise). Since the EPS model (black stars) does not depend on the smoothness factor, the results are not varying with ν and are only plotted for an easier visual comparison.

For the CORR model, the ratio R_i is close to 1 for nearly all values of ν by taking $\beta = 0.3$. The case $\beta = 0$ gives a ratio $R_i >> 1$. As expected, the EPS model is showing an overestimation of the precision, i.e. the standard deviation is smaller than the coordinate scatter.

With a noise factor $\beta = 0.3$ and for the value of interest $\nu = 5/6$, the Up and East component have a smaller quadratic deviation with the CORR model (but only at the sub-millimeter level), whereas the N component deviation giving a not significantly higher than the EPS model

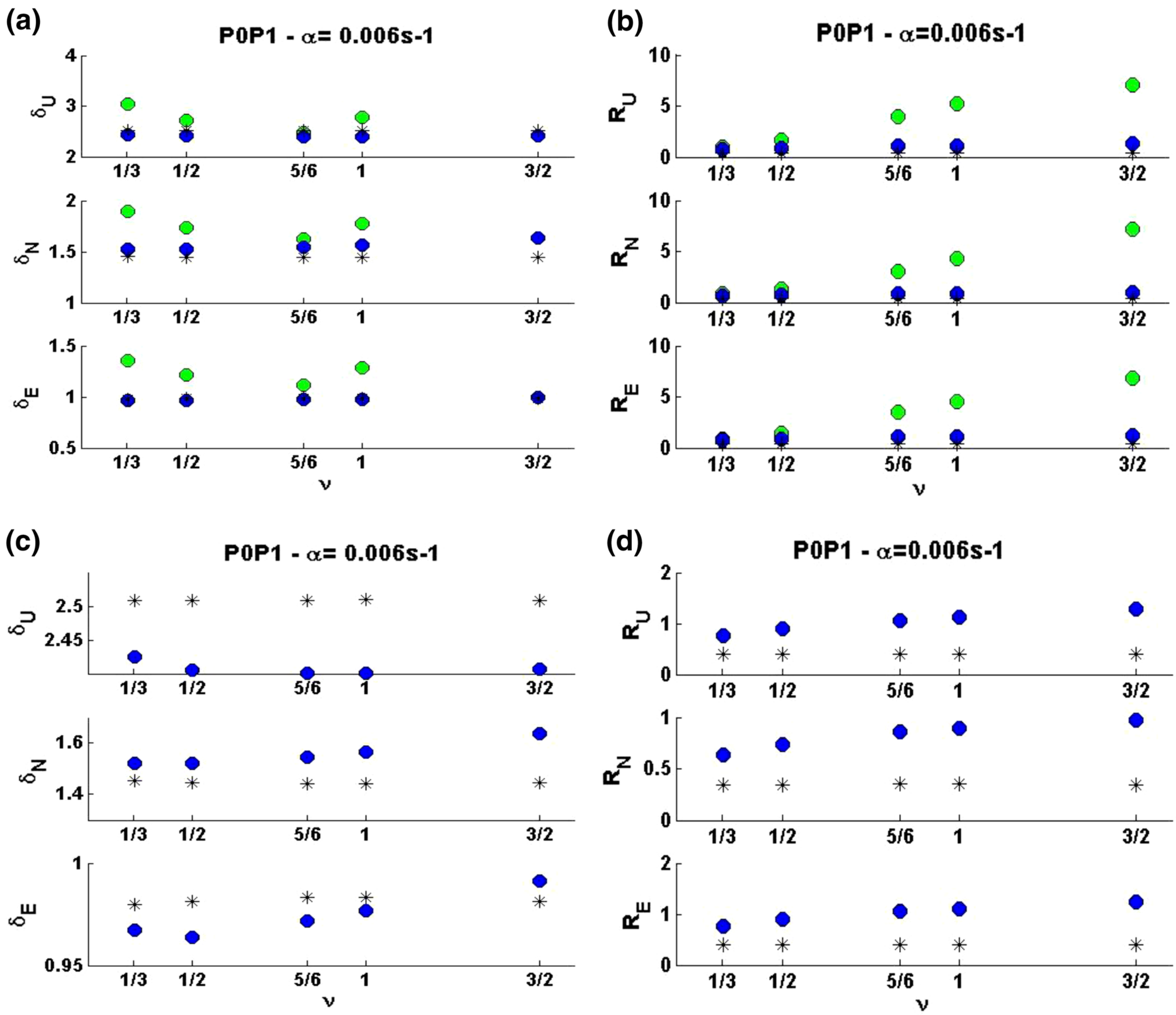


Fig. 8 Mean coordinate scatter δ_i (**a**) and ratio R_i (**b**) of the batch solutions are plotted for the N, E and U component versus smoothness factor ν for $\alpha = 0.006s^{-1}$. The values $\delta_i(\nu = 3/2)$ are higher than the other cases and not plotted for more readability. *Green circles* represent the CORR model with $\beta = 0$, *blue circles* $\beta = 0.3$ and *stars* the EPS model. **c** Zoom of **a**. **d** Zoom of **b**. Short baseline P0P1 (1,000 m)

(1.45 mm with EPS or 1.5 mm with CORR for $\nu = 5/6$). Higher values of the noise factor will lead to coordinate scatters that are comparable with the EPS model. The value of $\beta = 0.3$ gives the lowest quadratic deviation together with ratios of 1, which is coherent with the definition $\beta = \frac{\text{nugget}}{\text{sill}}$ of the structure function of the double differenced observations for the Seewinkel Network.

We can note moreover that the model with $\nu = 5/6$ is nearly equivalently performing than $\nu = 1/2$ (exponential model); $\nu = 1$ (AR(1) model) giving a lightly and not significant higher standard deviation for the E and N component. However, it should be noticed that the exponential model is giving a ratio R_i smaller than 1 for a comparable quadratic deviation, provided that the correlation length is taken accordingly. Thus, we should prefer the turbulence value of the smoothness parameter $\nu = 5/6$ which seems more reliable. Moreover, the physical interpretation is easier and allows an accurate estimation of the correlation time.

Varying the correlation time

In Fig. 9a, b, we varied the parameter $\alpha = \frac{\kappa_0 u}{a}$ for $\nu = 5/6$. It can be seen that value in the range $[6e^{-3}, 1e^{-2}]\,s^{-1}$ are giving ratio close to 1 for $\beta = 0.3$ as well as a lower quadratic deviation compared to the EPS model and the CORR model with $\beta = 0$. High values of α, i.e. small temporal correlations are leading asymptotically to the same result as the EPS model.

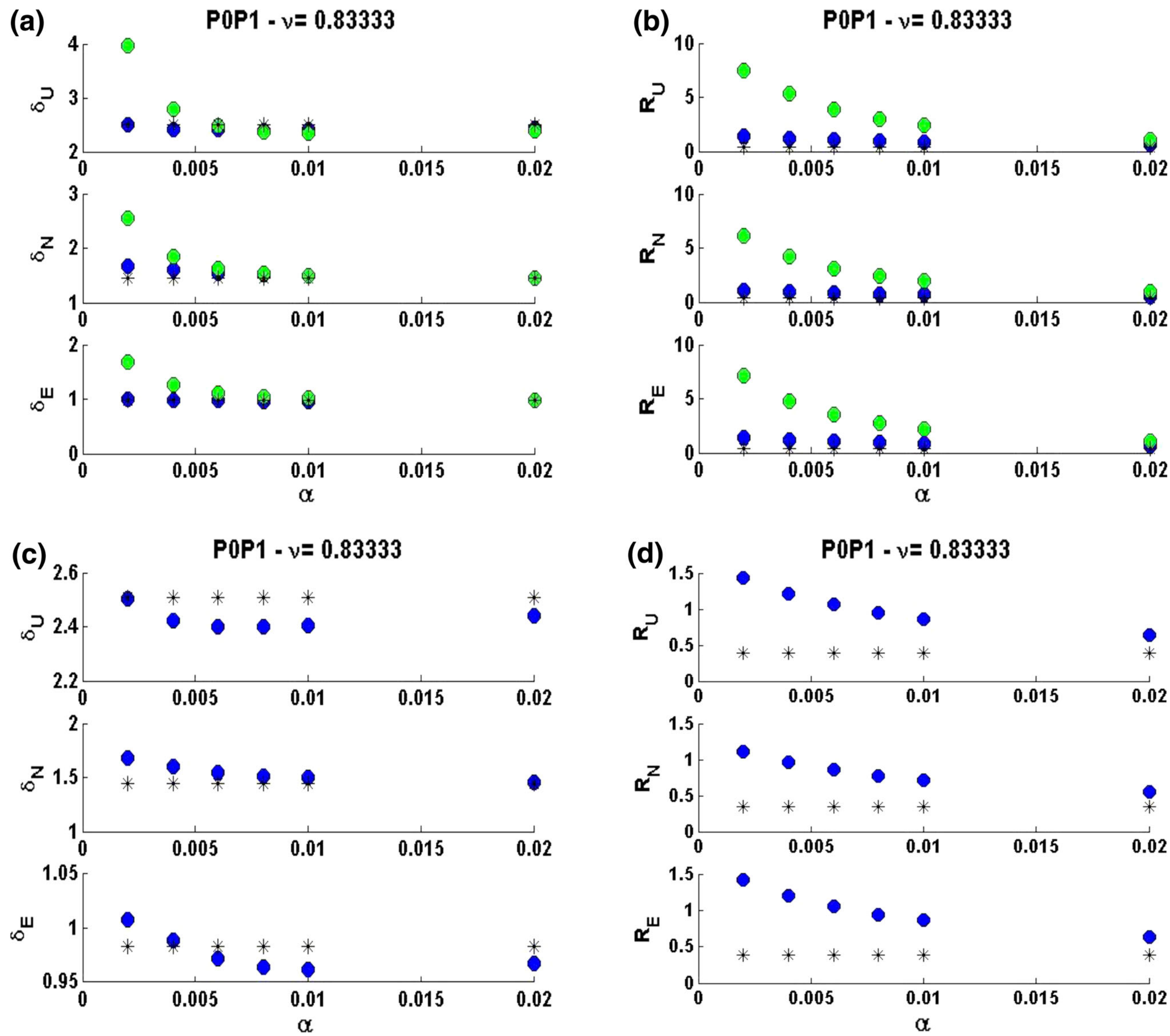


Fig. 9 Mean coordinate scatter δ_i (**a**) and ratio R_i (**b**) of the batch solutions are plotted for the N, E and U component versus α for $\nu = 5/6$. The values $\delta_i(\nu = 3/2)$ are higher than the other cases and not plotted for more readability. *Green circles* represent the CORR model with $\beta = 0$, *blue circles* $\beta = 0.3$ and *stars* the EPS model. **c** Zoom of **a**. **d** Zoom of **b**. Short baseline P0P1 (1,000 m)

From the study of the least-squares results (R_i and δ_i) for the baseline length 1,000 m, the values of $\nu = 5/6$ and $\alpha = 0,006\text{s}^{-1}$ seems to give accurate and relevant results. These values are coherent with values found by El-Rabbany (1994), Radovanovic (2001) or more recently Jansson and Persson (2013). It should be highlighted however that the differences between the different models are not very important as long as the Mátern parameters are taken in a reasonable interval.

1 s data rate

The results for the short baseline P0P1 at high data rate (1 s) were also estimated with 45 batches à 600 epochs. As before, the values of the smoothness parameters were changed as well as the value of the correlation time. It was found that the ratio R_i is close to 1 for $\nu = 5/6$. However, in this case, the noise factor was taken to 0.5, a slightly higher value as for the 30 s data rate. The coordinate scatter δ_i is smaller than those obtained with the EPS model, particularly for the Up and East components, although only in the sub-millimeter range. Once more, it was shown that varying the smoothness parameter in a reasonable range does not influence the results significantly.

The conclusions are the same as for the 30s rate: taking temporal correlations into account principally influence the

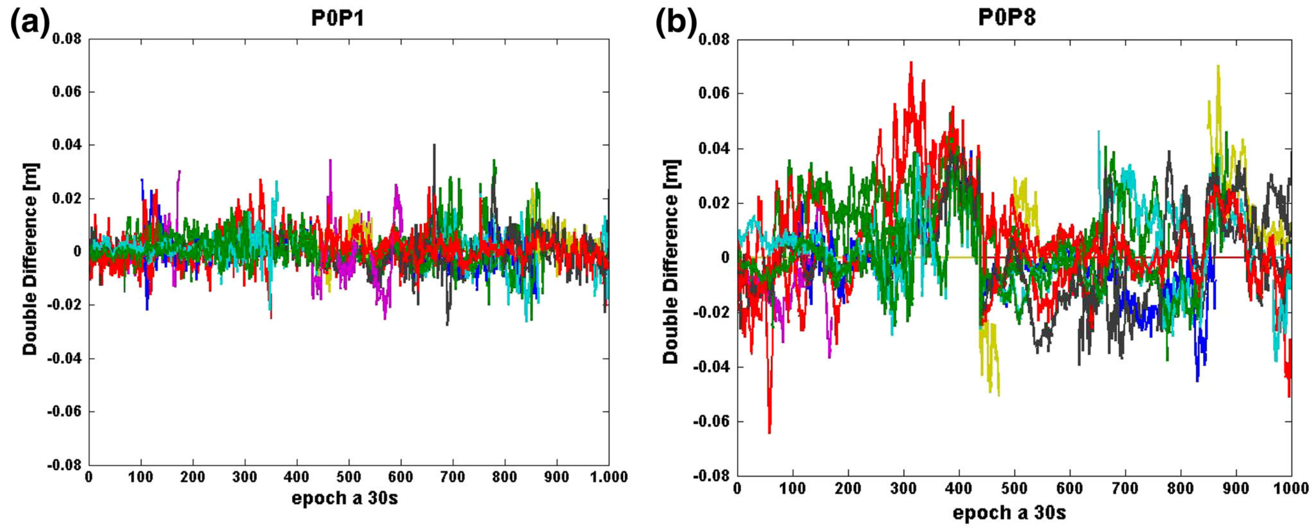


Fig. 10 Mean coordinate scatter δ_i (**a**) and ratio R_i (**b**) of the batch solution are plotted for the N, E and U component versus α for $\nu = 5/6$. *Green circles* represent the $CORR$ model with $\beta = 0.05$, *blue circles* $\beta = 0.3$ and *stars* the EPS model. Long baseline P0P8 (8,000 m)

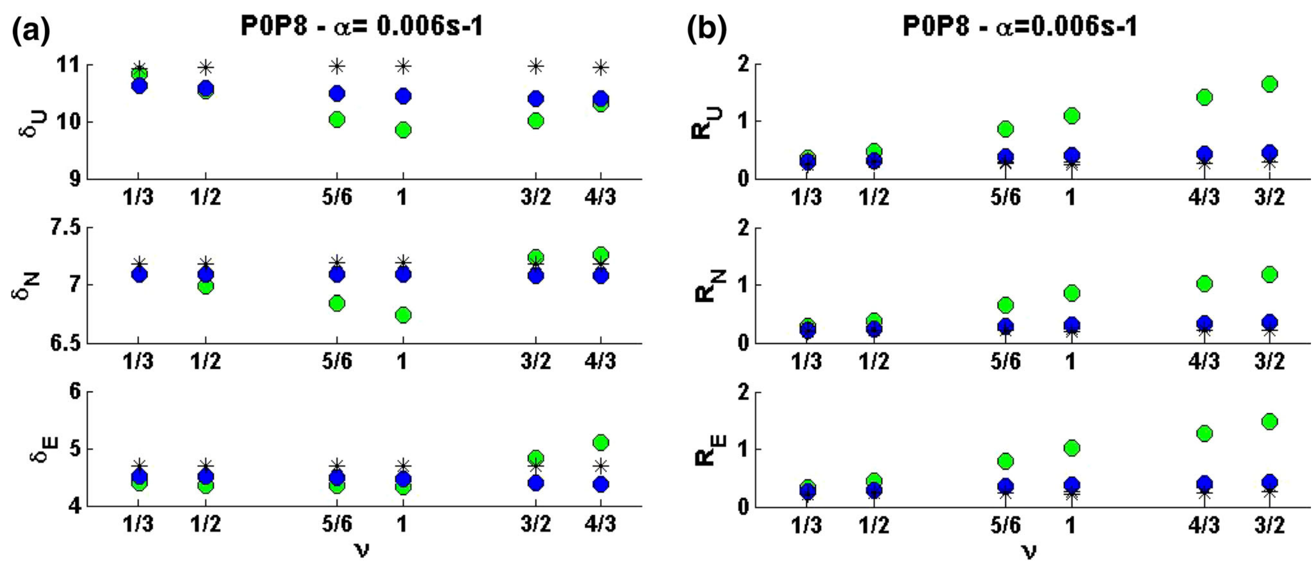


Fig. 11 Mean coordinate scatter δ_i (**a**) and ratio R_i (**b**) of the batch solution are plotted for the N, E and U component versus ν for $\alpha = 0.006s^{-1}$. *Green circles* represent the CORR model with $\beta = 0.05$, *blue circles* $\beta = 0.3$ and stars the EPS model. Long baseline P0P8 (8,000 m)

a posteriori variance. The spatial repartition of the parameter deviation (scatter) is not changing much as long as the Mátern parameters are not too far from 1. However, the a posteriori precision is more relevant and more accurate than with elevation-dependent diagonal covariance matrices.

4.4.2 Baseline P0P8 of 8,000 m

The same analysis was performed for the baseline P0P8 (8,000 m) with the same methodology. For a long baseline, other noise processes interact and the correlations are not only due to tropospheric fluctuations. Thus, in this part, we aim to show the impact of changing the Mátern covariance parameters α, ν on the least-squares adjustments. The results are presented in Figs. 11 and 10. As for the previous baseline P0P1, the value of the smoothness parameter (Fig. 11) $\nu = 5/6$ gives good results in term of ratio R_i and quadratic deviation δ_i which is smaller than with the EPS model. It should be mentioned that the noise factor of 0.05 gives the best results and not $\beta = 0.3$ as previously for the short baseline. It would mean that the nugget effect is close to 0, thus the cofactor of the observation noise is much smaller than for the short baseline of the Seewinkel Network.

The value of $\nu = 1$ provides lower values of δ_i, however with a light underestimation of the precision since $R_i > 1$. This is in agreement with the intuitive results that double

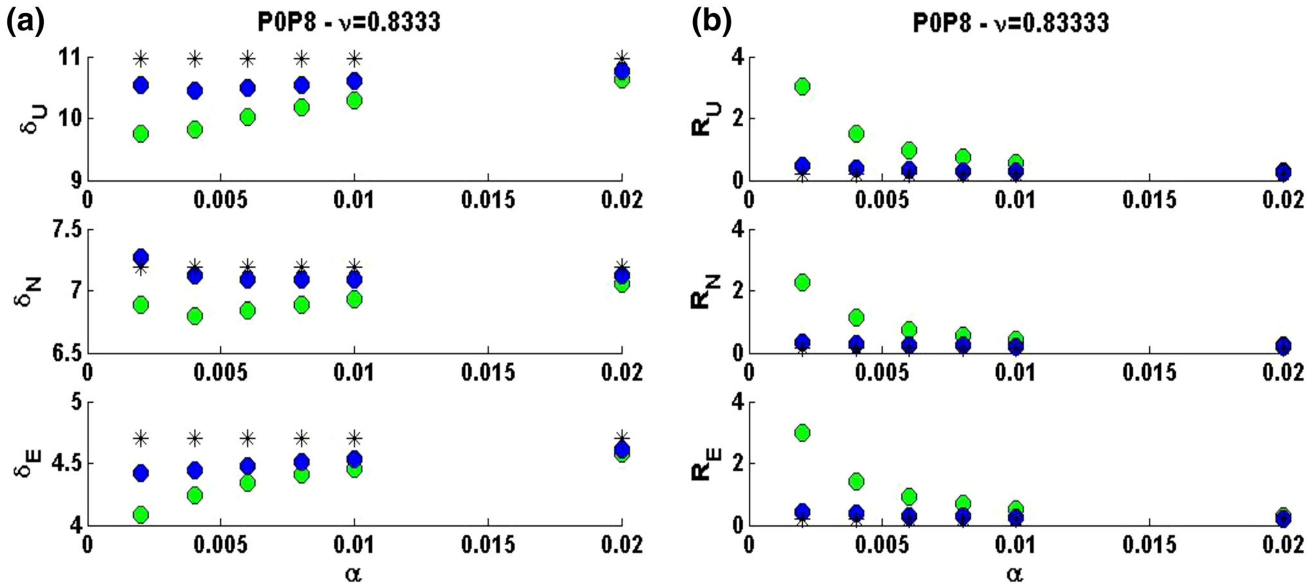


Fig. 12 **a** 8 h (30,000 s) of double differences for the 1,000 m baseline P0P1 and **b** 8 h (30,000 s) double differences for the 8,000 m baseline P0P8. Each satellite pair is color coded, the sampling rate is 30 s

differences from long baseline (Fig. 12b) are less smooth (Fig. 12a) than for shorter baseline since not all effects are canceled by double differencing, like, i.e. ionospheric propagation effects. This difference in term of standard deviation is however not very important (sub-millimeter level) and should be carefully interpreted by analyzing at the same time the values of R_i. This effect will be studied in the future for longer baselines.

Figure 10 presents the results of the least-squares adjustment when α is varied. As for the variation of the smoothness factor, β should be set close to 0. The value of $\alpha = 0.006\,\mathrm{s}^{-1}$ gives at the same time an improved coordinate scatter compared with the EPS model as well as a ratio of 1 for all three components. Smaller values of α $(0.002{-}0.004\,\mathrm{s}^{-1})$ gives also small values of δ_i, however, as with higher values of ν, with $R_i >> 1$.

The improvement of the CORR model with $\beta = 0.05$, $\nu = 5/6$, $\alpha = 0.006\ \mathrm{s}^{-1}$ is at the millimeter level for all three components in comparison with the EPS model, the ratio being at the same time more than 3 times better than with the standard model which remains a great improvement.

5 Conclusion

Temporal correlations of GNSS phase measurements due to tropospheric fluctuations are not taken in consideration in the currently used weighted least-squares models. The result is an overestimation of the a posteriori precision by up to a factor 10. Results of the Kolmogorov turbulence theory were used to develop a new and simplified model for the temporal correlations thanks to the Mátern covariance family which best suits to model both the temporal correlations for one satellite and for two satellites observed at one or two separated stations. The concepts of separation distance as well as inhomogeneity and anisotropy were taken in consideration. Using Taylor's frozen hypothesis, it was shown that a smoothness parameter of 5/6 as well as a Mátern correlation time between 125 and 200 s, in accordance with previous studies, should correctly model temporal correlations due to the tropospheric propagation. Moreover, it was stressed that all other parameters such as the structure constant, the depth or height of the troposphere as well as the vertical elongated parameter or the outer scale length should be carefully chosen.

In a case study, we used the data of the Seewinkel Network, specially designed for research on the effect of tropospheric fluctuations on GPS phase measurements, with weak multipath. Fully populated cofactor matrices were computed and compared with diagonal matrices of type $\frac{1}{\sin^2(\mathrm{El})}$ which are widely used in current processing softwares. To model other noise effects, an elevation-dependent diagonal matrix was added to the temporal cofactor matrices. It was shown that for the short baseline, fully populated cofactor matrices improves the quadratic deviation of the coordinates at the sub-millimeter level for short baseline and at the millimeter level for long baseline) compared with elevation dependant models. However, the a posteriori precision is more reliable than with diagonal matrices, i.e. no overestimation occurs. The results were similar for longer baseline (8,000 m) as well as for higher data rate (1 s), although for long baseline a higher smoothness factor than 5/6 could be taken in consideration. The new model leads globally to a slightly better and physically more relevant results in terms of quadratic deviation of the coordinates and a posteriori variance of the unknowns than the exponential and AR(1) model. Thus the proposed formula with a smoothness factor of $\nu = 5/6$ is promising, particularly for short baselines. As no double integrations are performed such as in other models (Treuhaft and Lany, Schön and Brunner) the computational time remains manageable. Some simplifications due to the Toeplitz like form of the covariance matrices could moreover lead to a faster implementation.

Modeling temporal correlations with the turbulence theory and Mátern cofactor function has lead to improved results in the least-squares solution, being at the same time technically feasible. In a next work, the impact on ambiguity resolutions will be shown as well as the influence of outliers on the parameters. Depending on the baseline length the choosing of the Mátern parameters could be moreover studied by minimum likelihood estimation.

Acknowledgments The authors gratefully acknowledge the funding by the DFG under the label SCHO1314/1-2. Fritz K. Brunner is warmly thanked for discussions on turbulence theory and for providing the GPS data of the Seewinkel Network. The valuable comments of three anonymous reviewers helped us improve significantly the manuscript.

Appendix: The Mátern covariance functions

A short introduction to the Mátern covariance family, also called Whittle Mátern covariance family, von Karman model (oceanography), Markov processes (geodesy) or autoregressive models (meteorology) is presented here, giving the principal features, vocabulary as well as dependencies. More details can be exemplarily found in Stein (1999), Mátern (1960), Guttorp and Gneiting (2005), Grafarend and Awange (2012). Mátern covariance functions have been concretely used by Handcock and Wallis (1994) to model meteorological fields. Fuentes (2002) also derived a non stationary family for the determination of air quality models.

A 2D autoregressive continuous process AR(1) called $Z(x, y)$ can be described by the stochastic differential equation:

$$\left(\frac{\partial}{\partial x^2}^2 + \frac{\partial}{\partial y^2}^2 - \alpha^2\right) Z(x, y) = \varepsilon(x, y), \tag{21}$$

where $\varepsilon(x, y)$ is white noise and α a constant. The corresponding spectral density is given by

$$W(\omega) \propto \frac{1}{(\omega^2 + \alpha^2)^2}, \tag{22}$$

with $\omega^2 = \omega_1^2 + \omega_2^2$ for the 2D case (Whittle 1954). The stationary covariance between two points $\mathbf{x}, \mathbf{x}'$ for this process is:

$$C(\mathbf{x}, \mathbf{x}') = C(r) = (\alpha r) K_1(\alpha r), \tag{23}$$

where K_1 is the modified Bessel function of 1st order, $r = \|\mathbf{x} - \mathbf{x}'\|$ for the isotropic case ($\|.\|$ being the norm of the vector). Whittle (1954) presented such a covariance function as a "natural spatial covariance" for the 2D case, as the exponential-based covariance functions are for one-dimensional processes.

Mátern (1960) used Whittle's result and derived for any dimension d a family of covariance functions based on an isotropic spectral density:

$$W(\omega) = \frac{2^{\nu-1}\phi \quad \Gamma(\nu + d/2)\,\alpha^{2\nu}}{\pi^{d/2}(\omega^2 + \alpha^2)^{\nu+d/2}}, \tag{24}$$

where $\omega^2 = \omega_1^2 + \omega_2^2 + \ldots + \omega_d^2$ is the angular frequency, Γ the Gamma function (Abramowitz and Segun 1972) and $\nu > 0, \alpha > 0, \phi > 0$ are constant parameters, d the dimension. The corresponding Mátern class of covariance functions is positive definite and reads:

$$C(r) = \phi\,(\alpha r)^{\nu}\,K_{\nu}(\alpha r)\,. \tag{25}$$

The parameter ν can be seen as a measure of the differentiability of the field (Stein 1999) thus "its smoothness". The constant α indicates how the correlations decay with increasing distance. Its inverse is usually called the correlation length in Kriging.

Smoothness parameter ν

Figure 13 highlights the influence of the smoothness parameter ν and the correlation length by simulating a random field/time series corresponding to the covariance function (Cressie 1993) using the eigenvalue decomposition of the corresponding Toeplitz covariance matrix (Vennebusch et al. 2010). The same random vector was used for each simulation.

The smoothness parameter ν was varied from 1/6 to 3. As ν increases, the time series are becoming less noisy for high frequency, the long periodic variations are predominating. The variance is decreasing with the smoothness parameter.

Correlation time

Figure 14 shows the influence of the parameter α. It was varied from 0.25 to 1 by keeping ν constant to 1 to simulate short and long correlation times.

Using the previous parametrization of the Mátern covariance family, the variance is not varying with the correlation time. The simulations of time series (Fig. 14b) highlight that changes of the correlation time are not acting on the smoothness of the field.

Other parametrizations

In the literature, further formulation of these covariance family functions are given (Handcock and Wallis 1994) where the parameter $\rho, \rho > 0$ is nearly independent of ν:

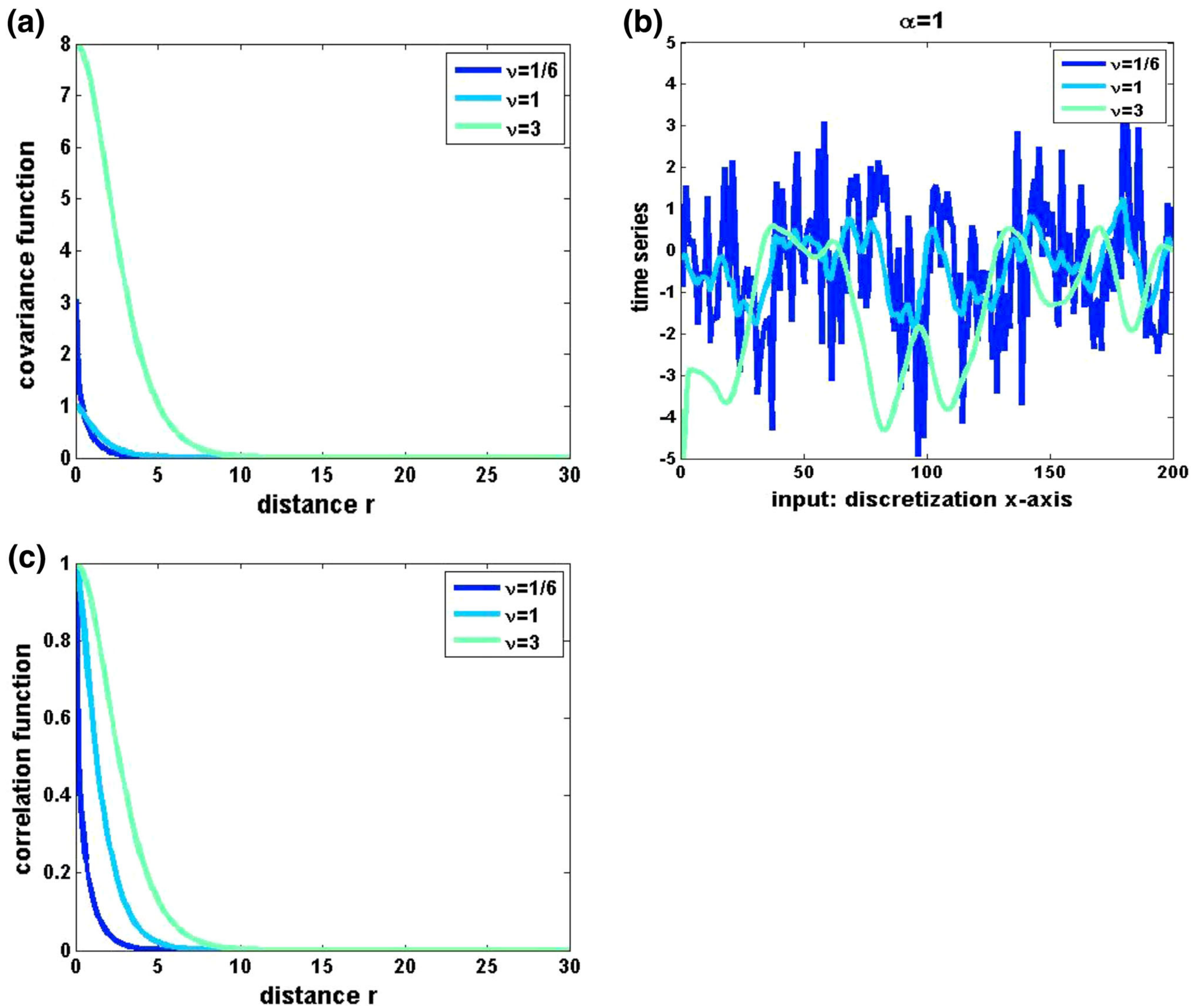


Fig. 13 **a** Covariance function (Màtern family) with $\alpha = 1$ by varying ν and **b** corresponding time series. The x axis was discretized of 200 equally spaced points. **c** Corresponding correlation function

$$C(r) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu} r}{\rho} \right)^{\nu} K_{\nu} \left(\frac{\sqrt{2\nu} r}{\rho} \right),$$

with a spectral density of the form:

$$W(\omega) = \frac{2^d \pi^{d/2} \Gamma(\nu + d/2)(2\nu)^{\nu}}{\Gamma(\nu) \rho^{2\nu}} \left(\frac{2\nu}{\rho^2} + \omega^2 \right)^{-(\nu+d/2)}.$$

This parametrization is said to be more stable when estimating the parameters ν, α with the maximum likelihood method (Stein 1999). We made use of it to develop our model for GPS phase correlations.

Shkarofsky (1968) presented a more general form of the Mátern model by introducing a shape parameter $\delta > 0$. The corresponding correlation function reads:

$$C(r) = \frac{1}{\delta^{\nu} K_{\nu}(\delta)} \left(\frac{r^2}{L^2} + \delta^2 \right)^{\nu/2} K_{\nu} \left(\sqrt{\frac{r^2}{L^2} + \delta^2} \right).$$

Thus with $\delta = 0$, the Mátern family is obtained.

If is half-integer, the covariance can be expressed in term of a product of an exponential and a polynomial of order p (Rasmussen and Williams 2006): for $\nu = 1/2$, the exponential model is obtained whereas for $\nu = 3/2$, $C(r) = \left(1 + \frac{\sqrt{3}r}{\rho}\right) e^{-\frac{\sqrt{3}r}{\rho}}$, which corresponds to a Markov process of second order and for $\nu = 5/2$, a Markov process of third order.

Advantage of the Mátern family

The advantages of the Mátern covariance functions' family for spatial interpolation are multiple, as developed in Stein (1999). Its flexibility to model the smoothness of physical processes (thus the rate of decay of the spectral density at high frequencies) is particularly useful, as well as the possibility to include non-stationarity or anisotropy (Fuentes 2002; Spöck

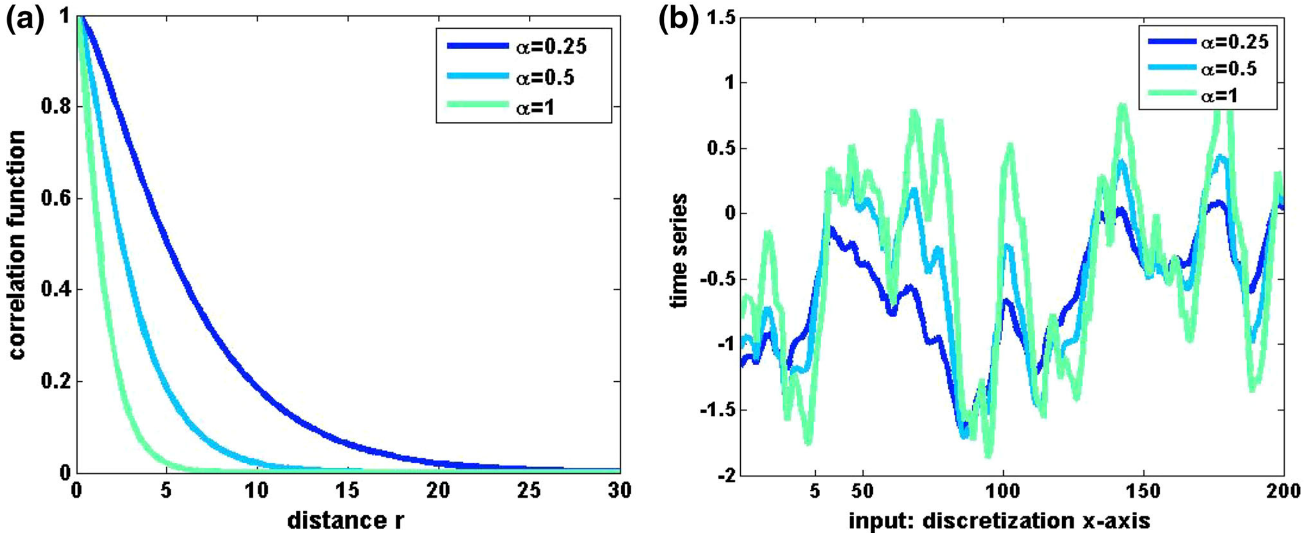


Fig. 14 **a** Covariance function (Mátern family) with $\nu = 1$ by varying α and **b** the corresponding time series. The x axis was discretized of 200 equally spaced points

and Pilz 2008). The degree of smoothness can be estimated a priori or being fixed in advance and the number of parameters to manage stays reasonable. The exponential ($\nu = \frac{1}{2}$) and Gaussian case ($\nu = \infty$) are two particular cases of this family although the last one that represents an infinitely differentiable field is concretely rarely found (Stein 1999; Handcock and Wallis 1994).

References

Abramowitz M, Segun IA (1972) Handbook of mathematical functions. Dover, New York edition

Beutler G, Bauersima I, Gurtner W, Rothacher M (1987) Correlations between simultaneous GPS double difference carrier phase observations in the multistation mode: Implementation considerations and first experiences. Manusc. Geod. 12(1):40–44

Böhm J, Schuh H (2013) Atmospheric effects in space geodesy. Springer, Berlin

Borre K, Tiberius C (2000) Time series analysis of GPS observables. In: Proceedings of ION GPS 2000, Salt Lake City, UT, USA, September 19–22, pp 1885–1894

Brunner FK, Hartinger H, Troyer L (1999) GPS signal diffraction modelling: the stochastic SIGMA-δ model. J Geod 73(5):259–267

Coulman CE, Vernin J (1991) Significance of anisotropy and the outer scale of turbulence for optical and radio seeing. Appl Opt 30(1):118–126

Cressie N (1993) Statistics for spatial data. Wiley, New York

Dach R, Brockmann E, Schaer S, Beutler G, Meindl M, Prange L, Bock H, Jäggi A, Ostini L (2009) GNSS processing at CODE: status report. J Geod 83(3–4):353–365

El-Rabbany A (1994) The effect of physical correlations on the ambiguity resolution and accuracy estimation in GPS differential positioning. PhD thesis, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada

Euler HJ, Goad CC (1991) On optimal filtering of GPS dual frequency observations without using orbit information. Bull Géodésique 65(2):130–143

Farge M (1992) Wavelet transform and their applications to turbulence. Ann Rev Fluid Mech 24:395–457

Fuentes M (2002) Spectral methods for nonstationary processes. Biometrika 89:197–210

Gage KS (1979) Evidence for a k to the -5/3 law inertial range in mesoscale two-dimensional turbulence. J Atmos Sci 36:1950–1954

Gradinarsky LP (2002) Sensing atmospheric water vapor using radio waves. Ph.D. thesis, School of Electrical Engineering, Chalmers University of Technology, Göteborg, Sweden

Grafarend EW (1976) Geodetic applications of stochastic processes. Phys Earth Planet Inter 12(2–3):151–179

Grafarend EW, Awange J (2012) Applications of linear and nonlinear models. Springer, Berlin

Guttorp P, Gneiting T (2005) On the Whittle–Mátern correlation family. NRCSE technical report series n°80

Handcock MS, Wallis JR (1994) An approach to statistical spatial–temporal modeling of meteorological fields. J Am Stat Assoc 89(426):368–378

Hartinger H, Brunner FK (1999) Variances of GPS phase observations: the SIGMA-ε model. GPS Sol 2(4):35–43

Howind J, Kutterer H, Heck B (1999) Impact of temporal correlations on GPS-derived relative point positions. J Geod 73(5):246–258

Hunt JCR, Morrison JF (2000) Eddy structure in turbulent boundary layers. Eur J Mech B Fluids 19:673–694

Ishimaru A (1997) Wave propagation and scattering in random media. IEEE Press and Oxford University Press, New York

Jansson P, Persson CG (2013) The effect of correlation on uncertainty estimates—with GPS examples. J Geod Sci 3(2):111–120

Jones RH, Vecchia AV (1993) Fitting continuous ARMA models to unequally spaced spatial data. J Am Stat Assoc 88(423):947–954

Khujadze G, Nguyen van yen R, Schneider K, Oberlack M, Farge M (2013) Coherent vorticity extraction in turbulent boundary layers using orthogonal wavelets . J. Phys. Conf. Ser. 318:022011

Kleijer F (2004) Tropospheric modeling and filtering for precise GPS leveling. PhD Thesis Netherlands Geodetic Commission, Publications on Geodesy 56

Kolmogorov NA (1941) Dissipation of energy in the locally isotropic turbulence. Proc USSR Acad Sci 32:16–18 (Russian)

Kraichnan RH (1974) On Kolmogorov's inertial range theories. J Fluid Mech 62(2):305–330

Kutterer H (1999) On the sensitivity of the results of least-squares adjustments concerning the stochastic model. J Geod 73:350–361

Lesieur M (2008) Turbul Fluids, 4th edn. Springer, The Netherlands

Luo X, Mayer M, Heck B (2012) Analysing time series of GNSS residuals by means of ARIMA processes. Int Assoc Geod Symp 137:129–134

Luo X, Mayer M, Heck B (2011) A realistic and easy-to-implement weighting model for GNSS phase observations. In: International union of geodesy and geophysics (IUGG), General Assembly 2011 - Earth on the Edge: Science for a Sustainable Plane, Melbourne, Australia, 28.06.-07.07.2011

Mátern B (1960) Spatial variation-stochastic models and their application to some problems in forest surveys and other sampling investigation. Medd. Statens Skogsforskningsinstitut 49(5)

Meurant G (1992) A review on the inverse of tridiagonal and block tridiagonal matrices. SIAM J Matrix Anal Appl 13(3):707–728

Monin AS, Yaglom AM (1975) Statistical fluid mechanics, vol 2. MIT Press, Cambridge

Nilsson T, Haas R (2010) Impact of atmospheric turbulence on geodetic very long baseline interferometry. J Geophys Res 115:B03407

Pany A, Böhm J, MacMillan D, Schuh H, Nilsson T, Wresnik J (2011) Monte Carlo simulations of the impact of troposphere, clock and measurement errors on the repeatability of VLBI positions. J Geod 85:39–50

Radovanovic RS (2001) Variance–covariance modeling of carrier phase errors for rigorous adjustment of local area networks, IAG 2001 Scientific Assembly. Budapest, Hungary, September 2–7, 2001

Rao C, Toutenburg H (1999) Linear models. Least-squares and alternatives, 2nd edn. Springer, New York

Rasmussen CE, Williams C (2006) Gaussian processes for machine learning. The MIT Press, New York

Romero-Wolf A, Jacobs CS, Ratcli JT (2012) Effects of tropospheric spatio-temporal correlated noise on the analysis of space geodetic data. In: IVS general meeting proceedings, Madrid, Spain, March 5–8, 2012

Satirapod C, Wang J, Rizos C (2003) Comparing different GPS data processing techniques for modelling residual systematic errors. J Surv Eng 129(4):129–135

Schön S, Kutterer H (2006) A comparative analysis of uncertainty modeling in GPS data analysis. In: Rizos C, Tregoning P (eds) Dynamic planet - monitoring and understanding a dynamic planet with geodetic and oceanographic tools. International Association of Geodesy Symposia, vol 130. Springer, Berlin, Heidelberg, New York pp 137–142

Schön S, Brunner FK (2007) Treatment of refractivity fluctuations by fully populated variance-covariance matrices. In: Proc. 1st colloquium scientific and fundamental aspects of the Galileo programme Toulouse Okt

Schön S, Brunner FK (2008a) Atmospheric turbulence theory applied to GPS carrier-phase data. J Geod 1:47–57

Schön S, Brunner FK (2008b) A proposal for modeling physical correlations of GPS phase observations. J Geod 82(10):601–612

Shrakrofsky IP (1968) Generalized turbulence space-correlation and wave-number spectrum-function pairs. Can J Phys 46:2133–2153

Spöck G, Pilz J (2008) Non-spatial modeling using harmonic analysis. VIII Int. Geostatistics congress, Santiago, 2008, pp 1–10

Stein ML (1999) Interpolation of spatial data. Some theory for kriging. Springer, New York

Stull RB (2009) An introduction to boundary layer meteorology. Springer, Berlin

Tatarskii VI (1971) Wave propagation in a turbulent medium. McGraw-Hill, New York

Taylor GI (1938) The spectrum of turbulence. In: Proceedings of the Royal Society London, Seires A CLXIV, pp 476–490

Treuhaft RN, Lanyi GE (1987) The effect of the dynamic wet troposphere on radio interferometric measurements. Radio Sci 22(2):251–265

Tikhonov AN, Goncharsky AV, Stepanov VV, Yagola AG (1995) Numerical methods for the solution of Ill-posed problems. Kluwer Academic Publishers, Dordrecht

Vennebusch M, Schön S, Weinbach U (2010) Temporal and spatial stochastic behavior of high-frequency slant tropospheric delays from simulations and real GPS data. Adv Space Res 47(10):1681–1690

Voitsekhovich VV (1995) Outer scale of turbulence: comparison of different models. J Opt Soc Am A 12(6):1346–1353

Wang J, Satirapod C, Rizos C (2002) Stochastic assessment of GPS carrier phase measurements for precise static relative positioning. J Geod 76(2):95–104

Wheelon AD (2001) Electromagnetic scintillation part I geometrical optics. Cambridge University Press, Cambridge

Whittle P (1954) On stationary processes in the plane. Biometrika 41(434):449

Wieser A, Brunner FK (2000) An extended weight model for GPS phase observations. Earth Planet Space 52:777–782

Williams S, Bock Y, Fang P, Jamason P, Nikolaidis RM, Prawirodirdjo L, Miller M, Johnson DJ (2004) Error analysis of continuous GPS position time series. J Geophys Res 109:B03412

J Geod (2016) 90:793–805
DOI 10.1007/s00190-016-0911-z

ORIGINAL ARTICLE

Taking correlations in GPS least squares adjustments into account with a diagonal covariance matrix

Gaël Kermarrec[1] · **Steffen Schön**[1]

Received: 25 November 2015 / Accepted: 23 April 2016 / Published online: 11 May 2016

Abstract Based on the results of Luati and Proietti (Ann Inst Stat Math 63:673–686, 2011) on an equivalence for a certain class of polynomial regressions between the diagonally weighted least squares (DWLS) and the generalized least squares (GLS) estimator, an alternative way to take correlations into account thanks to a diagonal covariance matrix is presented. The equivalent covariance matrix is much easier to compute than a diagonalization of the covariance matrix via eigenvalue decomposition which also implies a change of the least squares equations. This condensed matrix, for use in the least squares adjustment, can be seen as a diagonal or reduced version of the original matrix, its elements being simply the sums of the rows elements of the weighting matrix. The least squares results obtained with the equivalent diagonal matrices and those given by the fully populated covariance matrix are mathematically strictly equivalent for the mean estimator in terms of estimate and its a priori cofactor matrix. It is shown that this equivalence can be empirically extended to further classes of design matrices such as those used in GPS positioning (single point positioning, precise point positioning or relative positioning with double differences). Applying this new model to simulated time series of correlated observations, a significant reduction of the coordinate differences compared with the solutions computed with the commonly used diagonal elevation-dependent model was reached for the GPS relative positioning with double differences, single point positioning as well as precise point positioning cases. The estimate differences between the equivalent and classical model with fully populated covariance matrix were below the mm for all simulated GPS cases and below the sub-mm for the relative positioning with double differences. These results were confirmed by analyzing real data. Consequently, the equivalent diagonal covariance matrices, compared with the often used elevation-dependent diagonal covariance matrix is appropriate to take correlations in GPS least squares adjustment into account, yielding more accurate cofactor matrices of the unknown.

Keywords GPS · Correlations · Weighted least squares · Equivalent kernel · Mátern covariance family

1 Introduction

The functional model of the least squares adjustment in GPS positioning which links the GPS observables and the unknown model parameters is generally well described (Teunissen and Kleusberg 1998), while the currently used stochastic models remain improvable, resulting in over-optimistic parameter results, i.e., the formal variances of the estimated coordinates are too small compared to the actual coordinate scatter (El-Rabbany 1994; Radovanovic 2001). Current variance models assume either homoscedasticity (same variance for all observations) or heteroscedasticity (different variances but uncorrelated observations) using elevation-dependent models (Euler and Goad 1991; Dach et al. 2009), C/N0 (e.g., Wieser and Brunner 2000) or SNR-based models (Luo 2012a).

Physical correlations between GPS phase measurements are challenging to model and generally neglected whereas mathematical correlations for single or double differences are well studied and implemented in software (Beutler et al. 1986; Santos et al. 1997). Indeed, physical correlations depend on the receivers used (Borre and Tiberius 2000; Bona 2000) or on the environment through which GPS signals

✉ Gaël Kermarrec
gael.kermarrec@web.de

[1] Institut für Erdmessung (IfE), Leibniz Universität Hannover, Schneiderberg 50, 30167 Hannover, Germany

propagate. Especially, multipath (Luo 2012b) and turbulent refractivity fluctuations (Schön and Brunner 2008) are correlating GPS phase measurements. However, resulting fully populated variance–covariance matrices (VCM) for GPS double differences were empirically computed and used in least squares adjustments for different baseline lengths and data sets (Howind et al. 1999; Radovanovic 2001; Wang et al. 2002; Leandro et al. 2005; Satirapod et al. 2003; Jin et al. 2010). As inverting fully populated matrices is computationally demanding, different strategies were proposed by Howind et al. (1999) with an LU factorization of the inverse of the covariance matrix, Wang et al. (2002) who used an iterative procedure or Satirapod et al. (2003) with stochastic segmented methods. Klees and Broersen (2002) handled colored noise in least squares adjustments thanks to pre-conditioning techniques. However, a main issue is that normal equation stacking techniques are no longer feasible with fully populated VCM, i.e., standard scientific and commercial GNSS software packages rarely accept such matrices. Furthermore, quality control based on outlier rejection is more difficult.

Thus, in a first approach, physical correlations are often neglected and diagonal covariance matrices used for the raw observations. A consequence of this mis-modeling is an overestimation of the precision and sub-optimal parameter estimates. This situation could be improved by applying equivalent diagonal covariance matrices.

Based on statistical equivalence between least squares models (Rao and Toutenburg 1999; Krämer 1986), and the works of Luati and Proietti (2011), an equivalent diagonal covariance matrix that gives identical results as a fully populated VCM for the mean estimator case will be exposed. We will show that these results can be empirically extended to GPS adjustments thanks to the particular form of the design matrix. Subsequently, temporal and physical correlations can be taken into account assuming that an adequate covariance model is known.

The remainder of this paper is structured as follows: in Sect. 2, the mathematical equivalence between different least squares models in terms of estimates and variance estimators is shortly addressed. Next, using the mean estimator example, the computation of equivalent diagonal covariance matrices is explained in details. In Sect. 3, simulations studies are carried out to extend step by step the results to GPS positioning and check the feasibility of using the equivalence kernel.

2 Equivalences of least squares models

2.1 Mathematical background

Let

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon} \tag{1}$$

be the linear functional model, where $\mathbf{y}$ is the $n \times 1$ observation vector, $\mathbf{A}$ the non-stochastic $n \times u$ design matrix with full column rank ($rk(\mathbf{A}) = u$), $\mathbf{x}$ the $u \times 1$ parameter vector to be estimated, $\boldsymbol{\varepsilon}$ the $n \times 1$ observation error vector with $E(\boldsymbol{\varepsilon}) = \mathbf{0}$, $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathbf{T}}) = \sigma^2\mathbf{W}$, where $\mathbf{W}$ is a $n \times n$ positive definite fully populated cofactor matrix of the observations, σ^2 the a priori variance factor, and $E(.)$ denotes the mathematical expectation.

The Aitken theorem (Rao and Toutenburg 1999, p.105) states that the generalized least squares estimator (GLSE) is the best linear unbiased estimator (BLUE) for $\mathbf{x}$:

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathbf{T}}\mathbf{W}^{-1}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{W}^{-1}\mathbf{y}, \tag{2}$$

and the covariance matrix $V(\cdot)$ of the unknowns (a priori estimator) is given by

$$V(\hat{\mathbf{x}}) = \sigma^2(\mathbf{A}^{\mathbf{T}}\mathbf{W}^{-1}\mathbf{A})^{-1} = \sigma^2\mathbf{S}^{-1}, \tag{3}$$

with $\mathbf{S} = \mathbf{A}^{\mathbf{T}}\mathbf{W}^{-1}\mathbf{A}$ and $\mathbf{v} = \mathbf{y} - \mathbf{A}\hat{\mathbf{x}}$ being the $n \times 1$ residual vector.

$$\hat{V}(\hat{\mathbf{x}}) = \hat{\sigma}_0^2\mathbf{S}^{-1} \tag{4}$$

is the estimator for the parameter covariance matrix, where $\hat{\sigma}_0^2$ is called the a posteriori variance factor of the observations:

$$\hat{\sigma}_0^2 = \frac{(\mathbf{y} - \mathbf{A}\hat{\mathbf{x}})^{\mathbf{T}}\mathbf{W}^{-1}(\mathbf{y} - \mathbf{A}\hat{\mathbf{x}})}{n-u} = \frac{\mathbf{v}^{\mathbf{T}}\mathbf{W}^{-1}\mathbf{v}}{n-u}. \tag{5}$$

These estimators are unbiased, i.e., $E(\hat{\sigma}_0^2) = \sigma^2$, $E(\hat{V}(\hat{\mathbf{x}})) = \sigma^2\mathbf{S}^{-1}$ (Koch 1999; Rao and Toutenburg 1999) when the correct weight matrix $\mathbf{W}^{-1}$ is used.

The ordinary least squares estimator (OLSE) is obtained by taking $\mathbf{W} = \mathbf{I}$, thus identical variances for all observations (homoscedasticity) and no correlation between observations are assumed. For the GLSE, $\mathbf{W} = \mathbf{W}_F$ is the fully populated cofactor matrix of the observations.

Geodetic readers will refer to GLSE as "the classical least squares" and for OLSE, the least squares with identical weight. In the following, we will make use of both terminologies.

The GLSE can be transformed into OLSE by a whitening procedure (Koch 1999, p. 154; Schuh et al. 2014) using a Cholesky factorization of the cofactor matrix where $\mathbf{W}^{-1} = \mathbf{G}^{\mathbf{T}}\mathbf{G}$, $\mathbf{G}$ being a regular lower triangular matrix. Using the transformations $\mathbf{A}_{\mathbf{w}} = \mathbf{G}^{\mathbf{T}}\mathbf{A}, \mathbf{y}_{\mathbf{w}} = \mathbf{G}^{\mathbf{T}}\mathbf{y}, \boldsymbol{\varepsilon}_{\mathbf{w}} = \mathbf{G}^{\mathbf{T}}\boldsymbol{\varepsilon}$, Eq. 1 can be written: $\mathbf{y}_{\mathbf{w}} = \mathbf{A}_{\mathbf{w}}\boldsymbol{x} + \boldsymbol{\varepsilon}_{\mathbf{w}}$, with $E(\boldsymbol{\varepsilon}_w\boldsymbol{\varepsilon}_w^{\mathbf{T}}) = \sigma^2\mathbf{I}$.

Where $\mathbf{I}$ is the identity matrix. Thus, the observations have been transformed to be uncorrelated with equal variance, corresponding to the OLSE. For a concrete example of the use of whitening via Cholesky factorization, see exemplary Alkhatib and Schuh (2007). Following the same procedure, an eigenvalue decomposition of $\mathbf{W}^{-1}$ can alternatively be used. However, whitening is not always an optimal procedure, leading to inaccuracy for ill-conditioned Gaussian autocorrelation function (Koivunen and Kostinski 1999). Moreover, both the Cholesky and the eigenvalue decomposition are computational demanding ($0(n^3)$, Trefethen and Bau 1997), particularly for large matrix.

In both cases (whitened or not), an a priori knowledge of the covariance matrix is needed, or at least an estimation based for instance on the residuals. For more details on GLS, interested readers are referred exemplarily to Krämer (1986); Krämer and Donninger (1987), Baksalary (1988), works of Puntanen (1987), Puntanen and Styan (1989), Koch (1999), Rao and Toutenburg (1999), Niemeier (2008) or Grafarend and Awange (2012).

2.2 Equivalence between diagonally weighted least squares (DWLS) and general least squares (GLS)

In the next section, $\mathbf{W}_F$ denotes the fully populated cofactor matrix of full rank and $\mathbf{W}_E$ the equivalent diagonal cofactor matrix. The conditions under which the least squares estimate with $\mathbf{W}_E$ is identical to the one given by $\mathbf{W}_F$ are given in Luati and Proietti (2011) who showed that:

DWLS and GLS estimators are equivalent if and only if the $n \times u$ $\mathbf{A}$ matrix ($u < n$) can be decomposed as

$$\mathbf{A} = \mathbf{V}^{*}\mathbf{M} \tag{6}$$

where the u columns of $\mathbf{V}^{*}$ are eigenvectors of $\mathbf{W}_F\mathbf{W}_E^{-1}$ and $\mathbf{M}$ is a non singular matrix.

This theorem can be seen as a generalization of Zyskind (1967) results for the equivalence between OLS and GLS.

The diagonally weighted least squares can, therefore, be equivalent to the classical least squares if there are p-linear combinations of the columns of the design matrix that are eigenvectors of $\mathbf{W}_F\mathbf{K}$, where $\mathbf{K}$ is a diagonal matrix which elements provide the optimal kernel weights corresponding to $\mathbf{W}_F$.

A necessary and sufficient condition of equivalence between the GLS estimator and the DWLS estimator is that $\mathbf{W}_E\mathbf{W}_F^{-1}\mathbf{H} = \mathbf{H}\mathbf{W}_E\mathbf{W}_F^{-1}$ where $\mathbf{H}$ is a projection matrix onto the column space of $\mathbf{A}$ along the nullspace of $\mathbf{A}^{\mathbf{T}}\mathbf{W}_E^{-1}$.

In the next section, we will concentrate on particular cases of this theorem that could be applied to GPS positioning and assume the cofactor matrix $\mathbf{W}_F$ to be known.

3 Equivalence DWLS–GLS: particular cases and simulations

3.1 Computation of the equivalence matrix $\mathbf{W}_E$

In the general case of linear regression, Luati and Proietti (2011) showed that the previously defined matrix $\mathbf{M}$ can be chosen to be a sparse upper triangular matrix. Consequently, the linear combinations of the columns of the design matrix are provided by $\mathbf{AM^{-1}}$.

In the case of the mean estimator, the degree of the fitting polynomial is 0 and the particular design matrix $\mathbf{A} = \mathbf{1}$ which allows to choose a scalar matrix $\mathbf{M}$. Consequently, in this case, a necessary and sufficient condition for the equivalence between GLS and DWLS is that each element of the diagonal matrix $\mathbf{W}_E^{-1}$ is the sum of the row elements of the inverse of the fully populated cofactor matrix $\mathbf{W}_F^{-1}$ (Luati and Proietti 2011).

It should be pointed out that in a more general case of local polynomial regression, where only the first row of $\mathbf{A}$ is the vector $\mathbf{1}$, a necessary condition for $\mathbf{W}_E$ to yield equivalent results compared to the GLS is $\mathbf{W}_E^{-1}\mathbf{1} \propto \mathbf{W}_F^{-1}\mathbf{1}$, i.e., the diagonal elements of $\mathbf{W}_E^{-1}$ are proportional to the sum of the row elements of $\mathbf{W}_F^{-1}$ (Luati and Proietti 2011).

Figure 1 highlights graphically how to compute the equivalent cofactor matrix for the mean estimator case. Compared with the whitening procedure presented in 2.1, the use of an equivalent diagonal matrix, which can be seen as a reduced version of $\mathbf{W}_F$ valid in the least squares adjustment, implies neither a change of the least squares model nor a Cholesky transformation to obtain the correct estimates. The procedure is, therefore, more simple as the equivalent matrix gives mathematically the same estimate for the case $\mathbf{A} = \mathbf{1}$ compared with $\mathbf{W}_F$. However, it should be pointed out that the equivalent matrix will not whiten the observations.

Due to this equivalence, the estimates, their cofactor matrix as well as the residual vector are identical. However, the results are not identical for the estimator of the a posteriori variance factor of the observations $\hat{\sigma}_0^2 = \frac{\mathbf{v^T W^{-1} v}}{n-u}$ since the two cofactor matrices $\mathbf{W}_E$ and $\mathbf{W}_F$ have different structures. Indeed, $\frac{\mathbf{v^T W_F^{-1} v}}{n-u} = \frac{\mathbf{v^T W_E^{-1} v}}{n-u}$ would be only valid if $\mathbf{v^T W_F^{-1}} = \mathbf{v^T W_E^{-1}}$. The latter equation cannot be proven for an arbitrary residuals vector $\mathbf{v^T}$.

Closed form for the inversion of some covariance matrices

AR(1) process

As highlighted by Fig. 1, the inverse of the cofactor matrix has to be computed to determine the equivalent matrix. For some particular cases, such as the autoregressive process of order 1, an explicit inverse $\mathbf{W}_{AR(1)}^{-1}$ exists (Rao and Toutenburg 1999; Bierman 1977), provided that the autocorrelation coefficient ρ is known or estimated.

$$\mathbf{W}_{AR(1)}^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho & 0 & \dots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \ddots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix} \tag{7}$$

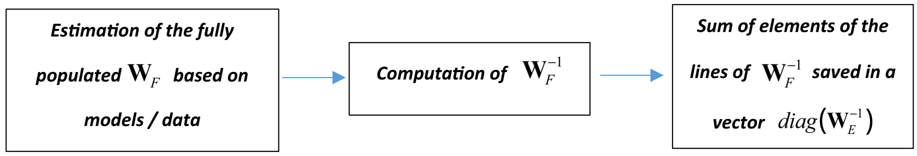


Fig. 1 Computation of $\mathbf{W}_E$

This greatly simplifies the computation of the equivalence matrix since the sum of the elements can be directly computed leading to:

$$\mathbf{W}_{AR(1)\_EQUI}^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1-\rho & 0 & 0 & \dots & 0 & 0 \\ 0 & (1-\rho)^2 & 0 & \ddots & 0 & 0 \\ 0 & 0 & (1-\rho)^2 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & (1-\rho)^2 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1-\rho \end{bmatrix}$$

Toeplitz matrices

Toeplitz matrices (Gray 2006) arise in the theory of weakly stationary processes where the main challenge is often to solve linear systems by inverting large matrices. Thus, since 1960, many authors have proposed two main ways to improve the classical Gaussian elimination methods or Cholesky decomposition which involves $0(n^3)$ operations by computing either the inverse of the Toeplitz matrix or its factorization (Levinson 1947; Durbin 1960; Trench 1964) or by factorizing the Toeplitz matrix directly (Barciss 1969). Algorithms using the fast Fourier transform belong to both classes and use only $0(n^2)$ (Brent et al. 1980) or more recently $0(n(\log n)^2)$ operations (Ammar and Gragg 1988). Other references and detailed procedure on inversion algorithms of Toeplitz matrices can be found in the study by Brent (1989).

In geodesy, such matrices are found exemplarily to compute the anomalous potential (Moritz 1980). Bottoni and Barzaghi (1993) proposed an algorithm based on the fast Fourier transform and the preconditioned conjugate gradient method. Thus, possibilities exist to speed up the inversion of Toeplitz matrices and can be used to compute the equivalent matrices in less operations.

For all other types of VCM matrices, we propose to invert and sum the VCM outside the program package and to reintroduce the diagonal weight matrix.

3.2 Case study: mean estimator with a Toeplitz covariance matrix

To have a better understanding of the equivalent model and its implication, an illustrative example was simulated. A **1**-vector design matrix is chosen to explain the general results of Luati and Proietti (2011) selecting a Toeplitz covariance matrix and a corresponding correlated observation vector. Using different weight matrices in the least squares adjustment, the equivalence is shown.

Stochastic model

The computation of the elements of the covariance matrix is based on the very general Mátern covariance family (Appendix) and reads

$$W_F(i,j) = \phi(\alpha\,|i-j|)^{\nu} K_{\nu}(\alpha\,|i-j|), \tag{8}$$

where $\nu > 0$, $\alpha > 0$ are constant parameters. The scalar parameter $\phi > 0$ is chosen so that the variance equals 1; ν can be seen as a measure of the differentiability of the field (Stein 1999) and is defined as "its smoothness". The constant α is the inverse of the Mátern correlation time (ICT) and indicates how the correlations decay with increasing distance or time (Journel and Huifbregts 1978). The modified Bessel function of order ν (Abramowitz and Segun 1972) is denoted by K_ν. Although no explicit formula for the inversion of Mátern covariance matrices is known, this covariance family was chosen because of its flexibility and concrete application in GPS computation as shown by Kermarrec and Schön (2014). The simulated covariance matrices are scaled to have all diagonal elements equal 1.

Functional model

The design matrix is $\mathbf{A} = \mathbf{1}$, thus a so-called "mean estimator". However, it is worth being aware that the result of the least squares procedure is the mean of the time series only for the homoscedasticity case, thus $\mathbf{W} = \mathbf{I}$.

Observations

An eigenvalue decomposition $\mathbf{W}_F = \mathbf{USU}^T$ of $\mathbf{W}_F$ was used to generate the observation time series

$$\mathbf{y}(t) = \mathbf{U}\sqrt{\mathbf{S}}\mathbf{n}(t) \tag{9}$$

where $\mathbf{n}(t) = N(\mathbf{0}, \mathbf{I})$ is a random vector of zero mean and with a standard deviation of 1. The ith column vector of $\mathbf{U}$ corresponds to the ith eigenvector of $\mathbf{W}_F$ and the diagonal element of the diagonal matrix $\mathbf{S}$ to the corresponding eigenvalues (Cressie 1993; Vennebusch et al. 2010). 10,000 simulations were run for a stationary time series having the Toeplitz covariance matrix of size $n = 1000$, the Mátern parameters used being $[\alpha_0, \nu_0] = [0.05, 1]$. It should be pointed out that due to the correlation structure, the simulated times series (as for instance depicted in Fig. 2a) are

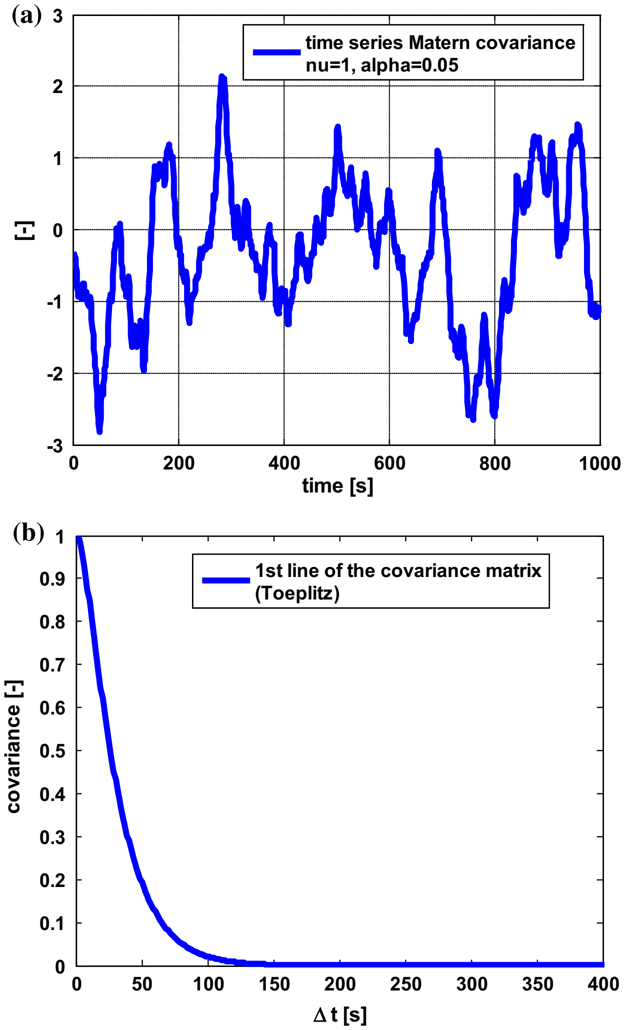


Fig. 2 **a** Correlated time series used for the mean estimator. The Mátern parameters are $[\alpha_0, \nu_0] = [0.05, 1]$. **b** The corresponding 400 first values of the line of the Toeplitz covariance matrix

neither zero mean nor have a standard deviation of 1. Only the underlying random vector has this property. Figure 2a, b presents, respectively, the time series used for the simulation which results are given in Table 1, as well as the first 400 values of the a line of the Toeplitz covariance matrix (Eq. 8).

Least squares estimation

Three kinds of least-squares cofactor matrices $\mathbf{W}$ are compared to study their influence on the least squares results:

- COV: the correct fully populated cofactor matrix $\mathbf{W}_F$, which was used for the computation of the time series
- EQUI: the equivalent diagonal kernel matrix $\mathbf{W}_E$ as presented in Sect. 3.1, see also Fig. 1
- IDEN: the identity covariance matrix

Table 1 Comparison of different covariance matrices models for the mean estimator with correlated data $[\alpha_0, \nu_0] = [0.05, 1]$

	$\hat{\mathbf{x}}$	$V(\hat{\mathbf{x}})$	$\hat{\sigma}_0$
COV	−0.3412	0.0593	0.9990
EQUI	−0.3412	0.0593	0.1256
IDEN	−0.3278	0.001	0.9921

Following estimators are compared:

- The estimate itself, i.e., the least-squares solution: $\hat{\mathbf{x}} = (\mathbf{A^T W^{-1} A})^{-1}\mathbf{A^T W^{-1} y}$,
- The a priori variance covariance matrix of the unknowns, $V(\hat{\mathbf{x}}) = \sigma^2(\mathbf{A^T W^{-1} A})^{-1}$,
- The a posteriori variance factor $\hat{\sigma}_0^2 = \frac{(\mathbf{y}-\mathbf{A}\hat{\mathbf{x}})^\mathbf{T}\mathbf{W}^{-1}(y-A\hat{\mathbf{x}})}{n-1} = \frac{\mathbf{v^T W^{-1} v}}{n-1}$.

To highlight how the EQUI matrices are built, we define two simple 3 × 3 covariance matrices which elements are given by:

$$\mathbf{W}_F^{-1} = \begin{bmatrix} p_{F11} & p_{F12} & p_{F13} \\ p_{F12} & p_{F22} & p_{F23} \\ p_{F13} & p_{F23} & p_{F33} \end{bmatrix}$$

and $\mathbf{W}_E^{-1} = \begin{bmatrix} p_{E11} & 0 & 0 \\ 0 & p_{E22} & 0 \\ 0 & 0 & p_{E33} \end{bmatrix}$, where $\mathbf{W}_E^{-1}$ and $\mathbf{W}_F^{-1}$ are the inverses of the equivalent matrix EQUI and COV, respectively.

The estimate equality we want to achieve reads:

$$\hat{\mathbf{x}} = (\mathbf{A^T W_F^{-1} A})^{-1}\mathbf{A^T W_F^{-1} y} \\ = (\mathbf{A^T W_E^{-1} A})^{-1}\mathbf{A^T W_E^{-1} y},$$

which is true if $\mathbf{A^T W_F^{-1}} = \mathbf{A^T W_E^{-1}}$.

For the mean estimator, this can be further written as:

$$[1 \quad 1 \quad 1] \begin{bmatrix} p_{F11} & p_{F12} & p_{F13} \\ p_{F12} & p_{F22} & p_{F23} \\ p_{F13} & p_{F23} & p_{F33} \end{bmatrix} = [1 \quad 1 \quad 1] \begin{bmatrix} p_{E11} & 0 & 0 \\ 0 & p_{E22} & 0 \\ 0 & 0 & p_{E33} \end{bmatrix}.$$

Thus, the elements of $\mathbf{W}_E^{-1}$ are simply:

$$p_{E11} = p_{F11} + p_{F12} + p_{F13}, \\ p_{E22} = p_{F12} + p_{F22} + p_{F23}, \\ p_{E33} = p_{F13} + p_{F23} + p_{F33}$$

which show concretely how the EQUI matrices are formed.

Results

The results of the simulations are presented in Table 1.

As expected from the results of Luati and Proietti (2011), the values of $V(\hat{\mathbf{x}})$ (Table 1—second column) and $\hat{\mathbf{x}}$ (Table 1—first column) are exactly identical for the EQUI and COV models, whereas the values obtained with IDEN, which correspond to the arithmetic mean value of the times series, are different.

As can be seen in Table 1—third column, the a posteriori variance factor for the EQUI case is much lower than the COV case although the estimate and variance of the unknown are identical in both models.

As our interest is in the potential empirical application of the equivalent model in GPS, no further developments are carried out to correct the a posteriori variance factor in the mean estimator case since no mathematical formulation of the inverses in the general case can be given.

4 Empirical extensions of the mean estimator results to further classes of design matrices

4.1 Motivation

As shown in Sect. 3, the mathematical equivalence is mainly based on the special structure of the design matrix $\mathbf{A} = \mathbf{1}$, the mean estimator case. Additional simulations were carried out with the same time series as presented in Fig. 2 using a design matrix filled with random numbers. As expected no equivalence for the estimate was given anymore, highlighting the requisite property of the design matrix (unit vector) for the mathematical equivalence.

For the particular case of GPS least squares adjustment and thanks to the slowly varying GPS constellation, the values of the columns of the design matrix sorted per satellite vary linearly for short observation spans of up to 600 s. Absolute values of the slope are between $5e^{-3}$ for low-elevation satellites and $5e^{-2}$ for high-elevation satellites (Schön and Kutterer 2006). Thus, the condition of the design matrix for the equivalence presented in Sect. 3 is quite well approximated. Let

$$\mathbf{A}^i = \begin{bmatrix} -\frac{\Delta x}{\rho_i}\frac{t_0}{t_0} & -\frac{\Delta y}{\rho_i}\frac{t_0}{t_0} & -\frac{\Delta z}{\rho_i}\frac{t_0}{t_0} \\ \vdots & \vdots & \vdots \\ -\frac{\Delta x}{\rho_i}\frac{t_n}{t_n} & -\frac{\Delta y}{\rho_i}\frac{t_n}{t_n} & -\frac{\Delta z}{\rho_i}\frac{t_n}{t_n} \end{bmatrix} \tag{10}$$

denotes the design matrix of the ith satellite sorted by epochs, where Δx, Δy, Δz are coordinate differences between the unknown ECEF coordinates of the observing site and the satellite i and ρ_i the geometric distance between the satellite i at epochs between t_0 and t_n. The whole design matrix reads: $\mathbf{A} = \begin{bmatrix} \mathbf{A}^1 \\ \vdots \\ \mathbf{A}^i \\ \mathbf{A}^m \end{bmatrix}$ where m is the number of visible satellites.

Since the strict equivalence as in Sect. 3.1. cannot be proven for GPS design matrices, simulations are used to motivate that the equivalent covariance matrices can also be applied for this type of matrices, giving more adequate results for absolute positioning than diagonal elevation-dependent covariance matrices. We start with simple cases and add step by step more elaborated design matrices. For illustration purposes, here, the minimum configuration with 4 satellites per epoch is considered for 1000 s with 1Hz data rate.

4.2 Simulations methodology

Design matrices

A design matrix (Eq. 10) for point positioning with 4 satellites is simulated, corresponding to typical short static positioning scenarios. Figure 3a shows the columns of the resulting design matrix of size $(4 \times 1000) \times 3$, sorted per satellite. The corresponding skyplot is presented in Fig. 3b. The design matrix corresponding to the geometry of the first station reads:

$$\mathbf{A}_1 = \begin{bmatrix} \mathbf{A}_1^1 \\ \mathbf{A}_1^2 \\ \mathbf{A}_1^3 \\ \mathbf{A}_1^4 \end{bmatrix}.$$

Using this particular design matrix, the resulting estimates of the least squares adjustment are corresponding to Earth Center Earth Fixed (ECEF) coordinates (Hoffmann-Wellenhof et al. 2001, chapter 8).

The feasibility of using the EQUI model (see Fig. 1 for how to build the corresponding matrix) is tested by changing the positioning mode resulting in different design matrices. Thus, to check the equivalence, five GPS positioning scenarios are simulated, the design matrix being gradually changed. Each time a slightly different matrix than the simple slope design matrices of Eq. 10 by adding tropospheric parameter estimation or clock estimation as well as double differencing is obtained.

Single point positioning (SP)

Besides the classic static single point positioning with epochwise clock estimation (SPEC), two other simulation cases are studied resulting from different clock modeling (Weinbach 2012):

- Single point positioning without clock parameter (SPWC)
- Single point positioning with one clock parameter (SP1C) for the observation interval

The design matrix is changed accordingly by adding an identity vector (SP1C) or identity matrices (SPEC) in the last columns, respectively.

- for the SPWC case: $\mathbf{A}_{\mathrm{SPWC}} = [\mathbf{A}_1]$
- for the SPEC case: $\mathbf{A}_{\mathrm{SPEC}} = [\mathbf{A}_1 \;\; \mathbf{MI}_{\mathrm{SPEC}}]$, where $\mathbf{MI}_{\mathrm{SPEC}} = \begin{bmatrix} \mathbf{I} \\ \mathbf{I} \\ \mathbf{I} \\ \mathbf{I} \end{bmatrix}$ and $\mathbf{I}$ the identity matrix of size $n \times n$,

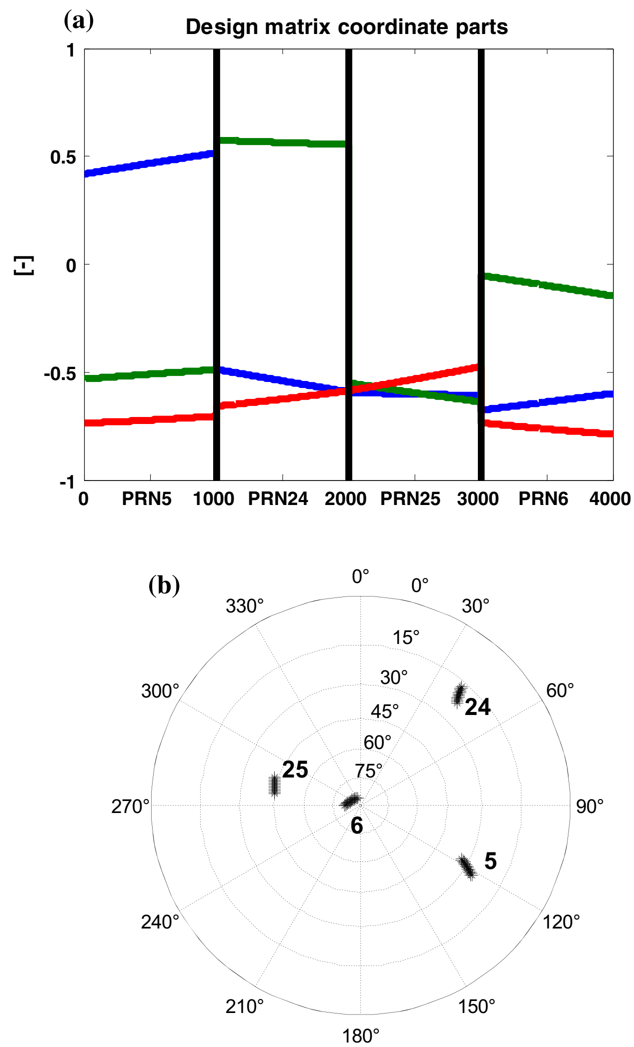


Fig. 3 **a** 3 columns (*blue* 1st column, *green* 2nd column and *red* 3rd column of the design matrix for the 3 coordinates, **b** the corresponding skyplot

with $n = 1000$, all satellites being present for the n epochs.

- for the SP1C case: $\mathbf{A}_{\text{SP1C}} = [\mathbf{A_1} \quad \mathbf{MI}_{\text{SP1C}}]$, where $\mathbf{MI}_{\text{SPEC}} = \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{bmatrix}$ and $\mathbf{1}$ the ones vector of size $n \times 1$

Precise point positioning (PPP)

The precise point positioning case is simulated using code and carrier phase measurements. The design matrix is adapted as follows: for the phase design matrix, a tropospheric parameter based on a simple mapping function $\frac{1}{\sin(El(t))}$ where $El(t)$ the elevation of the satellite varying with time t is estimated as well as epoch-wise clock parameters and one ambiguity per satellite. For the code design matrix, no ambiguity parameter has to be taken into account and only an epoch-wise clock and tropospheric parameters are computed.

Following the previous notation, the design matrix reads:

- for the PPP case: $\mathbf{A}_{\text{PPP}} = \begin{bmatrix} \mathbf{A}_{\text{phase}} \\ \mathbf{A}_{\text{code}} \end{bmatrix}$, with

$$A_{\text{phase}} = \begin{bmatrix} \mathbf{A}_1^1\mathbf{T}_1^1 & \overbrace{\begin{matrix}\mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0}\end{matrix}} & \overbrace{\mathbf{I}} \\ \mathbf{A}_1^2\mathbf{T}_1^2 & \begin{matrix}\mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0}\end{matrix} & \mathbf{I} \\ \mathbf{A}_1^3\mathbf{T}_1^3 & \begin{matrix}\mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0}\end{matrix} & \mathbf{I} \\ \mathbf{A}_1^4\mathbf{T}_1^4 & \underbrace{\begin{matrix}\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}\end{matrix}}_{\text{Ambiguity}} & \underbrace{\mathbf{I}}_{\text{clock}} \end{bmatrix}, \mathbf{A}_{\text{code}} = \begin{bmatrix} \mathbf{A}_1^1\mathbf{T}_1^1 \mathbf{0}\,\mathbf{0}\,\mathbf{0}\,\mathbf{0}\,\mathbf{I} \\ \mathbf{A}_1^2\mathbf{T}_1^2 \mathbf{0}\,\mathbf{0}\,\mathbf{0}\,\mathbf{0}\,\mathbf{I} \\ \mathbf{A}_1^3\mathbf{T}_1^3 \mathbf{0}\,\mathbf{0}\,\mathbf{0}\,\mathbf{0}\,\mathbf{I} \\ \mathbf{A}_1^4\mathbf{T}_1^4 \mathbf{0}\,\mathbf{0}\,\mathbf{0}\,\mathbf{0}\,\mathbf{I} \end{bmatrix}$$

$\mathbf{T}_1^i$ being the partial derivatives with respect to the tropospheric vector parameter for satellite i.

Double difference case (DD)

The design matrix is mathematically double differenced (Beutler et al. 1986) using the fourth satellite PRN6 as the reference satellite—the one having the highest elevation over the considered time period—no ambiguities were considered here.

- for the DD case $\mathbf{A}_{\text{DD}} = \begin{bmatrix} \mathbf{A}_1^1 - \mathbf{A}_1^4 \\ \mathbf{A}_1^2 - \mathbf{A}_1^4 \\ \mathbf{A}_1^3 - \mathbf{A}_1^4 \end{bmatrix}$, holding the coordinate of station 2 fixed.

Cofactor matrix $\mathbf{W}_F$

For this simulation, we used the result of Kermarrec and Schön (2014), which is an easy and physically relevant way to compute temporal correlations for GPS. The model is based on the Mátern covariance family (Appendix), scaled with an elevation-dependent factor. The covariance function C between two observations at time t and $t + \tau$ reads:

$$C_{it}^{it+\tau} = \frac{\delta}{\sin(El_i(t))\sin(El_i(t+\tau))}(\alpha\tau)^{5/6}K_{5/6}(\alpha\tau) \quad (11)$$

where El_i is the elevation of the satellite i, α and δ parameters derived from turbulence theory. The smoothness factor is taken to $\nu = 5/6$. For this example, a Mátern ICT of $\alpha = 0.05\ \text{s}^{-1}$ is chosen. Please refer to Kermarrec and Schön (2014) for more details on this model. No correlation between different satellites is taken into account, thus the resulting covariance matrix has a block diagonal structure. Since a scaling is not influencing the least squares results (Kutterer 1999), the cofactor matrix is computed.

Figure 4a is a representation of the covariance matrix used for the simulations. The elements are sorted per satellite and based on actual elevation.

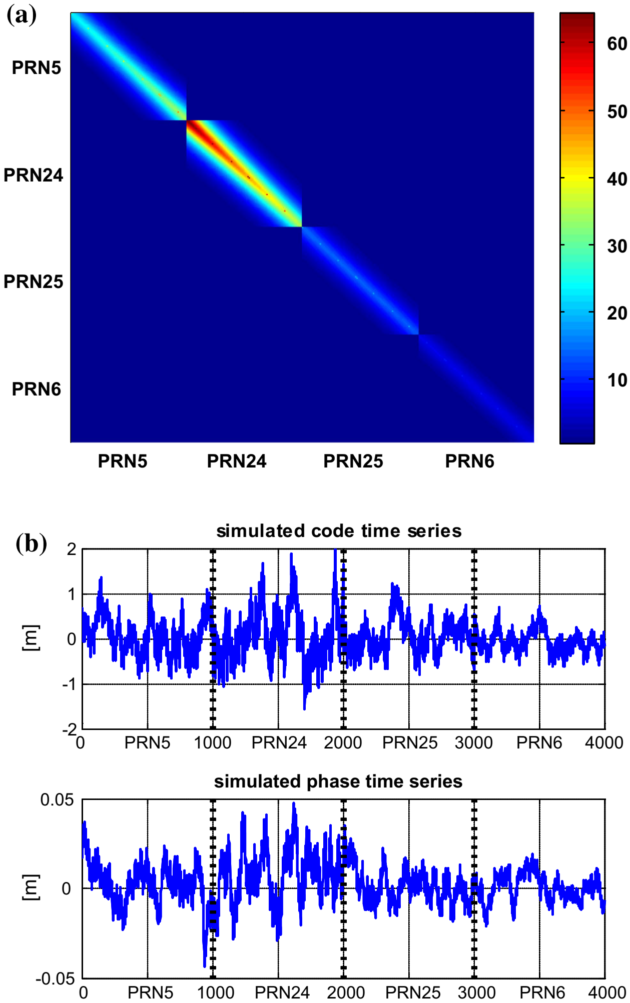


Fig. 4 **a** Cofactor matrix sorted per satellite (mm^2) computed with Eq. 12. The elements are based on the model by Kermarrec and Schön (2014). **b** Corresponding code and phase time series [m], sorted per satellites. 1000 epochs per satellite are generated

To be more realistic, an elevation-dependent noise matrix is added to the fully populated cofactor matrix, thus the global cofactor matrix of this model reads:

$$\mathbf{W}_F = f(\theta \mathbf{W}_{\text{matern}} + \beta \mathbf{\Gamma}), \qquad (12)$$

where $\mathbf{W}_{\text{matern}}$ is the cofactor matrix which elements are computed with Eq. 11 and $\mathbf{\Gamma}$ the diagonal elevation-dependent cofactor matrix which elements are $\gamma_{ii} = \frac{1}{\sin^2(El(i))}$. β, θ are positive factors taken as $\beta = 0.2, \theta = 0.8$ for these simulations. The factor f is taken as $f_{\text{phase}} = 2$ for phase measurements and $f_{\text{code}} = 2000$ for code measurement simulations. Subsequently, the variation of the magnitude of observations corresponds to typical values of code or phase residuals, respectively. It should be noted that the results for the empirical equivalence are not depending on these values.

For all the five scenarios, the cofactor matrix of the original observations for station 1 reads:

$$\mathbf{W}_{F1} = f\left(\theta\begin{bmatrix} \mathbf{C}_1^{1,1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_1^{2,2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_1^{3,3} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C}_1^{4,4} \end{bmatrix} + \beta\mathbf{\Gamma}\right), \text{ with}$$

$$\mathbf{C}_1^{i,j} = \begin{bmatrix} C_{i1}^{j1} & C_{i1}^{j2} & C_{i1}^{j3} & \ldots & C_{i1}^{jn} \\ C_{i2}^{j1} & C_{i2}^{j2} & C_{i2}^{j3} & & C_{i2}^{jn} \\ & \ldots & \ldots & & \\ C_{in}^{j1} & & & & C_{in}^{jn} \end{bmatrix}, \text{ computed with Eq. (11).}$$

$\mathbf{W}_{F1}$ can be either $\mathbf{W}_{F1,\text{code}}$ corresponding to $f = f_{\text{code}}$ or $\mathbf{W}_{F1,\text{phase}}$ computed with $f = f_{\text{phase}}$.

Single point positioning (SP)

For the three single point positioning cases, $\mathbf{W}_{F,\text{SP}} = \mathbf{W}_{F1,\text{code}}$ is the cofactor matrix of the undifferenced code observations.

Precise point positioning (PPP)

For the PPP case, $\mathbf{W}_{F,\text{PPP}} = \begin{bmatrix} \mathbf{W}_{F1,\text{phase}} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{F1,\text{code}} \end{bmatrix}$.

Double difference case (DD)

For the double difference positioning case, $\mathbf{W}_{F\text{DD}}$ is computed using the double difference operator:

$\mathbf{W}_{F,\text{DD}} = \mathbf{D}\begin{bmatrix} \mathbf{W}_{F1,\text{phase}} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{F2,\text{phase}} \end{bmatrix}\mathbf{D}^{\mathbf{T}}$, where $\mathbf{W}_{F2,\text{phase}}$ is the cofactor matrix of the undifferenced phase observations of station 2. The correlations between two different satellites as well as between two stations are neglected. For our scenario

$$\mathbf{D} = \begin{bmatrix} \mathbf{I} & \mathbf{0}\,\mathbf{0} & -\mathbf{I}\,-\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I}\,\mathbf{0} & -\mathbf{I}\;\;\mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0}\,\mathbf{I} & -\mathbf{I}\;\;\mathbf{0} & \mathbf{0} & -\mathbf{I} & \mathbf{I} \end{bmatrix}.$$

Simulating time series

As in Sect. 3.2, an eigenvalue decomposition of the VCM was used to generate the corresponding observation time series presented in Fig. 4b, top panel for the code and Fig 4b, bottom panel for the phase observations, respectively. For all simulations, the same time series were used and only the design matrix was adjusted to the different cases. The DD time series are corresponding to phase measurements, whereas SP corresponds to code and PP use both code and phase simulations.

As can be seen in Eq. 12, in the covariance function, the factor $\frac{1}{\sin^2(El(t))}$ where $El(t)$ is the elevation of the satellite is varying with time t. As a consequence, the resulting covariance matrix is not having a Toeplitz structure as seen in Fig. 4a and the covariance matrix has to be factorized. The Mátern kernel of Eq. (11) leading to a Toeplitz covariance matrix, the cofactor matrix can be rewritten as $\mathbf{W}_F = (\mathbf{\Gamma}^{1/2})\boldsymbol{T}(\mathbf{\Gamma}^{1/2})$

where $\boldsymbol{\Gamma}$ is the diagonal matrix containing the elevation dependency and T a Toeplitz matrix. A way to invert efficiently such block diagonal Toeplitz matrices is described exemplarily in Ataike (1973).

In Sect. 3.2, the impact of three kinds of covariance matrices, namely COV, EQUI, and IDEN were analyzed. In this section, the feasibility of using the EQUI model instead of the COV model in other cases than the mean estimation is our main objective. For sake of completeness, the results with the often used elevation-dependent diagonal covariance matrix are additionally indicated (model ELEV). Here, the diagonal elements of the covariance matrix correspond to the COV matrix.

4.3 Simulations results

The results of the simulation are presented in Table 2 for the Single Point Positioning as well as for PPP and relative positions with DD. The differences of the estimates $\hat{\mathbf{x}} - \hat{\mathbf{x}}_{\text{COV}}$, where $\hat{\mathbf{x}}$ are the estimates computed with the covariance matrix EQUI or ELEV are presented. The nominal values of the estimates are obtained with the COV model which corresponds to the least squares with fully populated VCM. Thus, the coordinates themselves are not given here since we focus on quantifying the empirical equivalence.

Table 2 (first part) highlights the feasibility of taking the EQUI model instead of the COV model in the least square adjustment for the Single Point Positioning cases with different clock modeling. Although the design matrix is gradually deviating from the simple slope case, differences between the EQUI and COV models below 0.2 mm are obtained for the X and Y components and below 1mm for the Z component (cf. SP1C and SPEC cases). A slight increase of the deviations with more complex models such as SPEC can be stated (difference of −1.19 mm for the Z component compared with the −0.10 mm value of the SPWC model).

On the contrary, it should be pointed out that the differences between the ELEV and COV models are larger than 10 cm for the Z component for the SP1C and SPEC models; smaller differences of 9 mm for the X component being obtained for the SPWC model. The values are all below the typical noise of the code observations of 30 cm^{-1} m. These results underline the necessity to take correlations into account to have estimates closer to the nominal values, which can be achieved when using the EQUI model.

Table 2 (second part) shows the results for positioning with carrier phase observations, namely PPP and relative positioning with DD. The performance of the equivalence model EQUI is high; the coordinate differences are smaller than 1 mm for PPP. The relative positioning with DD yields differences of the estimates below the sub-mm level although the complexity of the design matrix is increasing. These results are not comparable with the values obtained with the ELEV case which are reaching up to 11 mm (*X* value of the PPP case) or more than 2 mm for the DD case for the Z component.

As a consequence, both the feasibility of using the EQUI model in least squares adjustment corresponding to GPS scenarios as well as its superiority of the EQUI model with respect to the ELEV model are pointed out for all five simulation studies.

The presented results depend not only on the covariance model used but also on the simulated time series which is governed by the random vector (cf. Eq. 9). To validate the previous results, Monte Carlo simulations with 1000 runs with different random vectors and thus different observation time series with the same correlation structure were carried out. From this, the same order of agreement was found.

Table 2 Differences of the estimates (mm) for the EQUI model and ELEV model w.r.t the nominal solution obtained with the COV model, reference set $\alpha_0, \nu_0 = [0.05, 5/6]$

	SPWC		SP1C		SPEC	
	EQUI	ELEV	EQUI	ELEV	EQUI	ELEV
Differences of the estimates (mm)	0.17	−18.13	−0.06	−10.03	0.03	−8.98
	−0.07	−32.53	−0.25	−18.91	−0.17	−16.21
	−0.10	27.41	−1.12	107.52	−1.19	107.30
	PPP		DD			
	EQUI	ELEV	EQUI	ELEV		
Differences of the estimates (mm)	0.66	11.18	−0.00	0.10		
	0.11	−6.36	−0.01	1.21		
	−0.30	−8026	−0.06	2.00		

Single point positioning with 4 satellites, precise point positioning and relative positioning with double difference for 4 satellites

Table 3 Standard deviation [m] of the estimates—COV, EQUI and ELEV models

Standard deviation of the estimates (m) Case 4 satellites	SPWC	SP1C	SPEC	PPP	DD
COV					
X	0.10	0.11	0.11	0.07	0.0044
Y	0.13	0.15	0.15	0.10	0.0058
Z	0.10	0.48	0.48	0.07	0.0187
EQUI					
X	0.10	0.11	0.11	0.07	0.0044
Y	0.13	0.15	0.15	0.10	0.0058
Z	0.10	0.48	0.48	0.07	0.0188
ELEV					
X	0.01	0.02	0.02	0.01	0.0006
Y	0.02	0.02	0.02	0.01	0.0008
Z	0.01	0.07	0.07	0.01	0.0026

The reference set $\alpha_0, \nu_0 = [0.05, 5/6]$ was used. The minimum constellation with four satellites is considered

Cofactor matrices of the estimates

Table 3 gives the standard deviations of the estimates in [m] for the COV, EQUI and ELEV model for the three single point positioning cases, the PPP and relative positioning with DD. In all cases, the ELEV model yields approximately one order of magnitude smaller standard deviation than the models taking correlations into account. It highlights the known over-optimistic precision of the ELEV model (El-Rabbany 1994; Radovanovic 2001). A factor up to 10 between the ELEV and the COV models corresponds to previous findings with real case studies (Dach et al. 2007). It should be underlined that the COV and EQUI models give quasi-identical results for all simulated cases, the values being only different at the sub-mm level. Please note that other combination of the covariance matrices (Eq. 11) would have given different numerical values, however, without changing the basic statement on the empirical equivalence. This is also true for the coordinate differences.

Discussion

Further simulations were carried out with up to eight visible satellites as well as different satellites constellations and longer/shorter time spans. The results again reinforce the conclusions on the empirical equivalence between the EQUI and COV models. In addition, the Mátern parameters were successfully changed by simulating strong or weak observation correlations, without impacting the empirical equivalence. However, time spans have to be chosen in such way that the correlation function reaches the zero value inside the batch, otherwise no equivalence can be given anymore.

The results of all these simulations highlight that the EQUI model is empirically a good approximation of the correct estimates computed with the COV reference model, although no mathematical equivalence exists. Thus, using a diagonal matrix that takes correlations into account in a GPS least square adjustment is feasible. In the next part, a short case study with real data will confirm this statement.

4.4 Case study with EPN data

For the case study with real data, we assume that the estimated covariance matrices COV correspond to Eqs. 11 and 12 with a Mátern ICT of $\alpha = 0.05$ s^{-1}. Based on physical considerations, this model is a good approximation of the correlation structure of GPS observations (i.e., the observations are assumed to be uncorrelated after approximately 600 s following Schön and Brunner 2007), being at the same time easily computable. We do not aim here to study the influence on least squares results of different covariance matrices taking correlation into account.

Data from the EPN network (Bruyninx et al. 2012) are used to confirm the quasi-equivalence of the EQUI with the COV model for relative positioning with double differences. Two stations Zimmerwald 1 and 2 (ZIMM and ZIM2) in Switzerland, close to each other are chosen. The data are having a 30 s rate, applying a cutoff of 5°. No ionospheric anomaly during the observation period were found. Some multipath is present which appears not to be problematic for our comparison. The ambiguities are assumed to be solved in advance and the reference values of the coordinates of the stations are from the RINEX file values which are in this case the long-term station coordinates. No tropospheric parameters are estimated and the CODE reprocessed orbits and clocks are used (Dach et al. 2009). An a priori standard deviation of 2 mm was assumed for the L1 carrier phase measurements.

The estimates are computed with double differences for 1000 batches with maximum 100 observations per satellite and per batch on GPS day 50 of 2015.

Table 4 Mean and standard deviation of the difference of the estimates $\Delta\hat{\mathbf{x}}_{\text{EQUI}-\text{COV}} = \hat{\mathbf{x}}_{\mathbf{EQUI}} - \hat{\mathbf{x}}_{\mathbf{COV}}$ and $\Delta\hat{\mathbf{x}}_{\text{ELEV}-\text{COV}} = \hat{\mathbf{x}}_{\mathbf{ELEV}} - \hat{\mathbf{x}}_{\mathbf{COV}}$ (mm) computed with double differences

DD real case	Mean	Standard deviation	Min/max
$\Delta\hat{\mathbf{x}}_{\text{EQUI}-\text{COV}}$			
X	0.0046	0.0286	−0.1103/0.0592
Y	−0.0050	0.0050	−0.438/0.0425
Z	−0.0201	0.1409	−0.0976/0.2542
$\Delta\hat{\mathbf{x}}_{\text{ELEV}-\text{COV}}$			
X	0.0292	0.3442	−1.1912/1.1450
Y	0.0134	0.2127	0.5332/0.7009
Z	0.0564	0.6636	−1.2653/4.0437

Real data from the EPN network. The reference set for COV is $\alpha_0, \nu_0 = [0.05, 5/6]$

Table 5 Standard deviation (mm) of the estimates for the last batch—COV, EQUI and ELEV models

DD real case	Standard deviation of the estimates
COV	
X	1.9692
Y	1.0568
Z	1.8408
EQUI	
X	1.9722
Y	1.0577
Z	1.8495
ELEV	
X	0.6390
Y	0.3419
Z	0.5966

Real data from the EPN Network computed with double differences. The reference set for COV is $\alpha_0, \nu_0 = [0.05, 5/6]$

The mean, the standard deviation as well as min–max values (mm) of the estimates difference $\Delta\hat{\mathbf{x}}_{\text{EQUI}-\text{COV}} = \hat{\mathbf{x}}_{\text{EQUI}} - \hat{\mathbf{x}}_{\text{COV}}$ and $\Delta\hat{\mathbf{x}}_{\text{ELEV}-\text{COV}} = \hat{\mathbf{x}}_{\text{ELEV}} - \hat{\mathbf{x}}_{\text{COV}}$ computed over all the batches are presented in Table 4. The typical standard deviation of the estimates (mm) is given for the last batch as example in Table 5. As expected from the previous simulations, the mean of $\Delta\hat{\mathbf{x}}_{\text{EQUI}-\text{COV}}$ is 6 times, 2.7, and 2 times smaller than for $\Delta\hat{\mathbf{x}}_{\text{ELEV}-\text{COV}}$ for the X, Y, and Z component, respectively. It highlights that the equivalence is not only valid for a simple case with four satellites as assumed for the simulations but also for more complex design matrices corresponding to a full geometry with different elevations and lengths. The standard deviation is 12, 42 and 5 times smaller for $\Delta\hat{\mathbf{x}}_{\text{EQUI}-\text{COV}}$ than for $\Delta\hat{\mathbf{x}}_{\text{ELEV}-\text{COV}}$.

Moreover, Table 5 shows once more the overestimation of the ELEV model, the values being more than 3 times higher than the COV model. The nearly perfect equivalence between EQUI and COV models is confirmed, the differences being below 0.01mm.

The analysis of real data with a relative GPS positioning using double differences has shown that the EQUI model can be used instead of the COV model for computing both the coordinates and the co factor matrix of the estimates in a least squares adjustment.

5 Conclusion

Neglecting correlations between GPS phase measurements leads to an over-optimistic precision, for instance for the GPS least squares coordinate solution. Due to the difficulty to estimate or model correlations, diagonal covariance matrices with an elevation dependency are widely preferred. Indeed, the use of fully populated matrices yields computational burden due to inversion or restriction in algorithmic implementation.

An equivalent diagonal covariance matrix based on the work of Luati and Proietti (2011) for the mean estimator is presented as a way to take correlations into account without using fully populated VCM in the GPS least squares procedure. The model gives mathematically identical results in terms of parameter estimates as well as a priori variance estimator for the mean estimator. An empirical extension of these findings to classes of design matrices with linearly varying column values was successfully proposed. Prominent examples are the design matrices such as used in GPS adjustments for absolute positioning (SPP, PPP), and relative positioning with single or double differences. Although no mathematical equivalence is given anymore, we showed that, thanks to the particular structure of the GPS design matrices, the estimates obtained with the equivalent diagonal covariance matrices approximate those obtained by the fully populated covariance matrices. From our simulations with four satellites for periods of up to 20 min, the errors are below 1mm for the Single Point Positioning cases and PPP. Relative positioning with double differences case yields nearly exact estimates, below the sub-mm level. These solutions outperform results with the commonly used elevation-dependent diagonal matrices, which was confirmed by the study of EPN network data. Furthermore, the over-optimistic parameter variances are remedied, yielding adequate values that reflect the actual parameter scatter. Since the proposed equivalent model can be implemented in GPS software packages, taking correlations into account should not be avoided in future. Further studies will concentrate on sensitivity analysis of the Mátern parameters with real data.

Acknowledgments Parts of the work were funded by the DFG under the label SCHO1314/1-2, this is gratefully acknowledged by the authors. The European Permanent Network and contributing agencies are thanked for providing freely GNSS data and products. Valuable comments of three anonymous reviewers helped us improve significantly the manuscript.

Appendix: Mátern covariance family

First introduced by Mátern (1960), the covariance function reads:

$$C(r) = \phi(\alpha r)^{\nu} K_{\nu}(\alpha r). \quad (13)$$

Other parameterizations are possible as mentioned by Stein (1999), Shkarofsky (1968). Special cases arise when the smoothness factor is taken to:

- 1/2 :exponential covariance function
- 1: AR(1) process: autoregressive process of first order also called Markov process of first order. This process is often used in the field of geodesy to analyze gravitational fields (Meier 1981; Grafarend 1976) as well as to fit GPS covariance functions (Jansson and Persson 2013; Wang et al. 2002).
- infinity: squared exponential covariance function. Due to its indefinitely differentiability which is difficult to explain physically, this covariance function (Stein 1999; Handcock and Wallis 1994) should be avoided. Some examples of problems that can arise using this smoothness are shown in Stein (1999).

Because of its flexibility as well as the possibility to estimate the parameters via maximum likelihood (Stein 1999; Handcock and Wallis 1994 for meteorological data field), we adopted this covariance function for all our computation of covariance matrices used in GPS least squares procedure.

Kermarrec and Schön (2014) proposed to compute the elements of the GPS covariance matrices thanks to a close formula: $\sigma_{it}^{i(t+\tau)} = \frac{\delta}{\sin(El_i(t))\sin(El_i(t+\tau))}(\alpha\tau)^1 K_{\nu}(\alpha\tau)$, where $El_i(t)$ is the elevation of satellite i at t and $El_i(t+\tau)$ at $t+\tau$. Other weightings than elevation-dependent model could be chosen. The ranges $\alpha \in [0.005 - 0.025]$, $\nu \in \left[1/4^{-1}\right]$ were proposed for GPS time series.

References

Abramowitz M, Segun IA (1972) Handbook of mathematical functions. Dover, New York

Ammar GS, Gragg WB (1988) Superfast solution of real positive definite Toeplitz systems. SIAM J Matrix Anal Appl 9:61–76

Ataike H (1973) Block Toeplitz matrix inversion. SIAM J Appl Math 24(2):234–241

Alkhatib H, Schuh WD (2007) Integration of the Monte Carlo covariance estimation strategy into tailored solution procedures for large-scale least squares problems. JoG 81(1):53–66

Baksalary JK (1988) Criteria for the equality between ordinary least squares and best linear unbiaised estimators under certain linear models. Can J Stat 16(1):97–102

Barciss EH (1969) Numerical solution of linear equations of Toeplitz and vector Toeplitz systems. Numerische Mathematik 13:404–424

Beutler G, Bauersima I, Gurtner W, Rothacher M (1986) Correlations between simultaneous GPS double difference observations in the multistation mode: implementation considerations and first experiences. Manuscripta Geodaetica 12:40–44. Springer-Verlag, Berlin, Heidelberg, New York

Bierman GJ (1977) Factorization methods for discrete sequential estimation. Volume 123 of mathematics in science and engineering. Academic Press, New York

Bona P (2000) Precision, cross correlation, and time correlation of GPS phase and code observations. GPS Solut 4(2):3–13

Borre K, Tiberius C (2000) Time series analysis of GPS observables. In: Proceedings of ION GPS 2000, Salt Lake City, UT, USA, September 19–22, pp 1885–1894

Bottoni GP, Barzaghi R (1993) Fast collocation. Bull Géodésique 67(2):119–126

Brent R (1989) Old and new algorithms for Toeplitz systems. In: Luk FT (ed) Proceedings SPIE, volume 975, advanced algorithms and architectures for signal processing III. SPIE, Bellingham, pp 2–9

Brent RP, Gustavson FG, Yun DYY (1980) Fast solution of Toeplitz systems of equations and computation of Pade approximants. J Algorithms 1(259):295

Bruyninx C, Habrich H, Söhne W, Kenyeres A, Stangl G, Völksen C (2012) Enhancement of the EUREF Permanent Network services and products. In: Kenyon S, Pacino MC, Marti U (eds) Geodesy for Planet Earth. International association of geodesy symposia, vol 136. Springer, Berlin, Heidelberg, pp 27–35

Cressie N (1993) Statistics for spatial data. J. Wiley. &. Sons, Inc., New York

Dach R, Brockmann E, Schaer S, Beutler G, Meindl M, Prange L, Bock H, Jäggi A, Ostini L (2009) GNSS processing at CODE: status report. J Geod 83(3–4):353–365

Dach R, Hugentobler U, Fridez P, Meindl M (2007) Bernese GPS software version 5.0. Astronomical Institute, University of Bern, Switzerland

Durbin J (1960) The fitting of time series models. Rev Inst Int Stat 28:233–244

El-Rabbany A (1994) The effect of physical correlations on the ambiguity resolution and accuracy estimation in GPS differential positioning. PhD thesis, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada

Euler HJ, Goad CC (1991) On optimal filtering of GPS dual frequency observations without using orbit information. Bull Geod 65(2):130–143

Grafarend EW (1976) Geodetic applications of stochastic processes. Phys Earth Planet Inter 12(2–3):151–179

Grafarend EW, Awange J (2012) Applications of linear and nonlinear models. Springer, Berlin

Gray RM (2006) Toeplitz and circulant matrices: a review. Found Trends Commun Inform Theory 2(3):155–239

Handcock MS, Wallis JR (1994) An approach to statistical spatial–temporal modeling of meteorological fields. J Am Stat Assoc 89(426):368–378

Hoffmann-Wellenhof B, Lichtenegger H, Collins J (2001) GPS theory and practice, 5th edn. Springer Wien, New York

Howind J, Kutterer H, Heck B (1999) Impact of temporal correlations on GPS-derived relative point positions. J Geod 73(5):246–258

Jansson P, Persson CG (2013) The effect of correlation on uncertainty estimates—with GPS examples. J Geod Sci 3(2):111–120

Jin SG, Luo O, Ren C (2010) Effects of physical correlations on long-distance GPS positioning and zenith tropospheric delay estimates. Adv Space Res 46:190–195

Journel AG, Huifbregts CJ (1978) Mining geostatistics. Academic Press, New York

Kermarrec G, Schön S (2014) On the Mátern covariance family: a proposal for modeling temporal correlations based on turbulence theory. J Geod 88:1061–1079

Klees R, Broersen P (2002) How to handle colored observation noise in large-scale least-squares problems-building the optimal filter. DUP Science, Delft University Press, Delft (30 pages)

Koch KR (1999) Parameter estimation and hypothesis testing in linear models. Springer, Berlin

Koivunen AC, Kostinski AB (1999) The feasibility of data whitening to improve performance of weather radar. J Appl Meteor 38:741–749

Krämer W (1986) Least squares regression when the independent variable follows an ARIMA process. J Am Stat Assoc 81(393):150–154. doi:10.2307/2287982

Krämer W, Donninger C (1987) Spatial autocorrelation among errors and the relative efficiency of OLS in the linear regression model. J Am Stat Assoc 82(398):577–579

Kutterer H (1999) On the sensitivity of the results of least-squares adjustments concerning the stochastic model. J Geod 73:350–361

Leandro R, Santos M, Cove K (2005) An empirical approach for the estimation of GPS covariance matrix of observations. In: Proceeding ION 61st Annual Meeting, The MITRE Corporation and Draper Laboratory, 27–29 June 2005, Cambridge, MA

Levinson N (1947) The Wiener RMS error criterion in filter design and prediction. J Math Phys 25(1–4):261–278

Luo X, Mayer M, Heck B (2012) Analysing time series of GNSS residuals by means of ARIMA processes. Int Assoc Geod Symp 137:129–134

Luo X (2012) Extending the GPS stochastic model by means of signal quality measures and ARMA processes. PhD Karlsruhe Institute of Technology

Luati A, Proietti T (2011) On the equivalence of the weighted least squares and the generalised least squares estimators, with applications to kernel smoothing. Ann Inst Stat Math 63(4):673–686

Mátern B (1960) Spatial variation-Stochastic models and their application to some problems in forest surveys and other sampling investigation. Medd Statens Skogsforskningsinstitut 49(5):144

Meier S (1981) Planar geodetic covariance functions. Rev Geophys Space Phys 19(4):673–686

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

Niemeier W (2008) Adjustment computations, 2nd edn. Walter de Gruyter, New York

Puntanen S (1987) On the relative goodness of ordinary least squares estimation in the general linear model. Acta Univ Tamper Ser A 216

Puntanen S, Styan G (1989) The equality of the ordinary least squares estimator and the best linear unbiased estimator. Am Stat 43(3):153–161. doi:10.2307/2685062

Radovanovic RS (2001) Variance-covariance modeling of carrier phase errors for rigorous adjustment of local area networks. IAG 2001 Scientific Assembly, Budapest, Hungary, September 2–7, 2001

Rao C, Toutenburg H (1999) Linear models, least-squares and alternatives, 2nd edn. Springer, New York

Santos MC, Vanicek P, Langley RB (1997) Effect of mathematical correlation on GPS network computation. J Surv Eng 123(3):101–112

Satirapod C, Wang J, Rizos C (2003) Comparing different GPS data processing techniques for modelling residual systematic errors. J Surv Eng 129(4):129–135

Schön S, Kutterer H (2006) Uncertainty in GPS networks due to remaining systematic errors: the interval approach. J Geod 80:150–162

Schön S, Brunner FK (2007) Treatment of refractivity fluctuations by fully populated variance-covariance matrices. In: Proc. 1st Colloquium Scientific and Fundamental Aspects of the Galileo Programme Toulouse Okt

Schön S, Brunner FK (2008) Atmospheric turbulence theory applied to GPS carrier-phase data. J Geod 1:47–57

Schuh WD, Krasbutter I, Kargoll B (2014) Korrelierte Messung—was nun? Neuner H (Hrsg.): Zeitabhängige Messgrößen—Ihre Daten haben (Mehr-)Wert, DVW-Schriftenreihe 74. ISBN: 978-3-89639-970-0, S. 85–101

Shkarofsky IP (1968) Generalized turbulence space-correlation and wave-number spectrum-function pairs. Can J Phys 46:2133–2153

Stein ML (1999) Interpolation of spatial data. Some theory for kriging. Springer, New York

Teunissen PJG, Kleusberg A (1998) GPS Geod. Springer, Berlin

Trench WF (1964) An algorithm for the inversion of finite Toeplitz matrices. J Soc Indus Appl Math 12(3):515–522

Trefethen LN, Bau D (1997) Numerical linear algebra. Society for Industrial and Applied Mathematics, Philadelphia. ISBN 0898713617, 9780898713619

Vennebusch M, Schön S, Weinbach U (2010) Temporal and spatial stochastic behavior of high-frequency slant tropospheric delays from simulations and real GPS data. Adv Space Res 47(10):1681–1690

Wang J, Satirapod C, Rizos C (2002) Stochastic assessment of GPS carrier phase measurements for precise static relative positioning. J Geod 76(2):95–104

Weinbach T (2012) (2012): Feasibility and impact of receiver clock modeling in precise GPS data analysis. Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover ISSN 0174–1454, Nr. 303, Hannover 2013

Wieser A, Brunner FK (2000) An extended weight model for GPS phase observations. Earth Planet Space 52:777–782

Zyskind G (1967) On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. Ann Math Stat 38(4):1092–1109

J Geod (2017) 91:465–484
DOI 10.1007/s00190-016-0976-8

ORIGINAL ARTICLE

A priori fully populated covariance matrices in least-squares adjustment—case study: GPS relative positioning

Gaël Kermarrec[1] · Steffen Schön[1]

Received: 27 May 2016 / Accepted: 12 November 2016 / Published online: 9 December 2016

Abstract In this contribution, using the example of the Mátern covariance matrices, we study systematically the effect of apriori fully populated variance covariance matrices (VCM) in the Gauss–Markov model, by varying both the smoothness and the correlation length of the covariance function. Based on simulations where we consider a GPS relative positioning scenario with double differences, the true VCM is exactly known. Thus, an accurate study of parameters deviations with respect to the correlation structure is possible. By means of the mean-square error difference of the estimates obtained with the correct and the assumed VCM, the loss of efficiency when the correlation structure is missspecified is considered. The bias of the variance of unit weight is moreover analysed. By acting independently on the correlation length, the smoothness, the batch length, the noise level, or the design matrix, simulations allow to draw conclusions on the influence of these different factors on the least-squares results. Thanks to an adapted version of the Kermarrec–Schön model, fully populated VCM for GPS phase observations are computed where different correlation factors are resumed in a global covariance model with an elevation dependent weighting. Based on the data of the EPN network, two studies for different baseline lengths validate the conclusions of the simulations on the influence of the Mátern covariance parameters. A precise insight into the impact of apriori correlation structures when the VCM is entirely unknown highlights that both the correlation length and the smoothness defined in the Mátern model are important to get a lower loss of efficiency as well as a better estimation of the variance of unit weight. Consecutively, correlations, if present, should not be neglected for accurate test statistics. Therefore, a proposal is made to determine a mean value of the correlation structure based on a rough estimation of the Mátern parameters via maximum likelihood estimation for some chosen time series of observations. Variations around these mean values show to have little impact on the least-squares results. At the estimates level, the effect of varying the parameters of the fully populated VCM around these approximated values was confirmed to be nearly negligible (i.e. a mm level for strong correlations and a submm level otherwise).

Keywords Mátern covariance family · Correlations, smoothness · GPS · Double difference · Design matrix · Loss of efficiency · Mean-square errors · Variance of unit weight

✉ Gaël Kermarrec
gael.kermarrec@web.de

Steffen Schön
schoen@ife.uni-hannover.de

[1] Institut für Erdmessung (IfE), Leibniz Universität Hannover, Schneiderberg 50, 30167 Hannover, Germany

1 Introduction

The functional part of the least-squares adjustment is traditionally considered as deterministic and can be mathematically formulated. On the contrary, the stochastic model is usually unknown and simplified to diagonal matrices. Correlations between observations are therefore neglected due to the lack of knowledge of an approximate covariance model for the observations or the necessity to invert fully populated covariance matrices (Howind et al. 1999).

The effect of incorrect weights has been investigated in the statistical literature (we cite exemplary Rao and Toutenburg 1999; Watson 1967), particularly when the apriori covariance matrix is equal to the identity matrix (Puntanen and Styan 1986). The impact of incorrect covariance matrices on the least-squares estimates, cofactor matrices, mean-squared errors (MSE), outlier tests and the corresponding redundancy numbers or the variance of unit weight have been mathematically expressed by (Wolf 1961; Rao and Toutenburg 1999; Koch 1999; Kutterer 1999; Xu 1991, 2013; Hahn and Van Mierlo 1987). Kutterer (1999) proposed to use intervals and linearization to bound the difference of parameters estimated with perturbated covariance matrices. Xu (2013) investigated the effect of incorrect covariance matrices on the variance of unit weight, showing the dependency of the resulting bias. For the particular case of least-squares collocation, Xu (1991) analysed how incorrect prior weight matrices of the signals influence the mean square errors. Hahn and Van Mierlo (1987) studied the sensitivity of adjustment results when the precision of one observation is changed with special attention to the corresponding redundancy number or reliability. The derived formulas, particularly for the estimates, depend on the residuals—thus on the underlying data—making general conclusions on the impact of taking correlations into account in least-squares adjustment not straightforward.

Restricting ourselves in this study to the influence of approximate covariance matrices for the global positioning system (GPS) positioning case, El-Rabbany (1994), Wang et al. (1998), Bona (2000) and Tiberius and Kenselaar (2003) can be cited exemplarily for an overview of the previous work on temporal or physical correlations of the GPS observations. Luo (2012) or Luo et al. (2012) made use of the ARIMA or ARMA models to describe the correlation structure of the studentized and filtered residuals of positioning adjustment. Interesting dependencies with multipath or atmospheric conditions were shown. Different procedures based on the analysis of the residuals were also developed to compute the covariance matrices of GPS observations (Hannan 1970; Wang et al. 2002, 2005; Satirapod et al. 2003 for real-time kinematic positioning). Least-squares variance component estimation was used by Teunissen and Amiri-Simkooei (2008) and Amiri-Simkooei et al. (2009). However, this procedure is to some extend computational demanding and in the GPS positioning strategies often replaced by simpler model using for instance exponential functions (El-Rabbany 1994; Howind et al. 1999). Another possibility is to propose an apriori model based on physical considerations (Radovanovic 2001 for multipath; Schön and Brunner 2008; Kermarrec and Schön 2014 for tropospheric refraction).

The study of GPS observations heterosceadasticity (Bischoff et al. 2005) showed that the error variances of GPS observations are inhomogeneous. The variance component estimation (VCE) has been used to improve the stochastic model of GPS observations (Li et al. 2011; Tiberius and Kenselaar 2003; Amiri-Simkooei 2007; Amiri-Simkooei et al. 2016). Elevation dependent models (Vermeer 1997; Euler and Goad 1991; Dach et al. 2007), C/N0 (Wieser and Brunner 2000) or SNR-based models (Luo et al. 2013) have also been proposed.

Through all these studies, the goal of improving the stochastic model, by changing either the diagonal weighting or the correlation structure, is mainly to have a more accurate parameter estimation (i.e. lower root mean-square of the coordinates), more reliable test statistics and eventually better ambiguity resolution. However, a more concrete analysis that explains the mathematically derived formulas when changing the structure of the approximated covariance functions is missing for a GPS positioning case. Based on simulations where the underlying covariance model is exactly known, more general conclusions on the influence of varying the correlation structure of fully populated VCM on the least-squares results can be drawn to propose a simplified but accurate stochastic model for GPS observations.

In this contribution, we will systematically study the impact of changing the apriori covariance matrices on the mean-square error (MSE) of the estimates difference using the developments of Strand (1974). Moreover, the formula of Xu (2013) for the bias of the variance of unit weight will be more concretely investigated to quantify the impact of correlations on this parameter used for test statistics purpose (Teunissen 2000). We will restrict ourselves to the comparison of matrices with the same diagonal weighting. Following cases are particularly investigated: errors in the model (i.e. wrong correlation structure), simplifications by using only block diagonal structure or diagonal VCM (elevation dependent or identity). The Mátern model for describing correlations (Mátern 1960; Whittle 1954; Stein 1999) acts both on the smoothness and the correlation length of the time series. It was previously successfully applied by Kermarrec and Schön (2014) for modelling tropospheric refractivities and will consecutively be extended to describe in a general way temporal correlations and non-stationarity of GPS phase observations.

The choice of a Mátern function is suggested by analogy with the work of Luo (2012) on ARIMA models which were shown to fit the residuals of positioning adjustment, i.e. the differential equations of the corresponding underlying processes have some similarities (Rasmussen and Williams 2006). It should be pointed out that a precise description of the correlation structure for GPS data based on residuals analysis for particular cases is not our goal. Many different factors can create correlations (between observation types, or temporal for instance) and estimated from the observations (Amiri-Simkooei et al. 2013). In this contribution, a more general description of the correlations is here deliberately chosen in order to simplify and democratize the use of VCM in least-squares adjustment by applying a global covariance function. We follow there to some extend the work on kriging of Stein (1999). Thus, the corresponding covariance parameters will be varied in a physically plausible range.

The remainder of this paper is structured as follows: in a first part, we will define the mathematical notation and gives the general formulation of the MSE of the estimates difference as well as the bias of the variance of unit weight. In a second part, these formulas will be used for simulations where the true covariance matrix is known but approximated in different ways in the least-squares adjustment. A design matrix corresponding to GPS relative positioning with double differences will be used. Results for different batch sizes, different correlation structures and design matrices are presented. A parallel with the results of Hannan (1970) for the BLUP will be drawn in "Appendix 2" to explain some of the dependencies found in the simulations. In a third step, a real GPS data analysis for a short and long baseline positioning scenario will be presented validating the results of the simulations. Some propositions on how to choose the covariance matrices will conclude this contribution.

2 Mathematical background

In this section, the general notations as well as the main concept of linear weighted least-squares adjustment and of estimator comparison by means of Mean-Squared Error are summarized. Interested readers should refer to Rao and Toutenburg (1999), Koch (1999) and Grafarend and Awange (2012) for more details. As an elegant way to describe correlations, the Mátern covariance function is shortly presented.

2.1 Ordinary, general and feasible least square estimator

We assume a linear or linearized functional model

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \varepsilon, \tag{1}$$

$\mathbf{y}$ is the $n \times 1$ observation vector, $\mathbf{A}$ the non-stochastic $n \times u$ design matrix with full column rank ($\text{rk}(\mathbf{A}) = u$), $\mathbf{x}$ the $u \times 1$ parameter vector to be estimated, $\boldsymbol{\varepsilon}$ the $n \times 1$ observation error vector. The error term has zero mean and a normal distribution with $E\left(\boldsymbol{\varepsilon}\,\boldsymbol{\varepsilon}^{\mathbf{T}}\right) = \sigma^2\mathbf{W}_0$, where $\mathbf{W}_0$ is a $n \times n$ positive definite fully populated cofactor matrix of the observations, σ^2 the apriori variance factor, and $E\,(.)$ denotes the mathematical expectation.

The generalized least squares estimator (GLSE) reads:

$$\hat{\mathbf{x}}_0 = \left(\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{-1}\mathbf{y}, \tag{2}$$

and the cofactor matrix $\mathbf{W}_{0\hat{\mathbf{x}}\hat{\mathbf{x}}}$ of the unknows (apriori estimator) is given by:

$$\mathbf{W}_{0\hat{\mathbf{x}}\hat{\mathbf{x}}} = \left(\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{-1}\mathbf{A}\right)^{-1}, \tag{3}$$

Calling $\mathbf{v}_0 = \mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_0$ the $n \times 1$ residual vector, the aposteriori variance factor of the observations, also called the estimate of the variance of unit weight (Xu 2013; Koch 1999) is, therefore, given by:

$$\hat{\sigma}^2_{\mathbf{W_0}} = \frac{\left(\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_0\right)^{\mathbf{T}}\mathbf{W}_0^{-1}\left(\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_0\right)}{n-u} = \frac{\mathbf{v}_0^{\mathbf{T}}\mathbf{W}_0^{-1}\mathbf{v}_0}{n-u}. \tag{4}$$

This estimator is unbiased when the correct weight matrix $\mathbf{W}_0^{-1}$ is used, i.e. $E\left(\hat{\sigma}^2_{\mathbf{W_0}}\right) = \sigma^2$, where σ^2 is the apriori variance factor of the observations, Koch (1999) and Rao and Toutenburg (1999).

However, the GLSE is not a feasible estimator since the elements of the cofactor matrix $\mathbf{W}_0$ are unknown. Thus, the so-called Feasible Generalized Least Squares Estimator (FGLSE) is used in practise, where the true $\mathbf{W}_0$ is replaced by its estimate or assumption $\hat{\mathbf{W}}$ (Greene 2003). This terminology seems to be mostly used in econometrics, however, it allows a more accurate distinction of the different least-squares results when the covariance is unknown and estimated or totally known and will be used for the mean-squared error study.

There are $\frac{1}{2}n\,(n+1)$ unknowns and only $\frac{1}{2}\,(n-u)\,(n-u+1)$ observations in the stochastic model. As $\frac{1}{2}n\,(n+1) > \frac{1}{2}\,(n-u)\,(n-u+1)$ not all components can be estimated (Amiri-Simkooei 2007). A way to treat this problem is for instance to use an approximated model for the (co)variance of the observations by fitting exponential or polynomial functions (see exemplary Koch et al. 2010), as well as ARMA or ARIMA models (Luo et al. 2012) to the residuals; the covariance structure can be also found iteratively from the residuals (Wang et al. 2002).

The FGLSE can be transformed into an homosceaticitic model (identical variances for all observations and no correlation) by computing a $n \times n$ transformation matrix $\mathbf{T}$ such as the covariance matrix of the observations is $\sigma^2\mathbf{I}$ where $\mathbf{I}$ is the identity matrix. In practice, a Cholesky factorization of the covariance matrix $\hat{\mathbf{W}}$ is used (Koch 1999).

2.2 Misspecification of the covariance matrix

For different apriori covariance matrices, the least-squares results as defined in Sect. 2.1. will correspondingly change. In this section, we will shortly present the awaited bias by missspecifying the stochastic model in a least-squares model.

Let $\hat{\mathbf{W}} = \mathbf{W}_0 + \Delta\mathbf{W}$ represents the apriori or assumed cofactor matrix, where $\Delta\mathbf{W}$ is the difference between the assumed and the true cofactor matrix $\mathbf{W}_0$. The corresponding estimators are $\hat{\mathbf{x}}$, $\hat{\mathbf{x}}_0$, respectively.

A posteriori variance factor

Misspecification of the covariance matrix leads to errors in the a posteriori variance estimator (Watson 1955; Rao and Toutenburg 1999). It can be shown that:

$$E\left(\hat{\sigma}^2_{\hat{\mathbf{W}}}\right)(n-u) = \sigma^2 \mathrm{tr}\left(\mathbf{W_0}-\mathbf{A}\left(\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A^T}\hat{\mathbf{W}}^{-1}\right) + ...$$
$$...tr\begin{pmatrix}\sigma^2\mathbf{A}\left(\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A^T}\hat{\mathbf{W}}^{-1}(\mathbf{I}-2\mathbf{W_0})+... \\ \sigma^2\mathbf{A}\left(\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{W_0}\hat{\mathbf{W}}^{-1}\mathbf{A}\left(\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A^T}\end{pmatrix}$$

where tr denotes the trace of the matrix and $\mathbf{I}$ the identity matrix.

For the regression case and assuming that the regression vector forms an orthonormal set (i.e. $\mathbf{A^TA} = \mathbf{I}$), Watson (1955) obtained bounds of the variance factor showing a strong dependence with the roots of the design matrix. Following Xu (2013), the bias of the variance of unit weight due to the incorrect weights on the estimate can be further expressed as:

$$E\left(\hat{\sigma}^2_{\hat{\mathbf{W}}}\right) = \sigma^2 + \mathrm{tr}\left\{\left(\mathbf{I}-\hat{\mathbf{W}}^{-1}\mathbf{A}\left(\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A^T}\right)\hat{\mathbf{W}}^{-1}\Delta\mathbf{W}\right\}$$
$$\frac{\sigma^2}{n-u} = \sigma^2(1+TR) \tag{5}$$

Thus, when using an approximated cofactor matrix, a bias *TR* depending on $\Delta\mathbf{W}$ occurs, leading to either smaller or larger variances (Dufour 1989).

Influence on the estimated parameters

Kutterer (1999) derived the difference in the estimated parameter vector due to $\Delta\mathbf{W}$:

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}_0 - \left(\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A^T}\Delta\mathbf{P}\mathbf{v_0} = \hat{\mathbf{x}}_0 + \Delta\mathbf{x} \tag{6}$$

where $\Delta\mathbf{P} = \mathbf{P_0}-\hat{\mathbf{P}}$, $\hat{\mathbf{P}} = \hat{\mathbf{W}}^{-1}$, $\mathbf{P_0} = \mathbf{W_0^{-1}}$. This difference can be further linearized and expressed as sum of a first order term and a corresponding so called "error term" for small $\Delta\mathbf{P}$. Consecutively, bounds of the norms can be obtained. However, such bounds are rough, except for small variations around the original covariance matrices and imply an already accurate knowledge of the VCM. Such conditions are rarely reached, particularly when comparing OLS with FGLS estimator. Moreover, the differences $\hat{\mathbf{x}} - \hat{\mathbf{x}}_0$ depends on the residuals, and thus on the data sets used making this formula not adequate for a more general quantification on the estimates of the effect of changing the covariance matrices.

Mean-squared error MSE

Based on Xu (1991), Rao and Toutenburg (1999) or Strand (1974), we will here use the mean-squared error (MSE) to compare the estimators obtained with approximated VCM. As the mean-squared error is related to traces of matrices, for the next sections, the natural norm used will be the Frobenius norm defined for any square matrix as $\|\Delta\mathbf{W}\|_F = \mathrm{tr}\left(\Delta\mathbf{W}^T\Delta\mathbf{W}\right)$.

MSE is defined as the function $E\left[\left(\hat{\mathbf{x}}-\hat{\mathbf{x}}_0\right)^T\left(\hat{\mathbf{x}}-\hat{\mathbf{x}}_0\right)\right]$ or the average squared difference between the estimator $\hat{\mathbf{x}}$ and the parameter $\hat{\mathbf{x}}_0$. It can be interpreted as the sum of the variability of the estimator (or precision) and its bias (or accuracy). For the case of unbiased estimator such as the general least squares estimator, the MSE and its variance have the same values. The mean-squared error of an estimator (Wackerly et al. 2008) measures its performance: a MSE close to 0 means that the estimator predicts the parameter with perfect accuracy. MSE should only be used for comparative purposes, otherwise they are meaningless.

The loss of efficiency in estimating $\hat{\mathbf{x}}$ instead of $\hat{\mathbf{x}}_0$ is given by the nonnegative definite matrix (Rao and Toutenburg 1999, p 107; Strand 1974):

$$\mathbf{W}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}-\mathbf{W}_{\mathbf{0}\hat{\mathbf{x}}\hat{\mathbf{x}}} = \left(\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{W_0}\hat{\mathbf{W}}^{-1}\mathbf{A}$$
$$\left(\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1} - \left(\mathbf{A^T}\mathbf{W}_0^{-1}\mathbf{A}\right)^{-1}. \tag{7}$$

Equation (7) represents the difference of the cofactor matrices of the unknowns between the FGLS and the GLS estimator.

If we call $E(\mathbf{x_0}) = \mathbf{x_{ref}}$, using the unbiasness of the least squares estimator and by taking trace, Eq. (7) can be written as:

$$E\left[\left(\hat{\mathbf{x}}-\mathbf{x_{ref}}\right)^{\mathbf{T}}\left(\hat{\mathbf{x}}-\mathbf{x_{ref}}\right)\right] = E\left[\left(\hat{\mathbf{x}}-\mathbf{x_0}\right)^{\mathbf{T}}\left(\hat{\mathbf{x}}-\mathbf{x_0}\right)\right]$$
$$-E\left[\left(\mathbf{x_{ref}}-\mathbf{x_0}\right)^{\mathbf{T}}\left(\mathbf{x_{ref}}-\mathbf{x_0}\right)\right],$$

or alternatively $\mathrm{MSE}_{\hat{\mathbf{x}}-\mathbf{x}_{\mathrm{ref}}} = \mathrm{MSE}_{\hat{\mathbf{x}}-\mathbf{x}_0} - \mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{ref}}-\mathbf{x}_0}$.

The trace operator being commutative, we obtain

$$\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}} = \mathrm{tr}\left(\left[\hat{\mathbf{W}}^{-1}\mathbf{A}\left(\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-2}\mathbf{A^T}\hat{\mathbf{W}}^{-1}\right]\mathbf{W}_0\right)$$
$$-\mathrm{tr}\left(\left(\mathbf{A^T}\mathbf{W}_0^{-1}\mathbf{A}\right)^{-1}\right) \tag{8}$$

This expression can be simply written as

$$\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}} = \mathrm{tr}\left(\mathbf{P_1}\mathbf{W}_0\right)$$
$$-\mathrm{tr}\left(\left(\mathbf{A^T}\hat{\mathbf{W}}_0^{-1}\mathbf{A}\right)^{-1}\right), \tag{9}$$

where $\mathbf{P_1} = \left[\hat{\mathbf{W}}^{-1}\mathbf{A}\left(\mathbf{A}^T\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-2}\mathbf{A}^T\hat{\mathbf{W}}^{-1}\right]$ is a known matrix of size $(n \times n)$. Equation (8) highlights that the $\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}}$ is the difference of two unknown quantities and cannot be used easily without an exact knowledge of $\mathbf{W}_0$. However, using simple trace equality, Strand (1974) proved that Eq. (9) can be approximated. Moreover, based on the Rayleigh quotient, Strand (1974) developed a bound of this equation with known quantities (design matrix, $\|\Delta\mathbf{W}\|_F$ as well as the estimated covariance matrix) valid for small $\|\Delta\mathbf{W}\|_F$ which won't be developed here.

The main advantage of the MSE formulation is its non-dependency on the dataset allowing still a quantification of the mean-squared differences of the estimated parameters when covariance matrices are changed (i.e. for instance diagonal with fully populated VCM). For a more straightforward comparison of the estimators, ratios should be, however, formed (see also Stein 1999).

MSE differences and aposteriori variance factor ratios

In the case of simulation studies, $\mathbf{W}_0$ is known, thus the ratio

$$R_{\mathrm{MSE}} = \frac{\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}}}{\mathrm{MSE}_{\mathbf{x}_{\mathrm{ref}}-\hat{\mathbf{x}}_0}} = \frac{\mathrm{tr}\left(\left[\hat{\mathbf{W}}^{-1}\mathbf{A}\left(\mathbf{A}^T\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-2}\mathbf{A}^T\hat{\mathbf{W}}^{-1}\right]\mathbf{W}_0\right)}{\mathrm{tr}\left(\left(\mathbf{A}^T\mathbf{W}_0^{-1}\mathbf{A}\right)^{-1}\right)} - 1 \tag{10}$$

allows to analyse the influence of changes in the structure of the estimated covariance matrices; a minimum of the ratio being searched, deviations from it give some indications on the global qualities of the corresponding estimator.

The bias of the variance of unit weight can be further computed; an unbiaised $\hat{\sigma}^2_{\hat{\mathbf{W}}}$ corresponding to TR close to 0 being preferred. Correspondingly to Eq. (10), the ratio $R_{\hat{\sigma}} = \frac{E\left(\hat{\sigma}^2_{\hat{\mathbf{W}}}\right)}{\sigma^2} = 1 + \mathrm{TR}$ is formed. For test statistics we use the overall model test (Teunissen 2000) which reads:

$$\frac{\hat{\sigma}^2_{\hat{\mathbf{W}}}}{\sigma^2} > F_p\left(m-n, \infty, 0\right) \tag{11}$$

where $F_p\left(m-n, \infty, 0\right)$ is the central F-distribution having $n-u$ and ∞ degrees of freedom, and a probability p.

In case of real data study, as for instance by computing coordinates with least-squares adjustment from a GPS relative positioning scenario with double differences, $\mathbf{W}_0$ is unknown. Thus, as soon as neither $\Delta\mathbf{W}$ nor $\|\Delta\mathbf{W}\|_F$ are accessible, MSE values are not computable to quantify the loss of efficiency.

However, it is possible to compare two estimators by computing the root Mean-Squared of the parameters difference:

$$3\mathrm{drms}_{\mathbf{X_1}-\mathbf{X_2}} = \sqrt{\frac{\mathrm{trace}\left((\mathbf{X_1}-\mathbf{X_2})^{\mathbf{T}}(\mathbf{X_1}-\mathbf{X_2})\right)}{m}}, \tag{12}$$

where $\mathbf{X_1}, \mathbf{X_2}$ are the time series of coordinate of size $m \times 3$ computed with two different apriori cofactor matrices, $\hat{\mathbf{W}}_1$, $\hat{\mathbf{W}}_2$, respectively, but the same design matrix.

2.3 The Mátern covariance functions

In the following section, the Mátern covariance function used in this contribution to compute $\hat{\mathbf{W}}$ for GPS observations is shortly presented. Derived from an inverse polynomial spectral density function, Mátern covariance functions can model many physical processes (Mátern 1960; Guttorp and Gneiting 2005; Meier 1981) and can be extended to anisotropy and nonstationarity (Fuentes 2002; Spöck and Pilz 2008).

The isotropic and stationary covariance function can be parametrized as follows:

$$C\left(r\right) = \phi\left(\alpha r\right)^{\nu} K_{\nu}\left(\alpha r\right), \quad r = \|\mathbf{x}-\mathbf{y}\| \tag{13}$$

r is the Euclidean distance where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are two points (Yaglom 1987). The corresponding spectral density, related to the covariance function via the Wiener-Khintchine theorem (Chatfield 1989) is given by:

$$S\left(\omega\right) = \frac{\phi 2^{\nu-1}\Gamma\left(\nu+1/2\right)\alpha^{2\nu}}{\sqrt{\pi}\left(\alpha^2+\omega^2\right)^{\nu+1/2}} \tag{14}$$

where $\nu > 0$, $\alpha > 0$ are constant parameters, ω the angular frequency. The scalar parameter $\phi > 0$ is chosen so that the variance equals 1; ν is a measure of the Mean-Squared differentiability of the field (Stein 1999), defined as "its smoothness". The constant α is the inverse of the Mátern correlation time or length (ICT) and indicates how the correlations decay with increasing distance or time (Journel and Huifbregts 1978). A correspondence between correlation length and α is shortly given in Table 1. The modified Bessel function of order ν (Abramowitz and Segun 1972) is denoted by K_{ν}. When ν is half-integer, (i.e. $\nu = l + \frac{1}{2}$, where l is a non-negative integer), the Mátern covariance functions are easy to express as a product of an exponential and a polynomial of order l. For instance:

Table 1 Approximate correspondence between correlation length or time for different Mátern parameter sets

	$\alpha = 0.01$	$\alpha = 0.03$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.2$
Correlation time (s) (approximately)	600	300	180	120	60

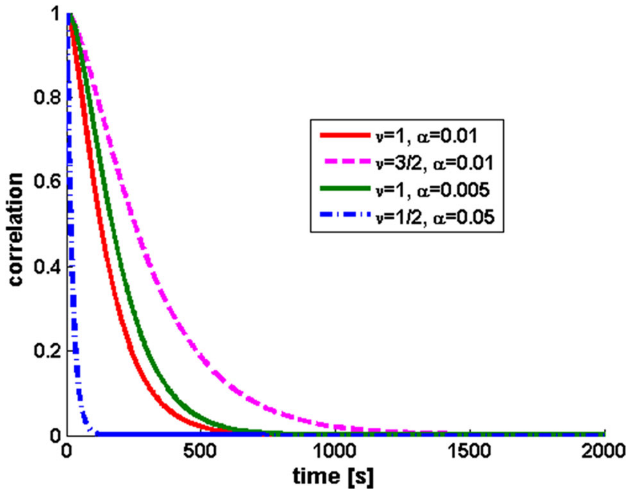


Fig. 1 Correlation functions for different Mátern parameter sets

- $\nu = 1/2$ corresponds to the exponential covariance function,
- $\nu = 1$ to the Markov process of first order,
- $\nu = \infty$ is the squared exponential covariance function. Although popular, this function corresponds to an infinitely differentiable random field at the origin and is thus said to be physically less plausible than other parametrization. Some examples of the problems that can arise by using this smoothness in least-squares or predictions are shown in Stein (1999) or Koivunen and Kostinski (1999).

Figure 1 shows the different correlation functions obtained by varying the Mátern parameters set $[\alpha, \nu]$. As can be seen, the differences between the correlation functions for different smoothness parameters are not visually obvious. However, the smoothness of the time series controls the rate of decay of the spectral density at high frequencies and plays an important role in interpolation problems. It should be chosen with care to avoid missspecification of the covariance structure, particularly at high frequencies. Consequences could be for instance a poor decorrelation of the observations and correspondingly residuals that are not white noise when used in a least-squares adjustment. Further examples of time series simulated with the Mátern covariance model can be found in Kermarrec and Schön (2014).

The elements of the corresponding cofactor matrices are given by:

$$\hat{\mathbf{W}}(r_1, r_2) = \left(\alpha \left|r_1^- r_2\right|\right)^{\nu} K_{\nu}\left(\alpha \left|r_1^- r_2\right|\right), \quad (15)$$

where (r_1, r_2) are two indices. The true cofactor matrix $\mathbf{W}_0$ is a $(n \times n)$ fully populated known matrix defined by $\mathbf{W}_0(r_1, r_2) = (\alpha_0 |r_1 - r_2|)^{\nu_0} K_{\nu_0}(\alpha_0 |r_1 - r_2|)$.

In the next section, cofactor matrices built on Eq. (15) are used in a least-squares adjustments (FGLSE) where the design matrix corresponds to a relative GPS positioning scenario. The Mátern parameters $[\alpha, \nu]$ of the $\hat{\mathbf{W}}$ are varied to show their concrete influence on the ratios as defined in Sect. 2.1.

3 Simulations using GPS positioning scenario

3.1 Stochastic and functional models

3.1.1 Stochastic model

The influence of changing the estimated covariance matrices in the least-squares adjustment is studied for a relative positioning scenario with double differences where we first assume the covariance matrix of the observations to be known. A first step explains the construction of the different fully populated covariance matrices whereas a second step describes the results for the MSE, variance of unit weight and ratios when changing the estimated VCM. General knowledge on positioning strategies with double differences can be found in Hoffmann-Wellenhof et al. (2001).

Based on the work of Kermarrec and Schön (2014) which proposed a modelization of tropospheric correlations for the GPS signals traveling the atmosphere, a Mátern function is taken to build the covariance matrices.

The covariance function C between 2 observations at time t and $t + \tau$ for one satellite i reads:

$$C^{i,i}(t, t+\tau) = \frac{\delta}{\sin(\mathrm{El}_i(t)) \sin(\mathrm{El}_i(t+\tau))} (\alpha\tau)^{\nu} K_{\nu}(\alpha\tau) \quad (16)$$

where El_i is the elevation of the satellite i. δ is a positiv factor so that the value of the covariance function is scaled to have a variance of 1 for a satellite at 90° elevation. For modelling correlations due to the travel of the GPS signals through the troposphere, the original model fixed $[\alpha, \nu] = [0.008, 5/6]$ based on turbulence theory.

We further extend the model by introducing a weighting factor ρ to model the covariance between different satellites i.e.

$$C^{i,j}(t, t+\tau) = \frac{\rho\delta}{\sin(\mathrm{El}_i(t)) \sin\left(\mathrm{El}_j(t+\tau)\right)} (\alpha\tau)^{\nu} K_{\nu}(\alpha\tau), i \neq j, \quad (17)$$

where El_j is the elevation of the satellite j.

Building the covariance matrix

The corresponding covariance matrix for a given station $\mathbf{C}^{i,j}$ is built with Eq. (15). ρ is taken to 0.1 which corresponds to a value that fits the model proposed by Kermarrec and Schön (2014) (i.e. correlations between different satellites 5–10 times lower than the covariance of one satellite with itself).

The fully populated VCM $\hat{\mathbf{W}}$ has a block structure. For a station A and s satellites, we have:

$$\hat{\mathbf{W}}_A = \begin{bmatrix} \mathbf{C}_\mathbf{A}^{1,1} & \mathbf{C}_\mathbf{A}^{1,2} & \mathbf{C}_\mathbf{A}^{1,3} & \cdots & \mathbf{C}_\mathbf{A}^{1,s} \\ & \mathbf{C}_\mathbf{A}^{2,2} & \mathbf{C}_\mathbf{A}^{2,3} & & \mathbf{C}_\mathbf{A}^{2,s} \\ & \cdots & \mathbf{C}_\mathbf{A}^{3,3} & & \\ & & & & \mathbf{C}_\mathbf{A}^{s,s} \end{bmatrix}.$$

For two stations A and B, the global covariance matrix for undifferenced observations reads:

$$\hat{\mathbf{W}}_{\text{UD}} = \begin{bmatrix} \hat{\mathbf{W}}_A & \hat{\mathbf{W}}_{A,.B} \\ \hat{\mathbf{W}}_{A,.B} & \hat{\mathbf{W}}_B \end{bmatrix}.$$

For sake of simplicity and because the correlations between observations of 2 different stations are much smaller than between the observations at one station only (Schön and Brunner 2008), we assume $\hat{\mathbf{W}}_{A,B} = \mathbf{0}$. $\hat{\mathbf{W}}_{\text{UD}}$ is a cofactor matrix of the undifferenced observations that does not account for mathematical correlations (Santos et al. 1997). Using the variance covariance propagation law (Beutler et al. 1987), the cofactor matrix for a relative positioning scenario with double differences reads finally:

$\hat{\mathbf{W}} = \mathbf{W} = M^{\mathbf{T}}\hat{\mathbf{W}}_{\text{UD}}\mathbf{M}$, where $\mathbf{M}$ is the matrix operator of double differencing.

In the same way, the true cofactor matrix of the observations is defined as $\mathbf{W}_0 = \mathbf{M}^{\mathbf{T}}\mathbf{W}_{0\text{UD}}\mathbf{M}$ where $\mathbf{W}_{0\text{UD}}$ is the true cofactor matrix of the undifferenced observations.

Choice of the model

Equations (16) and (17) allow to build sub-covariance matrices for all different satellites pair combination. It is based on a physical derived model (Wheelon 2001), generalized to become a "global" covariance model.

- The variance model chosen accounts for an elevation dependency which is the commonly used $1/\sin^2$ (El) model (Vermeer 1997).
- The Mátern covariance function describes temporal correlations via the time difference τ generated by different factors. Here, we let $[\alpha, \nu]$ varying freely. The $1/\sin^2$ (El) weighting factor is different for all satellite combinations and changes for each entry of the covariance matrix to model a small non-stationarity. This factor varies slowly as the elevation of the satellites.

This general model for correlation is built on models for tropospheric refractivities ($\nu = 5/6$) or multipath ($\nu = 1/2$ without elevation dependency) but takes advantage of the great flexibility of the Mátern model through the variations of the parameters ν and α in a physically plausible range. Thus, long or short correlations, as well as different smoothness can be simulated or modelled, i.e. the novelty being that all correlating factors are not described separately but mixed in a global model. Stein (1999) stressed the fact that having both parameters $[\alpha, \nu]$ allows a description of the high and low frequencies behaviour of the time series (or fields). Thus, better test statistics can be achieved compared with covariance functions having only one dependency. We cite exemplarily the spherical model or the squared exponential model, see Rasmussen and Williams (2006). Particularly the rate of decay of the spectral density at high frequency governed by ν is important for smooth processes as GPS carrier phase observations from long baselines (Sect. 4).

An estimation of the values of the parameters $[\alpha, \nu]$ can be done directly from the data using maximum likelihood estimation (MLE) or restricted MLE (RMLE)—see exemplarily Handcock and Wallis (1994). In the following work, we aim to study the impact of varying $[\alpha, \nu]$ around some "guessed" or roughly estimated values on the ratios presented in Sect. 2. It allows to make a proposal for an apriori variance covariance model based on a flexible covariance function. New insights on the works of Xu (1991, 2013) or Kutterer (1999) are consecutively given.

Although not exactly similar due to the differences between the underlying differential equations, the Mátern covariance family can be linked to the AR processes based on a comparison of the spectral density (Rasmussen and Williams 2006). Previously, Luo (2012), Luo et al. (2012) showed that AR models can fit the correlation structure of the residuals for GPS observations, the AR(1) model being insufficient for noise characterisation particularly if multipath is present. This motivates our modelization of the correlation of GPS observations with a Mátern family by varying the corresponding parameters. Moreover, in his work, Luo made use of a studentization of the residuals, the $1/\sin^2$ (El) weighting used in Eq. (16) can be considered to have the same effect. However, thanks to our model that accounts for different physical effects in a global and flexible way, the covariance doesn't need to be estimated each time from the data. This allows the use without many computational burden of fully populated VCM in least-squares adjustment and a concrete study of the impact on the ratios defined in Sect. 2. The philosophy behind this modelization is clearly different to the one used with iterative residual-based analysis.

Assumptions

In the resulting covariance matrices, all satellites are assumed to have the same correlation structure. Generally, in a positioning scenario, at least 6, but more often 8 or more satellites are present, with a relatively homogeneous sky distribution. Thus, we believe that an averaged value for $[\alpha, \nu]$ describe with sufficient accuracy the correlation structure of the GPS observations for all satellites of one session as an elevation

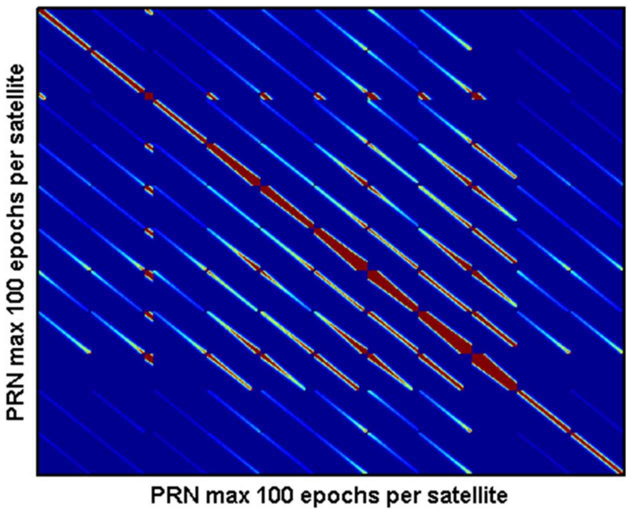


Fig. 2 Structure of the reference cofactor matrix (Mátern model with reference set $[\alpha_0, \nu_0] = [0.01, 1]$ and $\rho = 0.05$) sorted per satellite

dependency is already taken into account [Eq. (16)]. Similar argumentation is often employed for the exponential variance model (see exemplarly Luo et al. 2013).

As our case study focuses on L1 GPS observations, the description of correlations between different observation types (Teunissen et al. 1998) is not directly addressed. However, a global covariance model for L3 observations may also be resumed in a Matern covariance family. The variance model could, however, be adapted using for instance an exponential with elevation dependency (Euler and Goad 1991). This analysis is let to further study.

Parameter used for the simulations

For the simulations, a common apriori variance factor of 1 is assumed. The Mátern parameters set for $\mathbf{W}_{0\mathrm{UD}}$ are first chosen to $[\alpha_0, \nu_0] = [0.01, 1]$ corresponding to a physically plausible correlation length for GPS observations (Schön and Brunner 2008). This set remains unchanged for the computation of all covariance matrices, i.e. also sub matrices $\mathbf{C}_{\mathbf{A}}^{i,j}$ or $\mathbf{C}_{\mathbf{B}}^{i,j}$. Figure 2b is a qualitative representation of the structure of the undifferenced cofactor matrix. Taking stronger correlations as well as higher smoothness ($\alpha < 10^{-4}$ and $\nu > 3$) would lead to instabilities in the least-squares computation due to high condition number (up to 10^{10}) of the assumed covariance matrix.

Functional model

The design matrix is corresponding to a short baseline (<1 km). We took a maximum observation span per satellite of 100 epochs to avoid small sample size problems (i.e. the correlation function does not reach the 0 value inside the batch leading to instabilities).

3.2 Methodology

As we are searching for the behaviour of the loss of efficiency by missspecifying the correlation structure of the estimated covariance matrices or by neglecting correlations when $\mathbf{W}_{0\mathrm{UD}}$ is exactly known, the ratio R_{MSE} and the bias of the variance of unit weight are computed by mean of Eqs. (10) and (5), respectively. The relationship between the bias TR and $R_{\hat{\sigma}}$ is here $R_{\hat{\sigma}} = 1 + TR$.

Following to some extend the methodology of Koch et al. (2010), four scenarios for the computation of the estimated covariance matrices $\hat{\mathbf{W}}_{\mathrm{UD}}$ are chosen:

1. The correlation structure of $\hat{\mathbf{W}}_{\mathrm{UD}}$ is known (i.e. $[\alpha, \nu] = [\alpha_0, \nu_0]$), only the parameter ρ is varied.
2. ρ is taken to 0, and thus $\hat{\mathbf{W}}_{\mathrm{UD}}$ has a block diagonal structure (no correlation between different satellites). The Mátern parameters used to compute the block diagonal submatrices are changed around the reference values so that for every block diagonal sub-matrices, the same set $[\alpha] \in [0.001...0.1]$, $[\nu] \in [1/4....3/2]$ is used. The corresponding covariance matrices are called $\hat{\mathbf{W}}_{\mathrm{blockdiag}}$.
3. Diagonal VCM: The first possibility is $\hat{\mathbf{W}}_{\mathrm{UD}} = \hat{\mathbf{W}}_{\mathrm{ELEV}}$ which is corresponding to the commonly used elevation model for GPS data, neglecting correlations. The second case is $\hat{\mathbf{W}}_{\mathrm{UD}} = \mathbf{I}$**,** the identity matrix, when no correlation and equal variances are assumed.
4. A noise matrix is wrongly added to the apriori fully populated covariance matrix i.e. $\hat{\mathbf{W}}_{\mathrm{UD}} = \lambda \hat{\mathbf{W}}_{\mathrm{blockdiag}} + (1-\lambda)\, \hat{\mathbf{W}}_{\mathrm{ELEV}}, 0 \le \lambda \le 1$.

The influence of changing the true matrix by choosing another reference Mátern parameter set $[\alpha_0, \nu_0]$ or adding a noise covariance matrix are separately treated. The impact of the sample length and the design matrix are shortly addressed.

3.3 Results

3.3.1 Case 1: varying ρ

Figure 3 shows the variations of $R_{\hat{\sigma}}$ (top) and R_{MSE} (bottom) for two ρ_0, 0.05 (right) and 0.1 (left).

The loss of efficiency when the weighting parameter is misspecified is more important when $\rho > \rho_0 = 0.05$ than when $\rho < \rho_0$ as can be seen in Fig. 3d. If $\rho = 0$, the value R_{MSE} and $\sqrt{|\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}}|}$ reaches 0.006 and 0.11 mm, respectively, highlighting that the errors done by neglecting the correlations between different satellites is low.

When the weighting factor $\rho_0 = 0.1$ is considered, $R_{\mathrm{MSE}} = 0.025$ if the apriori weighting factor is 0 which corresponds to $\sqrt{|\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}}|} = 0.25$ mm. Similarly to the first

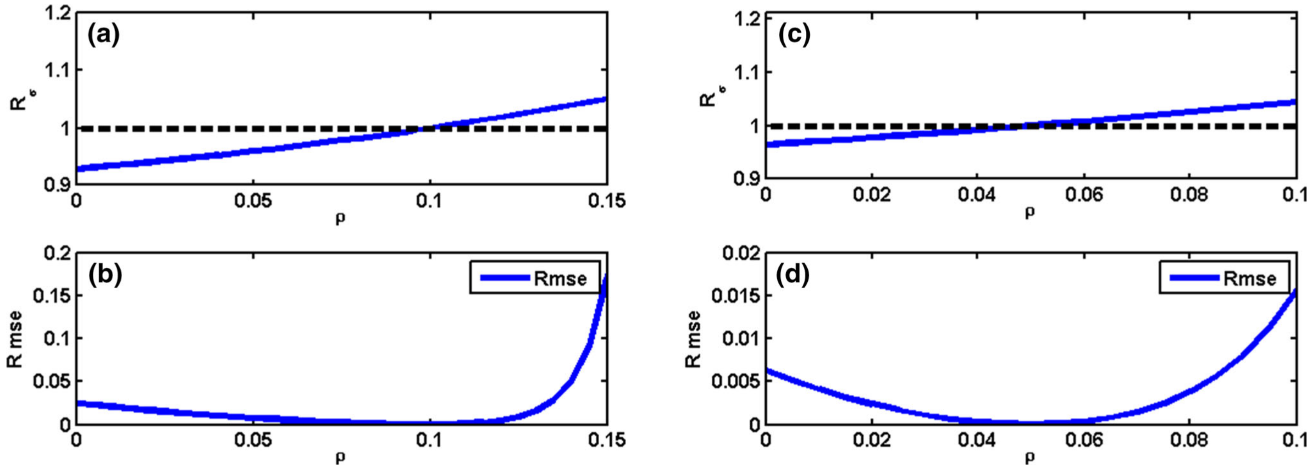


Fig. 3 Variations of the ratios with respect to changing ρ **a** $R_{\hat{\sigma}}$ versus ρ **b** R_{MSE} versus ρ for $\rho_0 = 0.1$ **c** $R_{\hat{\sigma}}$ versus ρ **d** R_{MSE} versus ρ for $\rho_0 = 0.05$. The Mátern parameters are $[\alpha, \nu] = [\alpha_0, \nu_0]$

case, if $\rho > \rho_0$, the loss of efficiency is quickly increasing and values up to $R_{\mathrm{MSE}} = 0.2$ and $\sqrt{\left|\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}}\right|} = 0.5\,\mathrm{mm}$ are found for $\rho = 0.15$. Moreover, instabilities—not represented here for sake of readability—occur for higher ρ (i.e. $\sqrt{\left|\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}}\right|} = 2$ mm for $\rho = 0.17$).

The top-plot of Fig. 3 shows the variations of $R_{\hat{\sigma}}$ with ρ. Here a more linear behaviour can be found. Neglecting correlations yields smaller values and thus statistical tests would pass, while too strong correlations yields higher values than 1, leading more easily to a rejection of the overall model test. An underestimated $\hat{\sigma}$ by using block diagonal VCM (Sect. 3.2, case 2) in the least-squares adjustment could be an indication that $\rho \neq 0$.

From these simulations, it follows that setting $\rho = 0$ and considering that $\hat{\mathbf{W}}_{\mathrm{UD}} = \hat{\mathbf{W}}_{\mathrm{blockdiag}}$ where only block diagonal matrices are computed with the reference Mátern set is a satisfactory approximation both for the R_{MSE} and $R_{\hat{\sigma}}$. It will be shown to remain valid also for higher correlation length (Sect. 3.3.5). Moreover, if the weighting factor is overestimated, the loss of efficiency is quickly increasing, leading to a higher inaccuracy at the estimates level as well as potentially instabilities due to the use of fully populated matrices with high condition number in the least-squares adjustment.

3.3.2 Case 2: block diagonal matrices: varying the estimated set $[\alpha, \nu]$

The variation of R_{MSE}, $R_{\hat{\sigma}}$ as well as $\|\Delta\mathbf{W}\|_F$ are presented in Fig. 4 when the Mátern sets $[\alpha, \nu]$ is varied, the reference set values being fixed to $[\alpha_0, \nu_0] = [0.01, 1]$ and $\rho_0 = 0.05$. Following 3.3.1, the block diagonal approximation is adopted. Figure 4a indicates that for high ν (i.e. $\nu > 0.8$), a rather exponential decrease of R_{MSE} with α is obtained, while for low ν, a small linear increase can be seen. Figure 4b shows, however, that $R_{\hat{\sigma}}$ is exponentially decreasing for all ν, the maximum values are obtained for low α (i.e. high correlation length). Thus, the overall model test is more easily rejected if erroneously correlations are considered.

Missspecifying $[\alpha, \nu]$

If the smoothness is correctly estimated ($\nu = \nu_0$) but α missspecified, the ratio R_{MSE} is increasing more strongly for lower value of α ($R_{\mathrm{MSE}} = 0.02$ for $\alpha = 0.008$) than for higher values (i.e. $R_{\mathrm{MSE}} = 0.012$ for $\alpha = 0.016$). The values of $\sqrt{\left|\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}}\right|}$ for these particular cases decrease from 0.26 to 0.21 mm, correspondingly.

The ratio $R_{\hat{\sigma}}$ is quickly increasing for $\alpha < \alpha_0$, higher values of the parameter (i.e. $\alpha > \alpha_0$) leading to a ratio close to the one given for $\nu < 1/2$.

Assuming that the correlation length is exactly known ($\alpha = \alpha_0$), a missspecification of the smoothness parameter in a range [0.4–1.3] can lead to a R_{MSE} two times higher (0.02) compared with the minimum value. For this case we found $\sqrt{\left|\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}}\right|} \in$ [0.14–0.2 mm]. The maximum is reached for small smoothness, thus for estimated matrices which covariance function at the origin is quickly decreasing (Eq. 13).

At the same time, the bias of the variance of unit weight is positive and increases strongly for $\nu > \nu_0$. However, the overall model test would reject such solutions. It should be noted that is $\nu < \nu_0$ the bias is negative, thus the quality of the solution is overestimated. Such a behaviour cannot be so clearly seen in the R_{MSE} highlighting the importance of studying both ratios.

In the GPS literature, the exponential model corresponding to $\nu = 1/2$ is quite popular [we cite exemplary El-Rabbany (1994)]. It is used to empirically fit the data for the computation of the correlation length. However, if

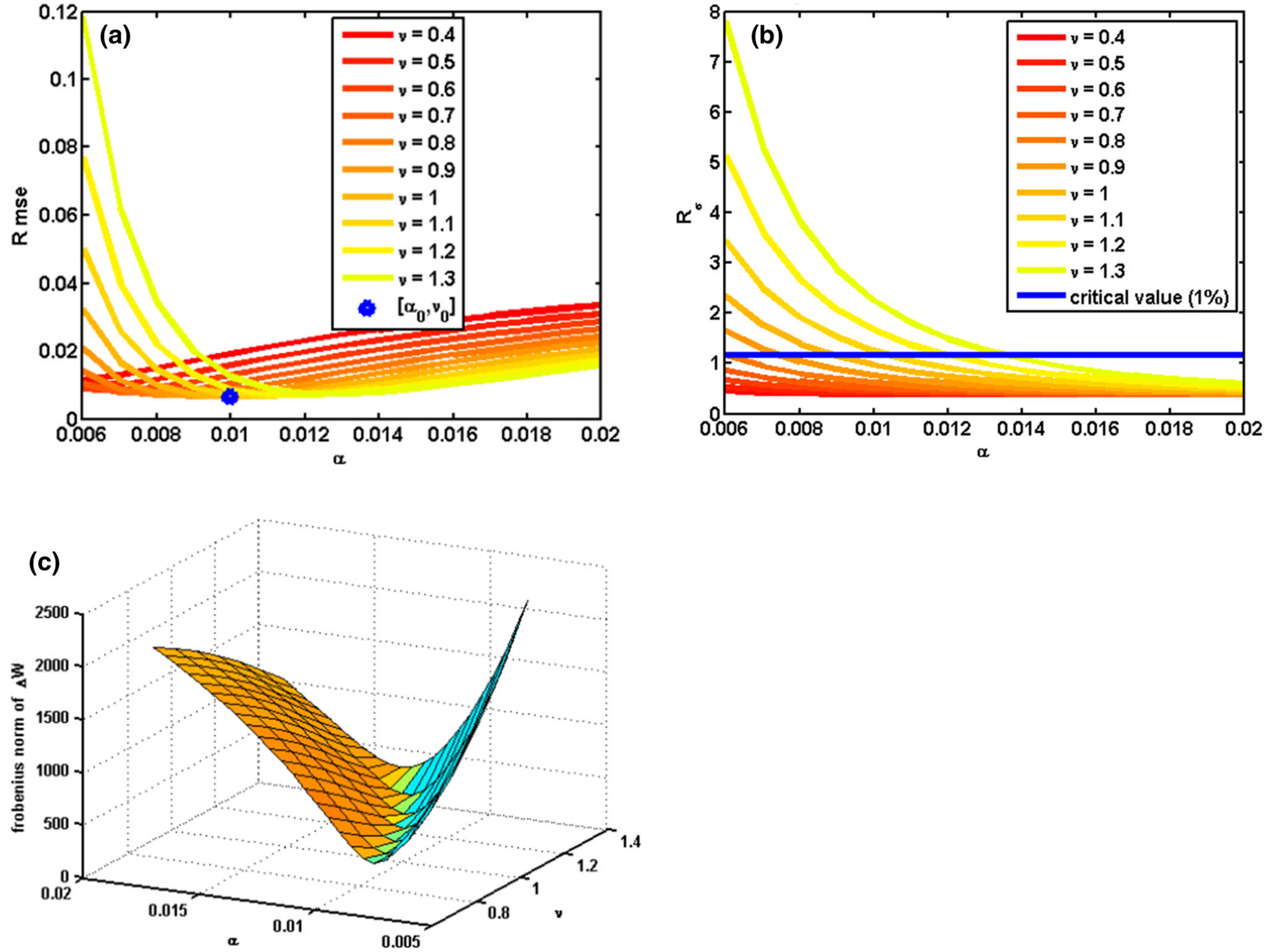


Fig. 4 **a** R_{MSE} versus α for different ν **b** bias TR of the variance of unit weight, the blue line is corresponding to TR $= 0$ **c** Frobenius norm of the matrix difference. The parameters $[\alpha, \nu]$ of $\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ are changed. The reference values are $[\alpha_0, \nu_0] = [0.01, 1]$

the data have a higher smoothness (i.e. $\nu = 1$), the loss of efficiency will be up to two times higher if at the same time $\alpha > \alpha_0$ (i.e. $R_{\mathrm{MSE}} = 0.02$ for $\alpha = 0.015$). The most important differences is highlighted by the behaviour of the bias of the variance of unit weight which is negative whichever α is taken in consideration. It points out that both the smoothness factor (behaviour of the covariance function at the origin) and the correlation length are playing an important role. However, when the covariance structure is absolutely unknown, a conservative set should be preferred ($\nu < 1$ and $\alpha > 0.05$ -i.e. less correlation), the bias of the variance of unit weight growing quickly if ν is too high. The same remark is pointed out in Rao (1967) for autoregressive processes.

Minimum of R_{MSE}

Figure 4a highlights that the minimum of R_{MSE} is reaching 0.01 for $[\alpha, \nu] = [\alpha_0, \nu_0]$. However, other Mátern sets would lead to a ratio R_{MSE} close to this minimum value. Figure 4b shows that for these same sets, the ratio $R_{\hat{\sigma}}$ would be 1 (i.e. no bias) and at the same time, the Frobenius norm of the matrices differences is minimal (Fig. 4c). Thus, for a given correlation structure, more than one Mátern set leads to optimal results in term of loss of efficiency and bias of the variance of unit weight.

3.3.3 Case 3 and 4: diagonal apriori VCM

Table 2 highlights that a higher loss of efficiency is obtained when changing the diagonal structure of the undifferenced covariance matrix. Indeed, with the identity covariance matrix R_{MSE} is more than 30 times higher than with an elevation dependent diagonal model. The solution would be moreover rejected by the overall model test since $R_{\hat{\sigma}} = 11$. A comparison of the diagonal elevation dependent model (ELEV) with block diagonal models shows that the bias of the variance of unit weight remains comparable, the loss of efficiency with $\hat{\mathbf{W}}_{\mathrm{ELEV}}$ being, however, 8 times higher than with $\hat{\mathbf{W}}_{\mathrm{blockdiag}}$.

Table 2 R_{MSE}, $\|\mathbf{\Delta W}\|_F$ and $R_{\hat{\sigma}}$ for diagonal estimated covariance matrices, short baseline scenario. Reference set $[\alpha_0, \nu_0] = [0.01, 1]$. The values for $\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ are given for comparison

Batch size 100 epochs/satellite $[\alpha_0, \nu_0] = [0.01, 1]$	$\hat{\mathbf{W}}_{\mathrm{ELEV}}$	$\mathbf{I}$	$\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ $[\alpha_0, \nu_0] = [0.01, 1]\rho = 0$
R_{MSE}	0.0493	1.9137	0.0066
$\|\Delta\mathbf{W}\|_F$	$3.4535e+03$	$3.7785e+03$	227.6539
$R_{\hat{\sigma}}$	0.9373	11.2989	0.9642

Following Sect. 2.2, a comparison between the ELEV and block diagonal model with the correct correlation structure can be done by building the difference $\sqrt{\left|\mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{BLOCKDIAG}}-\mathbf{x}_{\mathrm{ref}}} - \mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{ELEV}}-\mathbf{x}_{\mathrm{ref}}}\right|}$ which reaches a value of 0.23 mm, thus a difference between the 2 models at the submm level for this Mátern parameters set.

3.3.4 *Case 5: influence of noise matrices*

It is possible to add a diagonal covariance matrix to the apriori or assumed covariance matrices to simulate an elevation dependent noise, i.e. $\hat{\mathbf{W}}_{\mathrm{noise}} = \lambda\hat{\mathbf{W}}_{\mathrm{blockdiag}} + (1-\lambda)\,\hat{\mathbf{W}}_{\mathrm{ELEV}}, 0 \le \lambda \le 1$, which is often found in the analyses of measurements. If the true VCM is, however, free from noise ($\lambda_0 = 1$), the differences between the R_{MSE} with $\hat{\mathbf{W}}_{\mathrm{ELEV}}$ and $\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ are smaller than for case 4. R_{MSE} reached a value of 0.0492 if $\lambda_0 = 1$ with $\hat{\mathbf{W}} = \hat{\mathbf{W}}_{\mathrm{ELEV}}$ and 0.0066 with $\hat{\mathbf{W}} = \hat{\mathbf{W}}_{\mathrm{blockdiag}}$. However, if $\hat{\mathbf{W}} = \hat{\mathbf{W}}_{\mathrm{noise}}$ with $\lambda = 0.4$ is used, $R_{\mathrm{MSE}} = 0.0179$ (i.e. more weight is given to the diagonal and the solution becomes close to the one obtained with the ELEV model).

If the correct VCM is accounting for an elevation dependent noise as follows: $\hat{\mathbf{W}}_0 = \lambda_0\hat{\mathbf{W}}_{\mathrm{blockdiag}} + (1-\lambda_0)\,\hat{\mathbf{W}}_{\mathrm{ELEV}}$, $0 \le \lambda_0 \le 1$, a ratio $R_{\mathrm{MSE}} = 0.3128$ for $\lambda_0 = 0.2$ was found by taking $\hat{\mathbf{W}} = \hat{\mathbf{W}}_{\mathrm{blockdiag}}$ ($\lambda = 1$), and a bias of the variance of unit weight of 12.9386, which is a strong indication for a model missspecification. On the contrary, if $\hat{\mathbf{W}} = \hat{\mathbf{W}}_{\mathrm{ELEV}}$ ($\lambda = 0$), the $R_{\mathrm{MSE}} = 0.0112$ is 4 times lower than for the previous case, the bias of the variance of unit weight reaching -0.0125, thus close to 0. In this case, neglecting the noise matrix in the apriori VCM will lead to a higher loss of efficiency and bias than if only an elevation dependent diagonal matrix was set up, the block diagonal model being rejected by statistical tests of the variance of unit weight. However, a value of 0.40mm was computed for the difference $\sqrt{\left|\mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{BLOCKDIAG}}-\mathbf{x}_{ref}} - \mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{ELEV}}-\mathbf{x}_{ref}}\right|}$, highlighting that the error done at the estimate level remains anyway small. Thus, if there is an indication of a high noise level in the observations (i.e. $\lambda > 0.5$), taking $\hat{\mathbf{W}} = \hat{\mathbf{W}}_{\mathrm{ELEV}}$ is probably a better solution in term of loss of efficiency than trying to take correlation into account additionally.

3.3.5 *Change of the reference parameter set* $[\alpha_0, \nu_0]$

A higher loss of efficiency is reached if the reference set is changed to simulate longer correlations with $[\alpha_0, \nu_0] = [0.005, 1]$. Such a correlation structure could be a good modelization for longer baselines (i.e. >100 km), where non-modelled systematic effects such as the impact of the ionosphere, troposphere or multipath remain in the time series. For more clarity, we restrict ourselves to a comparison between the ELEV and some block diagonal model obtained by varying the Mátern parameter set of the block diagonal matrices. A low correlation scenario is also presented where $[\alpha_0, \nu_0] = [0.05, 1/2]$. The corresponding results for the two $[\alpha_0, \nu_0]$ are presented in Table 3.

As expected, changes of the reference Mátern parameter set $[\alpha_0, \nu_0]$ appears to act more strongly on the loss of efficiency R_{MSE} if $[\alpha_0, \nu_0] = [0.005, 1]$ than in the previous case, i.e. $[\alpha_0, \nu_0] = [0.01, 1]$. R_{MSE} is between 2 and 3 times higher than previously with $[\alpha_0, \nu_0] = [0.01, 1]$ for the two models $\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ and $\hat{\mathbf{W}}_{\mathrm{ELEV}}$. Moreover, the bias of the variance of unit weight is higher when the covariance matrices is misspecified. As mentioned previously, taking lower ν and higher α than the reference values is leading to results close to the one given by the elevation dependent diagonal model, too low α leading to a higher loss of efficiency. The value of $\sqrt{\left|\mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{BLOCKDIAG}}-\mathbf{x}_{\mathrm{ref}}} - \mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{ELEV}}-\mathbf{x}_{\mathrm{ref}}}\right|}$ is approximately 0.5 mm if α is not overestimated which is approximately 2 times higher than with $[\alpha_0, \nu_0] = [0.01, 1]$. However, underestimating α may lead to an erroneous MSE estimates difference of 2 mm with $[\alpha, \nu] = [0.001, 1]$.

For the low correlation scenario $[\alpha_0, \nu_0] = [0.05, 1/2]$, the loss of efficiency by missspecifying the covariance matrices is negligible (approximately 0.005 for the exemplary estimated covariance matrices). The MSE difference is at the same time much lower for this low correlation scenario (0.03 and 0.05 mm for the two relevant simulated cases). Thus, if the data are not strongly correlated, the estimates difference between a diagonal covariance model and a fully populated one (with a similar diagonal) will not exceed a submm level.

Table 3 R_{MSE}, $R_{\hat{\sigma}}$ and $\sqrt{|\mathrm{MSE}_{\hat{x}_{\mathrm{BLOCKDIAG}}-\mathbf{x}_{\mathrm{ref}}} - \mathrm{MSE}_{\hat{x}_{\mathrm{ELEV}}-\mathbf{x}_{\mathrm{ref}}}|}$ by changing the Mátern parameter set of reference for two correlation structures (low and high scenario). 100 epochs pro batch per satellite. The same design matrix was used as in Table 2

Batch size 100 epochs/satellite $[\alpha_0, \nu_0] = [0.005, 1]$	$\hat{\mathbf{W}}_{ELEV}$	$\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ $[\alpha, \nu] = [0.005, 1]$ $\rho = 0$	$\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ $[\alpha, \nu] = [0.01, 1]$ $\rho = 0$	$\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ $[\alpha, \nu] = [0.001, 1]$ $\rho = 0$
R_{MSE}	0.1112	0.0223	0.0478	1.1847
$R_{\hat{\sigma}}$	0.9104	0.9647	0.3462	20.7275
$\sqrt{\left\|\mathrm{MSE}_{\hat{\mathbf{x}}_{BLOCKDIAG}-\mathbf{x}_{ref}} - \mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{ELEV}}-\mathbf{x}_{\mathrm{ref}}}\right\|}$	xx	0.5486	0.4635	1.9072
Batch size 100 epochs/satellite $[\alpha_0, \nu_0] = [0.05, 1/2]$	$\hat{\mathbf{W}}_{\mathrm{ELEV}}$	$\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ $[\alpha, \nu] = [0.05, 1/2]$ $\rho = 0$	$\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ $[\alpha, \nu] = [0.05, 1]$ $\rho = 0$	$\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ $[\alpha, \nu] = [0.01, 1/2]$ $\rho = 0$
R_{MSE}	0.0051	0.0041	0.0053	0.0753
$R_{\hat{\sigma}}$	0.9623	0.9640	0.9640	2.5920
$\sqrt{\left\|\mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{BLOCKDIAG}}-\mathbf{x}_{ref}} - \mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{ELEV}}-\mathbf{x}_{\mathrm{ref}}}\right\|}$	xxx	0.0316	0.0592	0.1345

3.3.6 Influence of the sample length and the design matrix

Influence of the sample length

The influence of the sample length is comparable with the effect of taking a lower correlation length (higher α) or a lower smoothness factor (lower ν) into account.Simulations with batch size of up to 200 epochs a 30 s (i.e. 1.5 h of observations) were carried out showing that for all scenarios, the R_{MSE} is decreasing by increasing the batch length. For 200 epochs, R_{MSE} for $\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ and $\hat{\mathbf{W}}_{\mathrm{ELEV}}$ was 3 and 2.5 times smaller, respectively, than with 100 epochs, the bias of the variance of unit weight being slightly smaller than for short batch length, particularly for $\hat{\mathbf{W}}_{\mathrm{blockdiag}}$. The ratios with either $\hat{\mathbf{W}}_{\mathrm{blockdiag}}$ or $\hat{\mathbf{W}}_{\mathrm{ELEV}}$ are becoming similar with longer batch size (asymptotic behaviour). The difference $\sqrt{|\mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{BLOCKDIAG}}-\mathbf{x}_{\mathrm{ref}}} - \mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{ELEV}}-\mathbf{x}_{\mathrm{ref}}}|}$ was found to be 0.1 mm lower than with a batchsize of 100 epochs for the correct correlation structure.

Impact of changing the design matrix

The same simulations were carried out for different design matrices corresponding to different geometries or baseline lengths, leading to the same order of magnitude of the R_{MSE} for the Mátern parameter set of reference $[\alpha_0, \nu_0] = [0.01, 1]$. The corresponding results are not presented here for sake of shortness. The design matrix is not playing an important role in the loss of efficiency as long as approximated smoothness and correlation length values are known. Other scenarios were also tested, corresponding to single point positioning with or without clock estimation (Hoffmann-Wellenhof et al. 2001) leading to the same conclusions. The only way to have much higher R_{MSE} is to act artificially on the design matrix by letting for instance one column having random values. Please refer to "Appendix 2" for an interpretation of this "non-dependency". For the different types of design matrices which were simulated, the difference $\sqrt{|\mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{BLOCKDIAG}}-\mathbf{x}_{\mathrm{ref}}} - \mathrm{MSE}_{\hat{\mathbf{x}}_{\mathrm{ELEV}}-\mathbf{x}_{\mathrm{ref}}}|}$ varied between 0.2 and 0.8 mm for the Mátern parameter set of reference $[\alpha_0, \nu_0] = [0.01, 1]$, the higher results corresponding here to a 100 km baseline-length, highlighting the smallness of the MSE of the estimates differences between an elevation dependent diagonal model and fully populated models with the same diagonal. These important results give weight to the possibility of approximating the correlation structure of a dataset by estimating it for one batch and generalizing the value for the rest of the observations.

The next section is devoted to a case study with real data to confirm the conclusions of our simulations. $\mathbf{W}_0$ is from now on unknown, however, valuable assumptions about its structure are possible, based on known models or Maximum Likelihood Estimations of the Mátern parameters.

4 Case study GPS relative positioning

4.1 Case study short baseline

Methodology

L1 GPS data from the European Permanent Network EPN (Bruyninx et al. 2012) from two stations Zimmerwald 1 and 2 (ZIMM and ZIM2) in Switzerland are chosen as example for a short baseline positioning scenario. The distance is reaching a few meters so that distortions and specific problems due to long baselines (ambiguities resolution, ionosphere or troposphere) are mainly avoided. The observations have a 30s rate and a cutoff of 5° was taken. No ionospheric anomaly during the observation period was found. Some multipath is present which appears not to be problematic for our study since we are not searching for the "best" data, i.e. testing the influence of the covariance matrix on the results is the topic of our research. The North East Up (NEU) coordinates are computed for 51 consecutive GPS days -DOY100 until DOY150 of 2015. Batches of 100 epochs were computed starting at time GPS-SOD 10:10:10, shifted every following day of $-3'56''$ to have nearly the same geometry (i.e. the same design matrix). No change of the reference satellite occurs and the ambiguities are solved in advance. The design matrix corresponds to the previous simulations. The reference values are the long term station coordinates from the EPN solution.

The correct covariance structure of the observations is unknown and filled based on apriori knowledge, allowing a comparison of different matrices with different parameters following 3.2.2. The apriori variance factor is assessed to 2 mm. A critical value of 1.185 corresponding to the central F-distribution with $p = 0.1$ (i.e. 10%) is chosen (Teunissen 2000, p. 128). It should be pointed out that for these particular batches, the GLS estimates (i.e. the truth estimates) are unknown. Only values coming from the Feasible GLS with estimated covariance matrices are computable. NEU coordinates differences close to 0 are not obligatory correct for that particular dataset. For that reason, it is only possible to compare two estimated models with each other. Following Sect. 2.2, the root mean-squared of the coordinate difference between the ELEV and the block diagonal model is formed:

$$3\text{drms}_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}} = \sqrt{\frac{\text{trace}\left((\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}})^{\mathbf{T}}(\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}})\right)}{m}},$$

where $\mathbf{X}_{\text{blockdiag}}, \mathbf{X}_{\text{ELEV}}$ are two time series of NEU estimates of length $m = 51$ obtained with the same design matrix but different datasets as well as different estimated covariance matrices, i.e. $\hat{\mathbf{W}}_{\text{blockdiag}}$ and $\hat{\mathbf{W}}_{\text{ELEV}}$, respectively. The $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ value is expressed in [mm] and a natural comparator for a positioning case. We chose not to add a noise VCM, the MSE values and by analogy the $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ having shown to remain similar up to a few submm with or without noise VCM. Model missspecifications are anyway indicated by the overall model test as shown in Sect. 3.3.4.

As previously, the $\hat{\mathbf{W}}_{\text{blockdiag}}$ are changed by varying the Mátern parameter set in a plausible range, depending on the assumed correlation structure for GPS data (Schön and Brunner 2008). To confirm the small influence of the design matrix on the results, a second start time (GPS SOD 19:00:00) was computed with the same methodology. The results and conclusions being identical to the first time start, they are not presented for the sake of shortness.

Results

The corresponding results are presented in Fig. 5. The $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ versus ν for different α is plotted in Fig. 5a. It highlights how the differences between the two models behave with the apriori Mátern parameters. The ratio $R_{\hat{\sigma}} = \frac{E\left(\hat{\sigma}^2_{\hat{\mathbf{W}}}\right)}{\sigma^2}$ is shown in Fig. 5b with $\hat{\mathbf{W}} = \hat{\mathbf{W}}_{\text{blockdiag}}$, where $E\left(\hat{\sigma}^2_{\hat{\mathbf{W}}}\right)$ is corresponding to the mean value of $\hat{\sigma}^2_{\hat{\mathbf{W}}}$ over the 51 days. The blue line corresponds to the critical value 1.185 of the central F-distribution with $p = 0.1$. Figure 5c shows additionally the $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}}$ when the Mátern parameters are changed as well as the corresponding $3\text{drms}_{\mathbf{X}_{\text{ELEV}}}$.

The order of magnitude of $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ is below 0.1 mm when the Mátern parameter set is varied in a range of 0.5–1.2 for the smoothness ν and 0.025–0.1 s^{-1} for α. By comparing with the previous simulations (Table 3), this gives weight to a correlation structure corresponding to $\alpha \approx 0.05\ \text{s}^{-1}$ (i.e. low correlation scenario). The value of $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ being extremely small, a rough estimation of α should be sufficient to improve both the estimates as well as the variance of unit weight (Fig. 5b). However, $\hat{\sigma}_{\hat{\mathbf{W}}_{\text{ELEV}}}$ reaching already the apriori value of σ, it is questionable if taking correlations apriori into account for this short baseline, where the underlying data are very slightly correlated is necessary. Indeed, Fig. 5c points out that the difference between $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}}$ and $3\text{drms}_{\mathbf{X}_{\text{ELEV}}}$ is less than 0.05 mm and can be considered as negligible for most applications.

Nevertheless, some general remarks can be drawn from the subfigures: similarly to the simulations, Fig. 5a highlights that as ν becomes smaller, the value $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ tends to 0. Particularly for smaller α, the ELEV and block diagonal model are giving similar results. Thus, the choice of the smoothness parameter is as important as the correlation length (Stein 1999), particularly for observations with high frequencies. For this short baseline scenario, the ratio $R_{\hat{\sigma}}$ obtained by varying $\hat{\mathbf{W}}_{\text{blockdiag}}$ is below the critical value for a wide range of α if ν is smaller than 0.8. Thus, a smoothness parameter between $\nu = 1/2$ and $\nu = 1$ is reasonable together with a roughly estimated correlation length, the value of $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ being below the submm level. Higher ν by too small α (for instance $\nu > 1$ and $\alpha < 0.04$) leads to a rejection of the Mátern parameter sets (i.e. more "conservative" values should be preferred if the correlation structure is unknown as previously mentioned). Thus, approximated

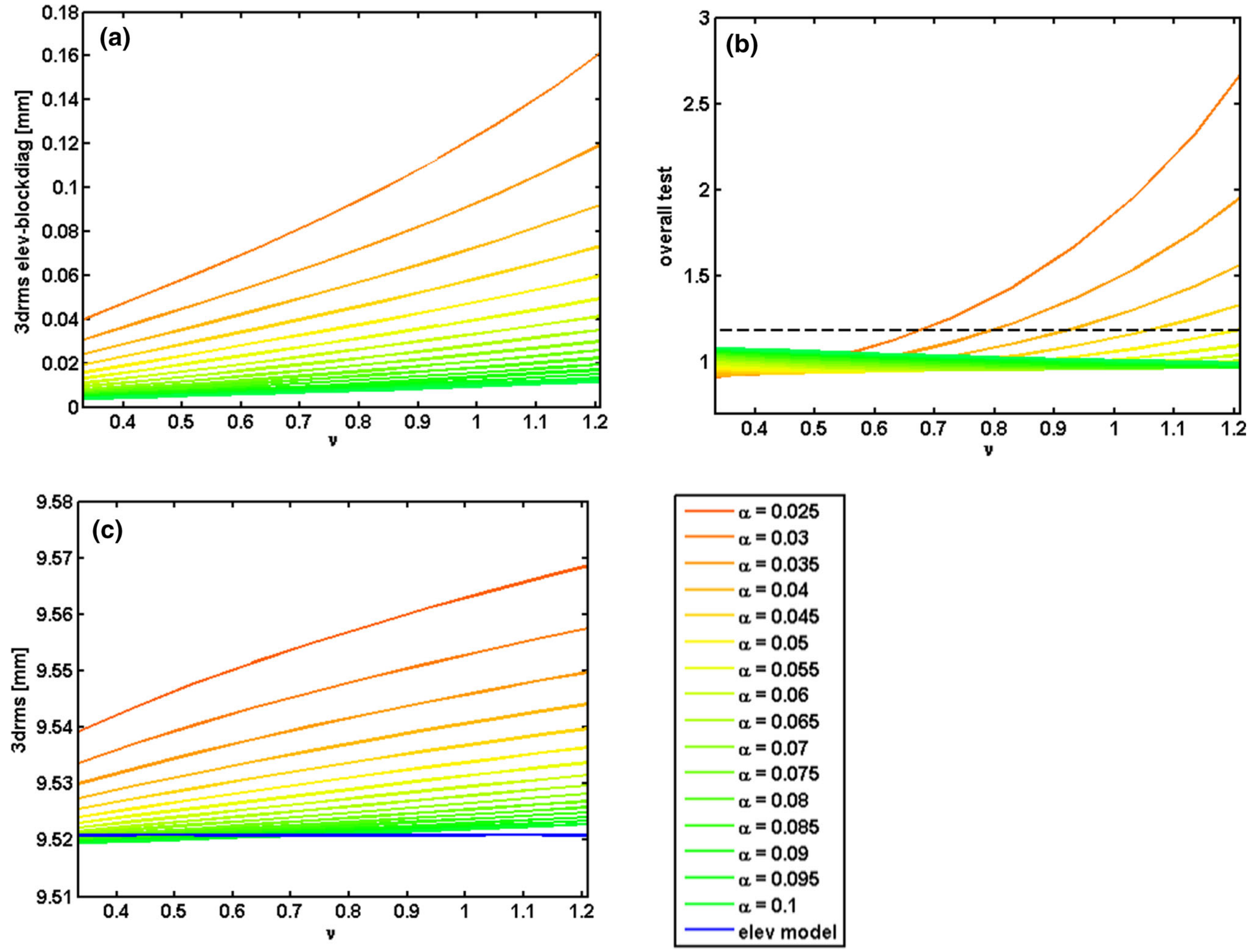


Fig. 5 **a** $3drms_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ (mm) versus ν, for different α **b** $R_{\hat{\sigma}}$ **c** $3drms_{\mathbf{X}_{\text{blockdiag}}}$ (mm) versus ν for different α. The *blue line* is corresponding to $3drms_{\mathbf{X}_{\text{ELEV}}}$, the *black dotted line* to the critical value of the central F-distribution with $p = 0.10$. Case short baseline

$[\alpha, \nu]$ for the global VCM are sufficient to have a slightly improved variance of unit weight with respect to the ELEV model.

Consecutively, a maximum likelihood estimation of $[\alpha, \nu]$ (Handcock and Wallis 1994; Stein 1999; Mardia and Watkins 1989) was computed exemplarily for one batch of 100 epochs and all 8 satellites present and gave approximate values between 0.1 and 0.05 s^{-1} for α and 1 and 1.05 for ν, thus in good agreement with the previous mentioned values.

4.2 Case study long baseline

Methodology

For this second case study, a longer baseline (approximately 80 km) between the EPN stations KRAW and ZYWI in Poland was computed. 51 days starting at GPS DOY200 of 2015 were chosen. An apriori variance factor of 8 mm is assumed. The results are presented in Fig. 6.

Results

From Fig. 6a it can be seen that the $3drms_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ value is linearly increasing with ν and approximately 3 times higher than for the short baseline length case for the same Mátern parameter set of $[0.02, 1]$. Following the results of the simulations, this gives weight to a higher correlation length than previously. Moreover, the effect of changing $[\alpha, \nu]$ on $3drms_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ and $R_{\hat{\sigma}}$ is more important than in the first case. The differences between the diagonal and the fully populated model are remaining, however, below the mm level.

Figure 6b shows clearly that $R_{\hat{\sigma}}$ is reaching a minimum, corresponding to a minimum of the bias. However, we suggest to reject Mátern sets on the left of the minimum, since it seems more logical that with increasing α and decreasing ν, $R_{\hat{\sigma}}$ tends to the value given with the ELEV model, as shown in the simulations. The set $[\alpha, \nu] = [0.015, 0.9]$ gives a $3drms_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ of approximately 0.8 mm with at the same time a minimum of $R_{\hat{\sigma}}$ of 1. This set is close to

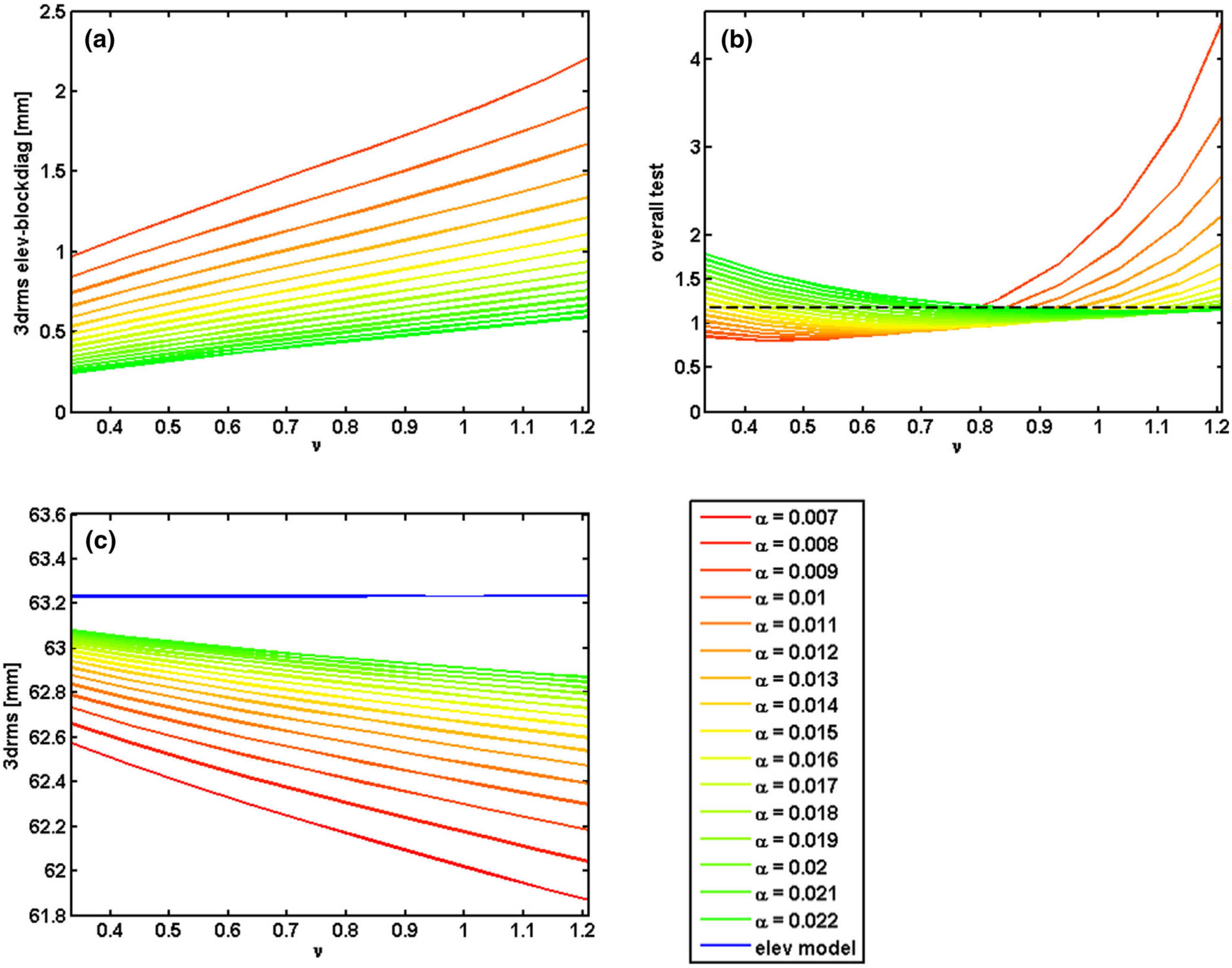


Fig. 6 **a** $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ (mm) versus ν, for different α **b** $R_{\hat{\sigma}}$ **c** $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}}$ (mm) versus ν for different α. The *blue line* is corresponding to $3\text{drms}_{\mathbf{X}_{\text{ELEV}}}$, the *black dotted line* to the critical value of the central F-distribution with $p = 0.10$. Case long baseline

the approximate mean values $[\alpha, \nu] = [0.018, 1]$ given by a maximum likelihood estimation for 8 time series of observations corresponding to satellites at different elevations from the first batch of observations. The corresponding value of $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ is 0.75 mm and is thus close to the guessed value. It should be noted that the elevation dependent diagonal model would have been rejected with the overall test for the assumed apriori variance of unit weight although the differences between the $3\text{drms}_{\mathbf{X}_{\text{blockdiag}}-\mathbf{X}_{\text{ELEV}}}$ are approximately 0.8 mm. Thus, taking correlation into account for longer baseline is valuable both in terms of loss of efficiency (i.e. the cofactor matrix of the estimates) and test statistics but not obligatory in terms of 3drms.

4.3 Proposal of how to choose the Mátern parameters of the estimated covariance matrix

The two case studies have shown that the differences in terms of estimates when different covariance matrices from fully populated (i.e. block diagonal) to diagonal are used do not exceed the mm to submm level. This holds true for a case corresponding to a 80 km baseline length where the observations were correlated. However, it would be too easy to conclude that taking correlations into account in least-squares adjustments for GPS positioning can be neglected, fully populated covariance matrices with correct Mátern parameters being important to obtain both a smaller loss of efficiency (i.e. a better covariance matrices of the estimates) as well as a better estimation of the variance of unit weight, as shown in the simulations and the case studies.

We recommend using the Mátern covariance family which includes smoothness and correlation length parameters and describe more accurately the behaviour of the observations at high and low frequencies. An elevation dependent weighting was added to model non-stationarity. Block diagonal matrices that do not account for correlations between different satellites lead to a negligible loss of efficiency compared with fully populated matrices. Moreover, the risk taken by missspecifying the weighting factor between different satellites is high and is avoided thanks to block diagonal VCM. However, for special cases of near orbiting satellite, a corresponding subcovariance matrix could be eventually computed for more accuracy. A good indication for this necessity being an underestimated ratio $R_{\hat{\sigma}}$.

For GPS observations, a smoothness value of 0.8 independent of the baseline length seems to be a good compromise. It is at the same time close to the value proposed by Kermarrec and Schön (2014). For short baselines, the data are nearly white noise and the correlation length small. From the simulations, approximated values around $\alpha = 0.05$ give a good estimation of both the a posteriori variance factor and the estimates. More exact values are unnecessary considering their impact at the estimates level as well as for the variance of unit weight. For longer baselines, -or as soon as relevant low frequencies can be proven in the double differenced observations- lower α should be preferred. Values of $0.008 \leq \alpha \leq 0.015$ are a compromise between stability of the least-squares adjustment and a more realistic VCM of the estimates (i.e. a smaller loss of efficiency and a more realistic variance of unit weight). In our long baseline scenario, the 3drms of the estimates difference when using the fully populated model and the diagonal elevation model was reaching 0.75–1.4 mm for this range of values. For more accuracy, a computation of the test statistic for the variance of unit weight by slightly varying the correlation length could be interesting, a value as close as possible to 1 for the overall test being searched.

For less computational burden, we propose to estimate the mean of the Mátern parameters corresponding to observations of satellites at different elevations for one batch via Maximum Likelihood. This is an easy way to obtain a sufficient approximation of $[\alpha, \nu]$ for a given dataset. A noise matrix should be eventually added to avoid an underestimated variance of unit weight with fully populated VCM. The amount of noise can be estimated from the observations and averaged. It should be noted that a neglection of such noise factor has only small impact on the estimates and on the loss of efficiency. For noisy datasets, it is preferable to use diagonal VCM.

Following our results, the smoothness of 1/2 used in the popular exponential covariance function is sub-optimal and yields results close to the one given by the diagonal elevation dependent model. Moreover, the estimated value of the smoothness with MLE are close to 1 with an uncertainty of +/0.05 in the two datasets used in this contribution. The a posteriori variance of unit weight was underestimated with an exponential model with respect to a higher smoothness and the correct correlation length, highlighting a high loss of efficiency although the differences was small (i.e. 0.2 mm) at the estimates level.

The w-statistics test (Teunissen 2000) which allows to detect outliers was also computed for both baselines. As shown in Li et al. (2016), an improved stochastic model leads to a lower false alarm of the w-statistics. Such results are related to the eigenvalues decomposition of the cofactor matrix of the estimates and a short explanation with a case study is presented in "Appendix 2". Similarly, the ambiguity resolution success rate will be improved (Amiri-Simkooei et al. 2016; Wang et al. 2000). Further contributions will concentrate on mathematical developments of these results as well as the impact of an improved diagonal weighting.

It should be mentioned that our study is based on trace of matrices and is giving a global indication of the loss of efficiency and the root mean-squared of the estimates differences. For these estimators, the design matrix was shown to play a minor role in the variations. The results for single parameters were not the aim of our contribution. Our results, particularly for the simulations, are general and not depending on the frequency. Therefore, they can be for instance extended to L2 or L3-measurements taking care of the chosen variance model.

5 Conclusion

The influence of changing fully populated covariance matrices in least-squares adjustments have been studied for both simulations and real cases corresponding to relative GPS positioning with double differences. The target parameters were the MSE, the 3drms as well as the aposteriori variance factor. Based on previous studies (Kermarrec and Schön 2014; Luo 2012), the Mátern family with an elevation dependent weighting was chosen to simulate fully populated VCM by varying the smoothness and the correlation length as well as the cross correlation factor. We used simulations with a reference set and showed the behaviour of the mean-squared errors of the difference of the estimates as well as the bias of the variance of unit weight for different scenarios. The results were compared with the one given with the commonly used $1/\sin^2$ diagonal variance model and are independent of the underlying dataset.

It was shown that neglecting the correlations between different satellites (i.e. block diagonal VCM) leads for our simulation scenario to a loss of efficiency close to 0. Compared with the values obtained when the cross correlation factor is two times higher than the reference value, the block diagonal approximation can be considered as less risky than wrong estimated Mátern parameters and should be, therefore, preferred.

By letting the Mátern parameters sets varying around the reference value, the minimum of loss of efficiency was shown not to be unique, corresponding to the bias of the variance of unit weight being minimal or close to zero. A parallel with the minimum of the Frobenius norm of the difference of the covariance matrices could be drawn. The influence of changing both the smoothness and the correlation length was studied. The behaviour of the ratio of the variance of unit weight was an important criterion to exclude some sets. The underlying geometry of the design matrix was not influencing the loss of efficiency. However, increasing the batch length or adding a noise matrix leaded to results closer to the one given by the corresponding elevation dependent model. The main loss of efficiency was obtained for strong correlated data when the estimates VCM is missspecified. In such cases, the root mean-squared of the coordinate difference could reach values between 0.5–1 mm, compared with values lower than 0.01 mm for low correlation scenarios.

Two real analyses with GPS data from the EPN network, with the same design matrix but different time series, confirmed the results of the simulations. The importance of the smoothness parameters was clearly shown, particularly for the long baseline scenario. Concluding the simulations and the real case study, we proposed an easy and simple way to determine the correlation structure of the GPS-observations for a use in a least-squares adjustment. It is based on the computation of a mean value of the Mátern parameter set for the observations of all satellites during one batch and all satellites. This value is integrated in a powerful covariance model developed for GPS observations. Even being a modelization, it allows to catch with enough accuracy the correlation structure and to benefit from their impact on the loss of efficiency.

Acknowledgements The authors gratefully acknowledge the EPN network and corresponding agencies for providing freely the data. Three anonymous reviewers are thanked for pointing out additional references as well as for their valuable comments which helped improve the original manuscript.

Appendix 1: BLUE and BLUP: on the only slight dependency of the ratios with the design matrix: an interpretation

Although not exactly similar, a parallel can be drawn between the results obtained with the best linear unbiased estimator (BLUE)—corresponding to our simulations using least-squares—and best linear unbiased predictor (BLUP). This parallel helps to explain the small dependencies of our results on the structure of the design matrix for relative GPS double difference positioning design for a given batch size and correlation structure.

GPS positioning strategies could be interpreted as a "coordinates interpolation". For the double differenced case, the position of the satellites is known. A parallel to kriging could be seen where interpolated values of the station can be modeled by a Gaussian process governed by prior covariances (Cressie 1993). It is clearly not exactly the case, prediction and estimation being two different statistical procedures—an estimator is using data to guess a parameter while a predictor uses the data to guess at some random values that are not part of the dataset. Moreover, in a relative positioning case with double differences, the resulting global covariance matrices are also mathematically correlated, thus the spectral density are not corresponding to simple Mátern cases. However, drawing this parallel having in mind its limitation allows a better understanding of the previous results.

Following Stein (1999) and Hannan (1970), a formula modelling the prediction error under wrong spectral density $S_1(\omega)$ can be obtained for the BLP (Best Linear Predictor) case. The MSE of the pseudo BLP reads:

$$\mathrm{MSE}_{\hat{\mathbf{x}}-\mathbf{x}_{\mathrm{ref}}} = \frac{4\pi^2 \int_{-\pi}^{\pi} S_0(\omega)/[S_1(\omega)]^2 d\omega}{\left\{\int_{-\pi}^{\pi} [S_1(\omega)]^{-1} d\omega\right\}^2}$$

where $S_0(\omega)$ is the correct spectral density. $\mathrm{MSE}_{\hat{\mathbf{x}}-\mathbf{x}_{\mathrm{ref}}}$ is the expected value of the BLP under the spectral density S_0 and S_1 the estimated spectral density.

If one assume a Mátern spectral density as in Eq. (14) with $[\alpha_0, \nu_0] = [0.01, 1]$, Fig. 7 highlights that the same shape as in Fig. 3a can be obtained by varying the Mátern parameters around the reference set. The prediction error is varying +/−0.01 around the value obtained with the reference Mátern parameter set. It should be noted that the nominal values obtained for $\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}}$ are not directly transferable to the previous GPS simulations since their meaning as well as the scaling of the spectral density are not exactly comparable. The prediction error can also be computed by taking a constant spectral density, i.e. simulating a white noise approximation. In both cases, the value $\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}}$ was found to be approximately 5 times the reference value of $\mathrm{MSE}_{\hat{x}_0-\mathbf{x}_{\mathrm{ref}}}$ which is corresponding to our previous results with the GPS double differenced design matrices for that Mátern parameters set.

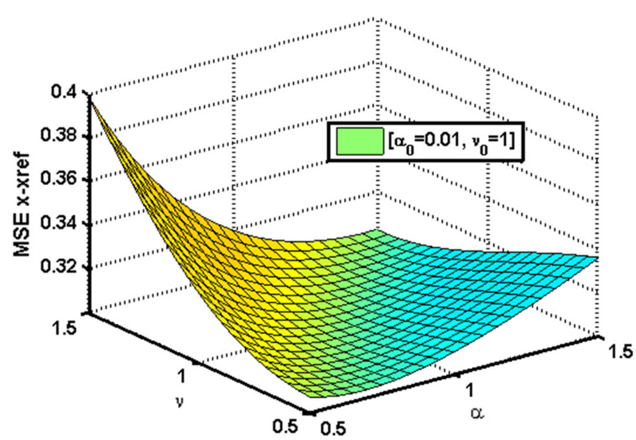


Fig. 7 Theoretical $\mathrm{MSE}_{\hat{x}-x_{\mathrm{ref}}}$ versus α, ν for the Mátern parameter set (a) $[\alpha_0, \nu_0] = [0.01, 1]$

Appendix 2: cofactor matrix of the estimates and correlations

To explain the results found for the statistical tests as well as for the 3drms of the estimates differences and its small impact with respect to the baseline length, a study of the eigenvalue decomposition of the involved matrices is here exemplarily done. For the short baseline scenario the correlation length were shown to be small or even negligible. Therefore, we concentrate on the long baseline scenario. The last batch corresponding to the last GPS DOY of our data set is here taken as an example to have a better insight on the effect of taking correlations into account in least-squares adjustment.

Following notations are used:

$$\hat{\mathbf{x}} = \mathbf{K}_{1/2}\hat{\mathbf{W}}^{-1/2}\mathbf{y} = \mathbf{K}_{1/2}\mathbf{y}_{\text{white}}$$

where $\mathbf{K}_{1/2} = \left(\mathbf{A}^{\mathrm{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\hat{\mathbf{W}}^{-1/2}$

$\mathbf{y}_{\text{white}}$ is the whitened observations vector, using either $\hat{\mathbf{W}}_{\text{ELEV}}$ or $\hat{\mathbf{W}}_{\text{blockdiag}}$, where the matrices account for double differences and correlations.

Figure 8b represents the eigenvalue repartitions of the square root of the two cofactor matrices. It can be seen that the eigenvalues of the block diagonal matrix that takes correlation into account have higher values than those of the elevation dependent diagonal matrix which explains mostly the differences seen when representing the whitened observation vectors $\mathbf{y}_{\text{white}}$ (Fig. 8a).

Figure 9 highlights the similarity between the $\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = \left(\mathbf{A}^{\mathrm{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}$ matrices. As can be seen, these matrices share the same eigenvectors, only the eigenvalues being different. Moreover, for a simple case with an AR(1) cofactor matrix, it can be seen that the structure of the GPS design matrices, which have a nearly linear behaviour (Kermarrec and Schön 2016) is responsible for this similarity, tests with purely random design matrices have shown that the eigenvectors are not similar anymore for such extreme cases.

At the estimate level, the differences will mainly come from the way the observations are whitened thanks to the estimated covariance matrix (Fig. 8a). Strong differences are not expected due to the similarity of the $\mathbf{K}_{1/2}$ matrices, and subsequently of the $\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}$ matrices.

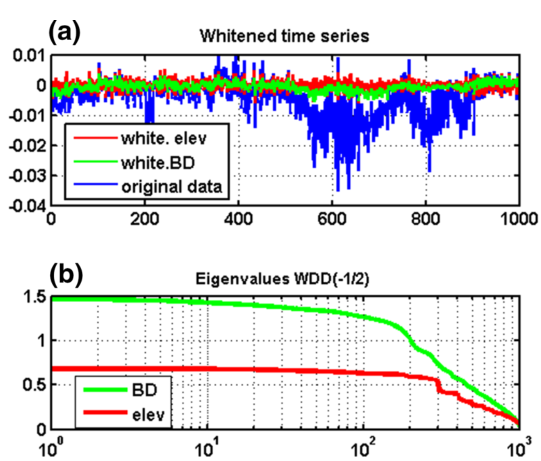


Fig. 8 **a** Whitened time series sorted per satellite (100 epochs maximum per satellite) **b** eigenvalues repartition of $\hat{\mathbf{W}}^{-1/2}$ (log plot)

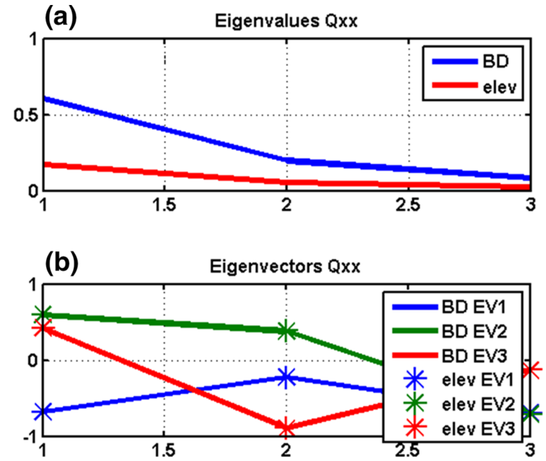


Fig. 9 **a** Eigenvalues of $\mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}$ for block diagonal matrices (BD) and **b** corresponding three eigenvectors

References

Abramowitz M, Segun IA (1972) Handbook of mathematical functions. Dover, New York

Amiri-Simkooei AR (2007) Least-squares variance component estimation: theory and applications. PhD thesis, Delft University of Technology, Delft

Amiri-Simkooei AR, Teunissen PJG, Tiberius C (2009) Application of least-squares variance component estimation to GPS observables. J Surv Eng 135(4):149–160

Amiri-Simkooei AR, Zangeneh-Nejad F, Asgari J (2013) Least-squares variance component estimation applied to GPS geometry based observation model. J Surv Eng 139(4):176–187

Amiri-Simkooei AR, Jazaeri S, Zangeneh-Nejad F, Asgari J (2016) Role of stochastic model on GPS integer ambiguity resolution success rate. GPS Solutions 20(1):51–61

Beutler G, Bauersima I, Gurtner W, Rothacher M (1987) Correlations between simultaneous GPS double difference carrier phase observations in the multistation mode: Implementation considerations and first experiences. Manuscr Geod 12:40–44

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

Bona P (2000) Precision, cross correlation, and time correlation of GPS phase and code observations. GPS Solutions 4(2):3–13

Bruyninx C, Habrich H, Söhne W, Kenyeres A, Stangl G, Völksen C (2012) Enhancement of the EUREF permanent network services and products. Geod Planet Earth IAG Symp Ser 136:27–35

Chatfield C (1989) The analysis of time series: an introduction, 4th edn. Chapman and Hall, London

Cressie N (1993) Statistics for spatial data. Wiley, New York

Dach R, Hugentobler U, Fridez P, Meindl M, Schildknecht T (2007) Bernese GPS Software Version 5.0. Astronomical Institute, University of Bern. Springer

Dufour JM (1989) Non linear hypotheses, inequality restrictions and non-nested hypotheses: exact simultaneous tests in linear regression. Econometrica 57:335–355

El-Rabbany A (1994) The effect of physical correlations on the ambiguity resolution and accuracy estimation in GPS differential positioning. PhD thesis, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada

Euler HJ, Goad CC (1991) On optimal filtering of GPS dual frequency observations without using orbit information. Bull Geod 65(2):130–143

Fuentes M (2002) Spectral methods for nonstationary processes. Biometrika 89:197–210

Grafarend EW, Awange J (2012) Applications of linear and nonlinear models. Springer, Berlin

Greene WH (2003) Econometric analysis, 5th edn. Prentice Hall, Upper Saddle River

Guttorp P, Gneiting T (2005) On the Whittle–Mátern correlation family. NRCSE technical report series no 80

Handcock MS, Wallis JR (1994) An approach to statistical spatial-temporal modeling of meteorological fields. J Am Stat Assoc 89(426):368–378

Hannan EJ (1970) Multiple time series. Wiley, New York

Hahn M, Van Mierlo J (1987) Abhängigkeit der Ausgleichungsergebnisse von der Genauigkeitsänderung einer Beobachtung. Zeitschr Vermess 3:105–115

Hoffmann-Wellenhof B, Lichtenegger H, Collins J (2001) GPS theory and practice, 5th revised edn. Springer, New York

Howind J, Kutterer H, Heck B (1999) Impact of temporal correlations on GPS-derived relative point positions. J Geod 73(5):246–258

Journel AG, Huifbregts CJ (1978) Mining geostatistics. Academic Press, New York

Kermarrec G, Schön S (2014) On the Mátern covariance family: a proposal for modeling temporal correlations based on turbulence theory. J Geod 88(11):1061–1079

Kermarrec G, Schön S (2016) Taking correlations in GPS least squares adjustments into account with a diagonal covariance matrix. J Geod 90(9):793–805

Koch KR (1999) Parameter estimation and hypothesis testing in linear models. Springer, Berlin

Koch KR, Kuhlmann H, Schuh WD (2010) Approximating covariance matrices estimated in multivariate models by estimated auto- and cross-covariances. J Geod 84(6):383–397

Koivunen AC, Kostinski AB (1999) The feasibility of data whitening to improve performance of weather radar. J Appl Meteor 38:741–749

Kutterer H (1999) On the sensitivity of the results of least-squares adjustments concerning the stochastic model. J Geod 73:350–361

Li B, Shen Y, Lou L (2011) Efficient estimation of variance and covariance components: a case study for GPS stochastic model evaluation. IEEE Trans Geosci Remote Sens 49(1):203–210

Li B, Lou L, Shen Y (2016) GNSS elevation-dependent stochastic modeling and its impacts on the statistic testing. J Surv Eng 142(2):04015012

Luo X, Mayer M, Heck B (2012) Analysing time series of GNSS residuals by means of ARIMA processes. Int Assoc Geod Symp 137:129–134

Luo X (2012) Extending the GPS stochastic model by means of signal quality measures and ARMA processes. PhD, Karlsruhe Institute of Technology

Luo X, Mayer M, Heck B, Awange JL (2013) A realistic and easy-to-implement weighting model for GNSS phase observations. IEEE Trans Geosci Remote Sens

Mardia KV, Watkins AJ (1989) On multimodality of the likelihood estimation of models for residual covariance in spatial regression. Biometrika 71(11):134–146

Mátern B (1960) Spatial variation: stochastic models and their application to some problems in forest surveys and other sampling investigation. Medd Statens Skogsforskningsinstitut 49(5):1–144

Meier S (1981) Planar geodetic covariance functions. Rev Geophys Space Phys 19(4):673–686

Puntanen S, Styan GPH (1986) The equality of the ordinary least squares estimator and the best linear unbiased estimator. Am Stat 43(3):153–161

Radovanovic RS (2001) Variance–covariance modeling of carrier phase errors for rigorous adjustment of local area networks. IAG 2001 Scientific Assembly. Budapest, Hungary, 2-7 Sept 2001

Rao C (1967) Least squares theory using an estimated dispersion matrix and its application to measurement of signals. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol 1. University of California Press, Berkeley, p 355–372

Rao C, Toutenburg H (1999) Linear models. Least-squares and alternatives, Springer, New York Second Edition

Rasmussen CE, Williams C (2006) Gaussian processes for machine learning. MIT Press, Cambridge

Santos MC, Vanicek P, Langley RB (1997) Effect of mathematical correlation on GPS network computation. J Surv Eng 123(3):101–112

Satirapod C, Wang J, Rizos C (2003) Comparing different GPS data processing techniques for modelling residual systematic errors. J Surv Eng 129(4):129–135

Schön S, Brunner FK (2008) Atmospheric turbulence theory applied to GPS carrier-phase data. J Geod 1:47–57

Spöck G, Pilz J (2008) Non-spatial modeling using harmonic analysis. VIII Int. Geostatistics congress, Santiago, p 1–10

Stein ML (1999) Interpolation of spatial data: some theory for kriging. Springer, New York

Strand ON (1974) Coefficient errors caused by using the wrong covariance matrix in the general linear model. Ann Stat 2(5):935–949

Teunissen PJG, Jonkman NF, Tiberius CCJM (1998) Weighting GPS dual frequency observations: bearing the cross of cross correlation. GPS Solutions 2(2):28–37

Teunissen PJG (2000) Testing theory and introduction. Series on mathematical geodesy and positioning. Delft University Press, The Netherlands

Teunissen PJG, Amiri-Simkooei AR (2008) Least-squares variance component estimation. J Geod 82:65–82

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

Vermeer M (1997) The precision of geodetic GPS and one way of improving it. J Geod 71:240–245

Dennis Wackerly, William Mendenhall, Scheaffer Richard L (2008) Mathematical statistics with applications, 7th edn. Thomson Higher Education, Belmont

Wang J, Stewart MP, Tsakiri M (1998) Stochastic modelling for GPS static baseline data processing. J Surv Eng 124:171–181

Wang J, Stewart P, Tsakiri M (2000) A comparative study of integer ambiguity validation procedures. Earth Planet Space 52:813–817

Wang J, Satirapod C, Rizos C (2002) Stochastic assessment of GPS carrier phase measurements for precise static relative positioning. J Geod 76(2):95–104

Wang J, Lee HK, Musa T, Rizos C (2005) Online stochastic modelling for network based GPS real-time kinematic positioning. J Glob Position Syst 4(1–2):113–119

Watson GS (1955) Serial correlation in regression analysis. Biometrika 42(3/4):327–341

Watson GS (1967) Linear least-squares regression. Ann Math. Stat 38:1679–1699

Wheelon AD (2001) Electromagnetic scintillation part I geometrical optics. Cambridge University Press, Cambridge

Whittle P (1954) On stationary processes in the plane. Biometrika 41:434–449
Wieser A, Brunner FK (2000) An extended weight model for GPS phase observations. Earth Planet Space 52:777–782
Wolf H (1961) Der Einfluss von Gewichtsänderungen auf die Ausgleichungsergebnisse. Z Vermess 86:361–362
Xu P (1991) Least squares collocation with incorrect prior information. Zeitschr Vermess 116:266–273
Xu P, Liu Y, Shen Y, Fukuda Y (2007) Estimability analysis of variance and covariance components. J Geod 81:593–602
Xu P (2013) The effect of incorrect weights on estimating the variance of unit weigth. Stud Geophy Geod 57:339–352
Yaglom AM (1987) Correlation theory of stationary and related random functions. Springer series in statistics. Springer, New York

Applied Mathematical Sciences, Vol. 11, 2017, no. 35, 1733 - 1737
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/ams.2017.75186

Possible Explanation of Empirical Values of the Matérn Smoothness Parameter for the Temporal Covariance of GPS Measurements

Gaël Kermarrec[1], Steffen Schön[1] and Vladik Kreinovich[2]

[1]Geodetic Institute
Leibniz University of Hannover
Schneiderberg 50, 30167 Hannover, Germany

[2]Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA

Abstract

The measurement errors of GPS measurements are largely due to the atmosphere, and the unpredictable part of these errors are due to the unpredictable (random) atmospheric phenomena, i.e., to turbulence. Turbulence-generated measurement errors should correspond to the smoothness parameter $\nu = 5/6$ in the Matérn covariance model. Because of this, we expected the empirical values of this smoothness parameter to be close to 5/6. When we estimated ν based on measurement results, we indeed got values close to 5/6, but interestingly, all our estimates were actually close to 1 (and slightly larger than 1). In this paper, we provide a possible explanation for this empirical phenomenon. This explanation is based on the fact that in the sensors, the quantity of interest is usually transformed into a current, and in electric circuits, current is a smooth function of time.

Mathematics Subject Classification: 86A30 62J10 76F05

Keywords: measurement errors, GPS measurements, geodesy, Matern covariance model, turbulence

1 Formulation of the Problem: An Empirical Fact That May Need Explaining

Temporal covariance: general idea. In many practical situations, we perform repeated measurements of the corresponding quantity (or quantities) at different moments of time.

Often, in data processing, measurement errors of different measurement results are assumed to be independent; see, e.g., [2]. In many cases, this assumption makes sense, since during the time between the two measurements the factors affecting the measurement change in a random way. However, when we make multiple repeated measurements, the time interval between the two consequent measurements is so small that at least some of these factors do not have time to change. As a result, there is a significant correlation between the measurement errors of two consequent measurements.

To properly process the results of the corresponding measurements, we need to know the covariance between measurement results obtained at different moments of time.

According to [3], in many practical applications, the covariance $C(T)$ between the measurements separated by time T is often well described by the following *Matérn model:*

$$C(T) = \varphi \cdot (\alpha \cdot T)^{\nu} \cdot K_{\nu}(\alpha \cdot T), \tag{1}$$

for appropriate parameters:

- α (whose meaning is that it is the inverse of the correlation time) and
- ν (that describes the smoothness of the resulting process).

Here, $K_{\nu}(x)$ is the modified Bessel function of the second type of order ν. In general, the Bessel function $J_{\alpha}(x)$ is defined as the solution to the differential equation

$$x^2 \cdot \frac{d^2y}{dx^2} + x \cdot \frac{dy}{dx} + (x^2 - \alpha^2) \cdot y = 0$$

for which $y(0) = 0$. Then, we define $I_{\alpha}(x) \stackrel{\text{def}}{=} \mathrm{i}^{-\alpha} \cdot J_{\alpha}(\mathrm{i} \cdot x)$, where $\mathrm{i} \stackrel{\text{def}}{=} \sqrt{-1}$, and

$$K_{\alpha}(x) \stackrel{\text{def}}{=} \frac{\pi}{2} \cdot \frac{I_{-\alpha}(x) - I_{\alpha}(x)}{\sin(\alpha \cdot x)}.$$

The Matérn covariance function can also be characterized by its Fourier transform – spectral density

$$S(\omega) = \frac{\varphi \cdot 2^{\nu-1} \cdot \Gamma(\nu + 1/2) \cdot \alpha^{2\nu}}{\sqrt{\pi} \cdot (\alpha^2 + \omega^2)^{\nu+1/2}},$$

where $\Gamma(x)$ is the gamma-function.

Temporal correlation of GPS measurements: what we expected. For GPS measurements, the measurement error mostly comes from the uncertainty of atmospheric propagation, and this uncertainty, in its turns, is caused by turbulence. For turbulence, we expect the power spectrum to have asymptotics $S(\omega) \sim \omega^{-8/3}$ which corresponds to $\nu = 5/6 \approx 0.83$; see, e.g., [1].

Of course, there are other factors affecting the measurement error, factors which are described by different smoothness values. Thus, we expected the empirical value of the smoothness parameter ν to be not necessarily equal to $\nu = 5/6$, but close to this value.

Temporal correlation of GPS measurements: what we actually observed. In our analysis of the results of GPS measurements, we did indeed get values close to 5/6 – in the sense that the value 5/6 was always within the confidence interval for ν. However, interestingly, all our Maximum Likelihood estimations of ν led to values between 1 and 1.05; see [1].

Again, as we have mentioned, this does not mean that there is any contradictions with the turbulence idea – the value 5/6 is still within the confidence interval for ν – but the fact that in all the cases, we get values close to 1 and always larger than 1 may need explaining.

2 A Possible Explanation

Sensors usually transform a value of a quantity into an electric current. Most sensors – whether they are photo-sensors, temperature sensors, piezoelectric sensors – transform the quantity that we want to measure into an electric current. Then, based on the value of the current, we estimate the value of the desired physical quantity.

The current smoothly depends on time. In the linear approximation, any system that processes electric circuits can be viewed as consisting of resistors, capacitors, and inductors, the basic elements of all possible electric circuits. Here:

- the voltage of the resistor is proportional to the current I,
- the voltage of the capacitor is proportional to the electric charge – i.e., to the integral of the current, and

- the voltage of the inductor is proportional to the time derivative $\frac{dI}{dt}$ of the current.

Since voltage has to be always finite, this implies that the derivative of the current is always finite – i.e., that the dependence of the current on time is always differentiable.

This implies that the dependence of the measured value on time is also differentiable. Since the measured value of the quantity is determined based on the value of the corresponding current, this implies that the measured quantity should also be a differentiable function of time.

How is this related to the smoothness parameter ν? It is known (see, e.g., [3], Section 2.4), that a random process described differentiable functions if and only if $\int \omega^2 \cdot S(\omega)\,d\omega < +\infty$.

For the Matérn covariance function, we have $S(\omega) \sim \omega^{-(2\nu+1)}$ for large ν, so $\int \omega^2 \cdot S(\omega)\,d\omega \sim \int \omega^{-(2\nu-1)}\,d\omega$. For $\omega \to \infty$, this integral is infinite when $2\nu - 1 \le 1$:

- when $2\nu - 1 < 1$, this integral is proportional to $\omega^{-(2\nu-2)} \to \infty$, and
- when $2\nu - 1 = 1$, this integral is proportional to $\ln(\omega) \to \infty$.

The integral is finite when $2\nu - 1 > 1$, i.e., when $\nu > 1$.

Thus, the fact that the dependence of the measured value on time is differentiable means that we should have $\nu > 1$.

So what value ν should we expect? The actual value ν should be close to 5/6. We want our measurements to be as accurate as possible, so we would like to have the value of the smoothness ν to be as close to 5/6 as possible. On the other hand, as we have mentioned, we will always have values of the smoothness parameter larger than 1.

Out of the values larger than 1, the smaller the value ν – i.e., the closer it is to 1 – the closer it is to 5/6.

So, for accurate measurements, we expect the corresponding value ν to be very close to 1 (but still larger than 1).

So, we have an explanation. This is exactly what we observe – we observe values ν which are close to 1 and larger than 1. Thus, we indeed get a possible explanation for the observed phenomenon.

Acknowledgements. This work was performed when Vladik was a visiting researcher with the Geodetic Institute of the Leibniz University of Hannover, a visit supported by the German Science Foundation. This work was also supported in part by NSF grant HRD-1242122.

References

[1] G. Kermarrec and S. Schön, A priori fully populated covariance matrices in least-squares adjustment – case study: GPS relative positioning, *Journal of Geodesy*, **91** (2017), no. 5, 465–484.
https://doi.org/10.1007/s00190-016-0976-8

[2] S. G. Rabinovich, *Measurement Errors and Uncertainties: Theory and Practice*, Springer Verlag, Berlin, 2005.
https://doi.org/10.1007/0-387-29143-1

[3] M. L. Stein, *Interpolation of Spatial Data: Some Theory for Kriging*, Springer Verlag, New York, 1999.
https://doi.org/10.1007/978-1-4612-1494-6

Received: June 8, 2017; Published: June 27, 2017

Acta Geod Geophys
DOI 10.1007/s40328-017-0209-5

ORIGINAL STUDY

On modelling GPS phase correlations: a parametric model

Gael Kermarrec[1] · Steffen Schön[1]

Received: 2 February 2017 / Accepted: 19 October 2017

Abstract Least-squares estimates are unbiased with minimal variance if the correct stochastic model is used. However, due to computational burden, diagonal variance covariance matrices (VCM) are often preferred where only the elevation dependency of the variance of GPS observations is described. This simplification that neglects correlations between measurements leads to a less efficient least-squares solution. In this contribution, an improved stochastic model based on a simple parametric function to model correlations between GPS phase observations is presented. Built on an adapted and flexible Mátern function accounting for spatiotemporal variabilities, its parameters are fixed thanks to maximum likelihood estimation. Consecutively, fully populated VCM can be computed that both model the correlations of one satellite with itself as well as the correlations between one satellite and other ones. The whitening of the observations thanks to such matrices is particularly effective, allowing a more homogeneous Fourier amplitude spectrum with respect to the one obtained by using diagonal VCM. Wrong Mátern parameters—as for instance too long correlation or too low smoothness—are shown to skew the least-squares solution impacting principally results of test statistics such as the apriori cofactor matrix of the estimates or the aposteriori variance factor. The effects at the estimates level are minimal as long as the correlation structure is not strongly wrongly estimated. Thus, taking correlations into account in least-squares adjustment for positioning leads to a more realistic precision and better distributed test statistics such as the overall model test and should not be neglected. Our simple proposal shows an improvement in that direction with respect to often empirical used model.

✉ Gael Kermarrec
gael.kermarrec@web.de

Steffen Schön
schoen@ife.uni-hannover.de

[1] Institut für Erdmessung (IfE), Leibniz Universität Hannover, Schneiderberg 50, 30167 Hannover, Germany

Published online: 31 October 2017

Keywords Mátern covariance function · Correlation · GPS · Realistic stochastic model

1 Introduction

Compared with the well-described functional model for GPS positioning, the stochastic model of GPS phase observations still remains improvable. A first approach to face this challenge and assess the unknown variance–covariance matrix (VCM) of the observations is to use variance–covariance estimation (VCE) techniques. We cite exemplarily Teunissen and Amiri-Simkooei (2008), Tiberius and Kenselaar (2000), Amiri-Simkooei et al. (2009, 2016), Bona (2000), Li et al. (2008, 2016). Iterative procedures have also been developed based on the whitening of the least-squares (LS) residuals (Wang et al. 2002; Satirapod et al. 2003; Leandro et al. 2005; Jin et al. 2010). Besides this somehow computational demanding approach, a second one is based on modelling the co-variance of GPS phase observations or residuals. If heteroscedasticity is widely assumed (Bischoff et al. 2005) and taken into account thanks to an elevation dependent function for the variance -mostly cosine or exponential variance, correlations are mostly disregarded due to a lack of knowledge of the correlation structure. The corresponding Variance Covariance Matrices (VCM) are thus diagonal and easier to handle in processing software. The main disadvantage of this simplification is the biased least squares solution (Koch 1999; Grafarend and Awange 2012). The proposals to model correlations of GPS phase observations are often limited to an exponential function (El-Rabbany 1994; Howind et al. 1999). The approximated correlation length is estimated by fitting the autocorrelation function of LS residuals with least-squares which was shown to be a non-optimal method to assess an accurate correlation structure (Stein 1999). In Radovanovic (2001), a linear combination of VCM accounting for correlations due to multipath described with an exponential function and a noise matrix was proposed. Correlations due to tropospheric refractivities were treated by Schön and Brunner (2008) and Kermarrec and Schön (2014) whereas the physical modelization of temporal correlations of GPS observations due to the ionosphere was empirical estimated by e.g. Wild et al. (1989). The noise of high frequency, short duration GPS observations is addressed in Moschas and Stiros (2013). From all these studies, it becomes evident that trying to model independently all correlation factors is not a straightforward task. Thus, a simple and understandable correlation model should be an useful tool to popularize the use of fully populated VCM in order to impact positively the LS solution.

In this contribution, we propose an innovative way to model elevation dependent correlations of GPS phase observations thanks to only one parametric covariance function. Using Maximum Likelihood Estimation (MLE) to determine its parameters, a wide range of correlations can be modelled without expressing them independently. Moreover, a spatiotemporal dependency allows an individual weighting of the covariance function for each satellite. Consecutively, fully populated VCM can be integrated in the weighted least-squares positioning adjustment. Besides whitening the observations, they lead to an improvement of the precision of the LS solution which becomes more realistic. The corresponding effects of this new stochastic model will be detailed thanks to a particular case study of a 80 km baseline for L1 and L3 observations, the ambiguities being fixed in advance. The conclusions will be extended to other baseline lengths.

The remainder of this papier is structured as follows: the first section provides a brief summary of the mathematical concepts of least-squares and physical correlations, introducing the Mátern model. The second section describes our proposal for GPS phase correlations. A case study concludes the contribution in a third section, giving an insight on what can be achieved thanks to this new model.

2 Mathematical concepts

2.1 Least-squares principles

When estimating a position with Global Positioning System (GPS) observations, unknown parameters such as coordinates or integer ambiguities are computed using a linearized weighted least-squares model. The corresponding functional model reads

$$\mathbf{l} = \mathbf{A}\mathbf{x} + \mathbf{v}, \quad (1)$$

where $\mathbf{l}$ corresponds to the $n \times 1$ observation vector (i.e. Observed Minus Computed vector), $\mathbf{A}$ the non-stochastic $n \times u$ design matrix with full column rank ($rk(\mathbf{A}) = u$), $\mathbf{x}$ the $u \times 1$ parameter vector to be estimated. When ambiguities are estimated, $\mathbf{A}$ and $\mathbf{x}$ are partitioned into a coordinates and ambiguities part, i.e. $\mathbf{A} = [\mathbf{A}_c, \mathbf{A}_{amb}]$ and $\mathbf{x} = [\mathbf{x}_c, \mathbf{x}_{amb}]$, respectively. $\mathbf{v}$ is the $n \times 1$ vector of the random errors. We let $E(\mathbf{v}) = \mathbf{0}$, $E(\mathbf{v}\mathbf{v}^{\mathbf{T}}) = \sigma_0^2\mathbf{W}$, where $\mathbf{W}$ is a $n \times n$ positive definite fully populated cofactor matrix, σ_0^2 the apriori variance factor and $E(.)$ denotes the mathematical expectation.

To solve for Eq. (1), the cost function $\|\mathbf{v}\|_{\mathbf{W}}^2 = (\mathbf{l} - \mathbf{A}\mathbf{x})^{\mathbf{T}}\mathbf{W}(\mathbf{l} - \mathbf{A}\mathbf{x}) = \|(\mathbf{l} - \mathbf{A}\mathbf{x})\|_{\mathbf{W}}^2$ is minimized (Koch 1999, Misra and Enge 2012) and under the previous assumptions, the estimates of the unknown $\hat{\mathbf{x}}$ are obtained by:

$$\hat{\mathbf{x}} = \left(\mathbf{A}^{\mathbf{T}}\mathbf{W}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{W}^{-1}\mathbf{l} \quad (2)$$

The apriori cofactor matrix of the estimated vector is given by

$$\mathbf{Q}_{\hat{\mathbf{x}}} = \left(\mathbf{A}^{\mathbf{T}}\mathbf{W}^{-1}\mathbf{A}\right)^{-1} \quad (3)$$

Furthermore the aposteriori variance factor of the observations $\hat{\sigma}_0^2$ is expressed as

$$\hat{\sigma}_0^2 = \frac{(\mathbf{l} - \mathbf{A}\hat{\mathbf{x}})^{\mathbf{T}}\mathbf{W}^{-1}(\mathbf{l} - \mathbf{A}\hat{\mathbf{x}})}{n - u} = \frac{\mathbf{v}^{\mathbf{T}}\mathbf{W}^{-1}\mathbf{v}}{n - u}. \quad (4)$$

These estimators are unbiased when the correct weight matrix $\mathbf{W}^{-1}$ is used. Unfortunately, this matrix is in most cases unknown. As a consequence, the so-called feasible weighted least-squares (Greene 2003) is used and $\mathbf{W}$ is replaced by its estimates $\hat{\mathbf{W}}$ which we call in the following the apriori cofactor matrix of the observations.

2.2 Mathematical correlations

In order to eliminate nuisance parameters such as the receiver clock bias, the vector of double differenced carrier phase observations between 2 stations is formed (Seeber 2003). Thus, mathematical correlations have to be taken into account during differencing and the final cofactor matrix of the observations reads (Santos et al. 1997)

$$\hat{\mathbf{W}} = \mathbf{M}\hat{\mathbf{W}}_{UD}\mathbf{M}^{\mathbf{T}} \quad (5)$$

where $\mathbf{M}$ is the matrix operator of double differencing and $\hat{\mathbf{W}}_{UD}$ the undifferenced cofactor matrix of the observations.

For two stations A and B, the global cofactor matrix reads

$$\hat{\mathbf{W}}_{UD} = \begin{bmatrix} \hat{\mathbf{W}}_{A} & \hat{\mathbf{W}}_{AB} \\ \hat{\mathbf{W}}_{BA} & \hat{\mathbf{W}}_{B} \end{bmatrix} \quad (6)$$

where $\hat{\mathbf{W}}_{A}$ and $\hat{\mathbf{W}}_{B}$ are the cofactor matrices of the observations corresponding to station A and B, respectively, $\hat{\mathbf{W}}_{AB}$ and $\hat{\mathbf{W}}_{BA}$ the correlations matrices between observations of the 2 stations. Following Schön and Brunner (2008), we let in this contribution $\hat{\mathbf{W}}_{AB} = \hat{\mathbf{W}}_{BA} = \mathbf{0}$. However, the proposed model for correlations (Sect. 3) is general enough to allow for the computation of these two matrices.

Additionally, the ionospheric-free linear combination of the carrier phase measurements can be formed by linear combination of the carrier phase observations of L1 and L2 (Misra and Enge 2012). Besides the fact that L3 ambiguities are no longer integers, the noise is increased by a factor of 3 with respect to L1 and L2 observations.

2.3 Temporal correlations

Mathematical correlations from double differencing can be easily modelled. However, temporal correlations coming from multipath, ionospheric or tropospheric variations as well as the receivers themselves have to be taken into account in $\hat{\mathbf{W}}_{UD}$. For sake of simplicity, we skip in the following the subscript UD and $\hat{\mathbf{W}}$ designs the undifferenced cofactor matrix of the observations which can be computed thanks to a covariance function. A necessary and sufficient condition for a family of functions to be a class of covariance functions is the positive definiteness (Yaglom 1987). However, this condition is not easy to check directly and for this reason, a range of standard families, positive definite and flexible enough to be used widely have been identified. In the following, we introduce shortly a general covariance family called the Mátern covariance function (Mátern 1960; Stein 1999; Gelfand et al. 2010).

Empirically, the temporal correlation function $C(t)$ of stationary processes decreases as the time t increases. In addition, different applications may exhibit different degrees of smoothness, this parameter being related to the behaviour of the correlation function at the origin (Stein 1999). The popular Mátern family meets the requirement of flexibility and is defined as

$$C(t) = \gamma(\alpha t)^{v} K_{v}(\alpha t). \quad (7)$$

where γ is a scalar. v is called the smoothness of the time series. The inverse of the Mátern correlation time α indicates how the correlations decay with increasing time (Journel and Huifbregts 1978). The modified Bessel function of order v (Abramowitz and Segun 1972) is denoted by K_{v}.

The corresponding spectral density is given by

$$S(\omega) = \frac{2^{v-1}\gamma\,\Gamma\left(v + d/2\right)\alpha^{2v}}{\pi^{d/2}\left(\omega^{2} + \alpha^{2}\right)^{v+d/2}} \quad (8)$$

where $\omega^2 = \omega_1^2 + \omega_2^2 + \cdots + \omega_d^2$ is the angular frequency, Γ the Gamma function (Abramowitz and Segun 1972). The dimension of the field d is 1 in case of time series of observations.

From Eq. (8), the behaviour of $S(\omega)$ by letting $\omega \to 0$ is both influenced by the smoothness ν and the correlation parameter α. For high frequencies, i.e. $\omega \to \infty$, the role of ν is more important. Figure 1 shows for different parameter sets $[\alpha, \nu]$ the corresponding covariance function (top) and its realization (bottom). The time series are simulated thanks to an eigenvalue decomposition of the Toeplitz covariance matrix built using Eq. (7). Smooth time series are corresponding to high ν (Fig. 1 left). A visual determination of the smoothness of the corresponding time series is difficult to assess from the correlation function itself. On the other hand, α is related to the correlation length as highlighted in Fig. 1 (right), i.e. as α increases, the correlation length decreases.

The Mátern parameters can be estimated from the data via Maximum Likelihood Estimation (Handcock and Wallis 1994), or fixed aprori (Stein 1999) which was implicitly done by Howind et al. (1999), El-Rabbany (1994) when using an exponential function to model carrier phase correlations of GPS observations. Indeed, particular cases corresponding to a smoothness of $^1/_2, 1, \infty$ are known in geodesy as the exponential covariance function, the first order Markov or the Gaussian model respectively (Whittle 1954; Grafarend and Awange 2012; Meier 1981). Other parametrizations of the Mátern covariance function presented in Eq. (7) exist as well as covariance functions that model hole effects or small negative correlations based on exponentially damped cosine functions (Zastavnyi 1993). Such functions will not be used here as the correlation length and smoothness are

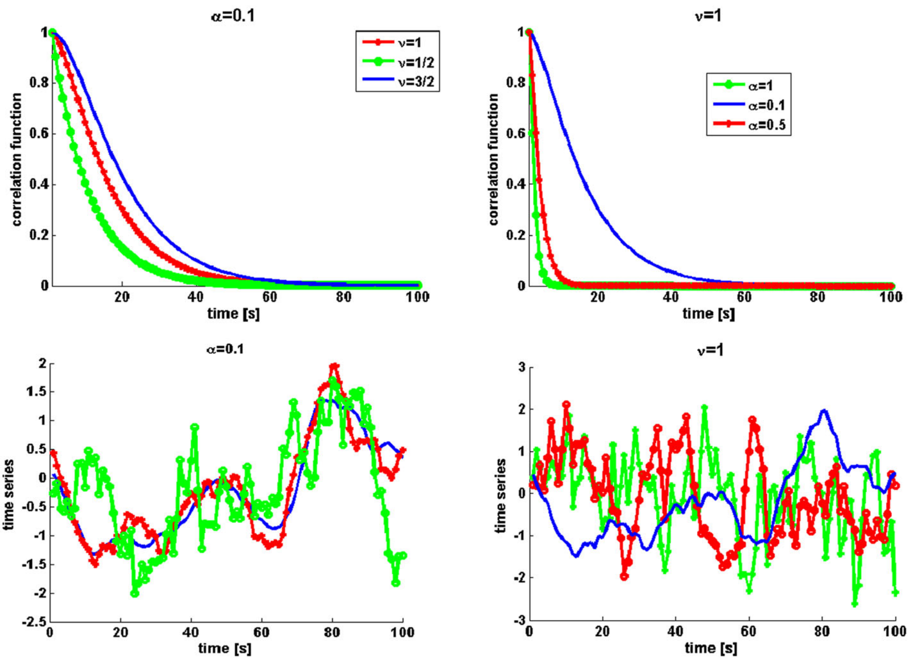


Fig. 1 Example of Mátern covariance functions by varying the parameter set $[\alpha, \nu]$. Top: correlation function for $\alpha = 0.1$ by varying ν (left) and correlation function for $\nu = 1$ by varying α (right). Bottom: corresponding time series

more important parameters than trying to take small cosine variations in consideration that may not be physically plausible. A simple, realistic and easy to use modelling of the correlations is our goal, the corresponding covariance matrices being further processed in weighted least-squares adjustments.

3 A model for correlations of GPS phase observations

3.1 Introduction

The proposed function is based on the Mátern covariance family and is directly inspired from Wheelon (2001) and Kermarrec and Schön (2014) who derived a function for correlations between satellite measurements due to turbulent fluctuations of the index of refractivity. It can be seen as an extension of this model to other kind of elevation dependent correlation factors. We note moreover that Luo (2012) showed that the correlation structure of pre-processed residuals from GPS positioning adjustment can be modelled thanks to AR or ARIMA processes. Although not exactly corresponding (Rasmussen and Williams 2006), the underlying differential equations of ARIMA and Mátern processes exhibits similarities as the spectral densities are for both cases rational and polynomial. Exemplarily, the AR(2) model can be expressed with a Mátern covariance function.

3.2 Proposal for modelling phase correlations

The proposed covariance function C between 2 observations of satellites i and j at time t and $t + \tau$ reads:

$$C_{it}^{jt+\tau} = \frac{\rho\delta}{\sin(El_i(t))\sin\left(El_j(t+\tau)\right)}(\alpha\tau)^{\nu}K_{\nu}(\alpha\tau) \tag{9}$$

where El_i and El_j are the elevations of the satellite i and j respectively. δ is a scaling parameter so that the variance equals 1 for satellites at 90° elevation using small argument approximations (Schön and Brunner 2008). ρ is a weighting factor which models the covariance between different satellites. We take here $\rho = 1$ for $i = j$ and $\rho = 0.1$ else, see Kermarrec and Schön 2014, 2017 for more details. A possible effect of underestimating ρ for $i \neq j$ in LS adjustments was shown to be a smaller aposteriori variance factor with respect to the apriori value.

This covariance function is derived from a spectral density function and remains therefore positive definite as long as the elevation is not 0°.

The variance of our model is based on the commonly used $1/\sin^2(El)$ function. In order to account for non-stationarity of the covariance (i.e. spatiotemporal dependency), a weighting factor $1/\sin(El_i(t))\sin\left(El_j(t+\tau)\right)$ is used to compute inter-satellites correlations. This factor is derived from GPS path signals through the atmosphere (Wheelon 2001).

3.3 Building fully populated VCM

Apriori fully populated covariance matrices accounting for correlations of one satellite with itself or with other one at one station can be built thanks to Eq. (9). The resulting cofactor matrix for station A as defined in Eq. (5) is given by:

$$\hat{\mathbf{W}}_A = \begin{bmatrix} \mathbf{C}_{\mathbf{A}}^{1,1} & \mathbf{C}_{\mathbf{A}}^{1,2} & \mathbf{C}_{\mathbf{A}}^{1,3} & \cdots & \mathbf{C}_{\mathbf{A}}^{1,s} \\ & \mathbf{C}_{\mathbf{A}}^{2,2} & \mathbf{C}_{\mathbf{A}}^{2,3} & & \mathbf{C}_{\mathbf{A}}^{2,s} \\ & \cdots & \mathbf{C}_{\mathbf{A}}^{3,3} & & \\ & & & & \mathbf{C}_{\mathbf{A}}^{s,s} \end{bmatrix} \quad (10)$$

where the subscript A stays for station A and the matrices $\mathbf{C}_{\mathbf{A}}^{i,j}$ are computed thanks to Eq. (9).

For a better visualization, Fig. 2 (left) presents an example of the obtained structure of the resulting fully populated covariance matrices sorted per PRN. It corresponds to a standard GPS constellation (Fig. 2 right) with 10 satellites, one batch having 100 observations and a data rate of 30 s. Depending on the processing strategy of the observations, double differenced matrices should be formed following Eq. (5).

3.4 Estimating the parameters of the covariance function from the observations

Usually, the parameters $[\alpha, v]$ are estimated by Maximum Likelihood (Stein 1999). Due to the complexity of the fully populated VCM for GPS observations, we propose two ways of estimating $[\alpha, v]$

- Estimation of one set of parameters $[\alpha, v]$ for each satellite by MLE for a batch of 1 h of observations. This set is kept and has an estimated validity of 4–6 h depending on the variations of atmospheric conditions and satellite geometry.

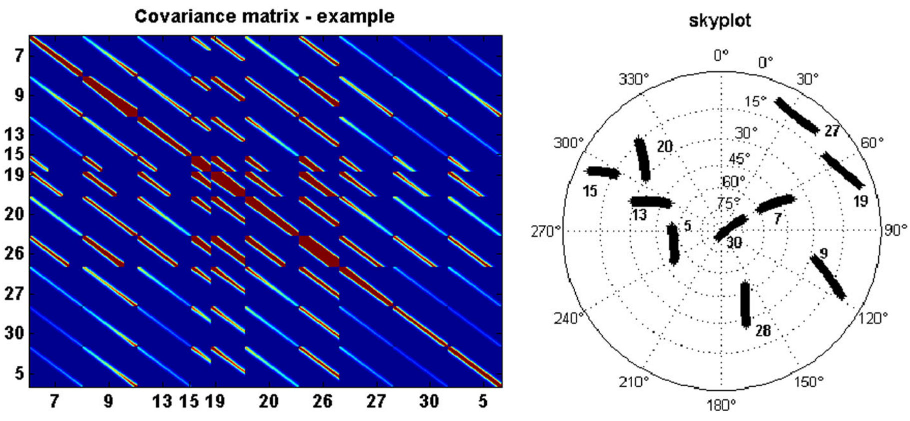


Fig. 2 An example of a fully populated covariance matrix computed with the proposed model (left). The matrix is composed of different non Toeplitz block matrices sorted per satellites. The block diagonal VCM corresponds to the temporal correlations for one satellite with itself whereas the submatrices describe the correlations between one satellite and all other ones. For this example we used $[\alpha, v] = [0.01, 1]$. The last block matrix is corresponding to the VCM of the reference satellite PRN5. The corresponding sky distribution is presented on the right

- An alternative and less computation demanding procedure is only based on the computation of the Mátern parameters by MLE for the satellite of reference. The corresponding values found were shown to be closer to a diagonal model (i.e. lower correlation length) than the previous solution and thus α can be decreased by 0.005 s^{-1} to account for this effect.

As we empirically found that values of $v > 2$ may lead to some computational problems, we propose not to allow the estimated smoothness parameter to be greater than 3/2. Due to the non-orthogonality of the Mátern parameters, this is similar to taking a smaller α, i.e. a longer correlation length.

For medium and long baselines, the double differenced observations may contain unmodelled effects. Thus, instead of computing the correlation structure based on the observations at one station, we suggest to directly estimate $[\alpha, v]$ from the double differenced observations.

In order to account for additive noise, a more general form of the covariance matrix for a station A can be used, i.e. $\hat{\mathbf{W}}' = a_{noise}\hat{\mathbf{W}}_A + (1 - a_{noise})\hat{\mathbf{W}}_{noise}$, with a_{noise} having to be estimated with $[\alpha, v]$ by MLE. For elevation dependent noise, $\hat{\mathbf{W}}_{noise} = \hat{\mathbf{W}}_{ELEV,A}$ (called the ELEV matrix) represents the elevation dependent diagonal covariance matrix corresponding to a $1/\sin^2(El)$ variance model. Depending on the noise structure that one wish to model $\hat{\mathbf{W}}_{noise}$ can also be replaced by the identity matrix. Intuitively, adding noise matrices will act on stabilizing the fully populated covariance matrices, similarly to a Tikhonov regularization (Tikhonov et al. 1995). It also impact the least-squares results, i.e. they become closer to the one given if only the corresponding diagonal VCM would have been used (Kermarrec and Schön 2017). Due to the non-orthogonality of the Mátern parameters (Gelfand et al. 2010), estimating a noise matrix will moreover lead in a shift of the corresponding set. Thus, if noise is wrongly taken into account, it will result in a smaller correlation length with noise matrix by same smoothness.

3.5 Comments on the proposed model

Our model is flexible and accounts for many factors (non-stationarity, elevation dependency, smoothness, correlation length), modelling implicitly many causes of correlations without expressing them individually. However, as every model, it remains a simplification of the correct but unknown correlation structure. It may be pointed out that also with LS-VCE procedures, simplifications are often necessary to have positive definite VCM by assuming for instance Toeplitz covariance matrices. In order to assess how the Mátern parameters influence the least-squares results, interested readers can consult the results of a sensitivity analysis based on simulations and two case studies for long and short baselines in Kermarrec and Schön (2017).

In the next section, we will focus on the effect of this new model both at the observations level and on some least-squares quantities such as the cofactor matrix of the estimates (Eq. 3) corresponding to the error ellipsoid, the aposteriori variance factor (Eq. 4) and the estimates (Eq. 2). A real positioning scenario is studied, the true covariance matrix of the observations being from now unknown.

4 Using the correlation model in a least-squares adjustment: a case study

4.1 Description of the data set

GPS L1 data from the European Permanent Network EPN (Bruyninx et al. 2012) from two stations KRAW and ZYWI are chosen as example for a medium baseline (80 km) positioning scenario. The observations have a 30 s rate, a cutoff of 3° was applied. The North East Up (NEU) coordinates are computed with double differences for 20 consecutive batches starting at GPS day DOY220, GPS-SOD 6000 s. Each batch represents 100 epochs (i.e. 3000 s). This number of epochs was chosen for two reasons. Firstly, it ensures that the correlation reaches the 0-value inside the batch so that the inversion of the VCM is accurate. Secondly, the correlation structure of small batches of observations is more difficult to assess and the values found may vary more strongly from batch to batch.

The ambiguities are solved in advance thanks to the Lambda method. The results of the case study are not impacted by the integer fixing method. Please note that they are not comparable with those found in Kermarrec and Schön (2017) for the same baseline, the methodology being different (i.e. same design matrix, different days).

The reference values for the station coordinates are the long term values from the EPN solution. The ionosphere-free linear combination L3 was additionally computed. The data were not sophistically filtered (i.e. for instance against multipath effects) in order to keep low frequencies variations in the measurements to study the whitening potential of our fully populated VCM. Because the batch length is shorter than 1 h, no tropospheric parameter was estimated. Nevertheless, estimating this parameter additionally was shown not to influence strongly the conclusions of our case study.

A realistic apriori variance factor σ_0 for double differenced observations of 4 mm was taken into account. A critical value of 4.7 mm corresponding to the Central F-distribution with $p = 0.1$ was chosen for the overall model test (Teunissen 2000). We assume that the GPS phase observations are normally distributed (Luo et al. 2011).

The correct covariance structure of the observations is unknown and computed following the methodology presented in Sect. 3. The results from three stochastic models are compared: the ID model corresponding to an identity VCM, the ELEV model where correlations are disregarded and the variance corresponds to a ${}^{1}\!/_{\sin^2(El)}$ variance model and the proposed correlation model with different parameter sets $[\alpha, \nu]$. This strategy aims to study the impact of a misspecification of the stochastic model.

The following quantities are analysed: $E(\hat{\sigma}_0)$, i.e. the mean of the aposteriori variance of unit weight over all m batches, the mean of the 3Drms of the estimates defined for one batch as $3Drms = \sqrt{\frac{trace\left(\hat{\mathbf{x}}^{\mathbf{T}}\hat{\mathbf{x}}\right)}{3}}$. The behaviour of $\mathbf{Q}_{\hat{\mathbf{x}}_c}$ and $\hat{\sigma}_0\mathbf{Q}_{\hat{\mathbf{x}}_c}$ are added exemplarily for one batch.

As the variations of $[\alpha, \nu]$ for the 20 batches of interest were below $\pm[0.005, 0.1]$ leading to negligible variations of the estimated parameters following Kermarrec and Schön (2017), the correlation structure estimated by MLE was fixed to $[\alpha, \nu]_0 = [0.012, 1.1]$ for the entire time span. The noise factor was set to 0, making use of the non-orthogonality of the Mátern parameters (Gelfand et al. 2010).

4.2 Impact on the observations

When taking correlation into account in the least-squares adjustment, the effect on the whitening of the residuals is an important criterion (Wang et al. 2002).

The fully populated matrices obtained with our model can be shown to be able to whiten the double differenced GPS observations better than the corresponding diagonal matrices, particularly in the low frequency domain. We define a whitened time series $\mathbf{l}_{\mathbf{white}}$ as $\mathbf{l}_{\mathbf{white}} = \hat{\mathbf{W}}^{-1/2}\mathbf{l}$ where $\mathbf{l}$ is the original time series. $\hat{\mathbf{W}}$ is the estimated double differenced VCM corresponding to the ID, ELEV or correlation models.

In order to have an insight on how VCM act on correlated observations, the amplitude Fourier spectra of two whitened Double Differenced time series are presented in Fig. 3. The case study from the last section (Fig. 2) is carried out. Figure 3 (left) corresponds to double differences using a low elevation satellite (10° elevation) whereas Fig. 3 (right) is using a satellite at 50° elevation. Both were observed during 200 epochs which corresponds to approximately 2 h of observations. Due to the elevation dependency of our model, the whitened double differenced time series are studentized (Luo 2012) allowing comparisons between ELEV and correlation models. Thus, the decomposition of the original—non studentized-time series is only given exemplarily. An exact explanation of why particular frequencies are present or not (i.e. multipath, site specific effects) is here on purpose not proposed.

4.2.1 *Impact of the smoothness v*

From Fig. 3, the impact of the smoothness factor v on the whitening can be seen, i.e. the noise at high frequencies is strongly increased with fully populated VCM with respect to the ELEV model (pink line). A VCM with a smoothness of ½ corresponding to an exponential correlation model with an elevation dependent variance is less able to whiten the time series than when a smoothness of 1 is considered (blue versus red lines). This result is coherent with the values found by MLE (Sect. 4.1) and particularly visible for the low elevation satellite (Fig. 3 left). Indeed, for the same $\alpha = 0.005$, a more efficient

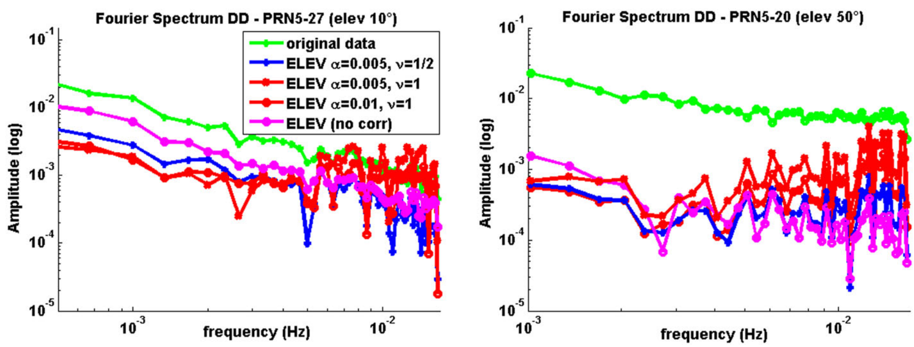


Fig. 3 Fourier decompositions of different whitened double differenced observations versus log-frequency (Hz). (Left) corresponds to a satellite starting at 10° elevation and (right) at 50° elevation. The amplitude is given as log scale. Different correlation lengths and smoothnesses were used to compute the VCM. The decomposition obtained with the diagonal VCM ELEV as well as the one of the original time series are given additionally. The length of the time series used is 200 epochs a 30 s

filtering of the low frequencies is obtained with $v = 1$ (red line) than with $v = 1/2$ (blue line).

4.2.2 Impact of the correlation parameter α

Following Eq. (8), the parameter α plays a more important role at low frequencies. This becomes evident in Fig. 3 (right) by comparing the two red lines ($\alpha = 0.005$ and $\alpha = 0.01$) corresponding to a common smoothness of 1. However, the smoothness still impacts the frequency content of the whitened time series at low frequencies [see Fig. 3 (left), blue and red lines].

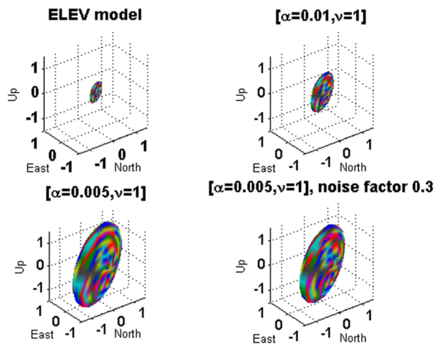
Compared with the whitening obtained with the ELEV model, it can be concluded that the fully populated VCM act both on filtering the low frequencies and increasing the high frequencies content. A lower smoothness of ½ is suboptimal for both cases. Adding a noise matrix to the VCM (not shown) would have given a higher amplitude for the low frequencies part of the spectrum, following the results given with the ELEV model.

This case study is an example. However, other batches of observations with different lengths and geometries were computed without changing the previous conclusions on the effect of $[\alpha, v]$. An exact white noise time series should not be expected and will never be obtained as the apriori VCM are not corresponding exactly to the true VCM of the observations.

4.3 Impact on the least-squares results

4.3.1 Cofactor matrix of the estimates

In this section, the impact of the fully populated covariance matrices of the observations on the cofactor matrix of the estimates is considered. One batch corresponding to a standard geometry is taken as example. In Fig. 4 (left), the 3D apriori error ellipsoids are plotted for different covariance matrices. The corresponding volumes are given in Fig. 4 (right) for $\mathbf{Q}_{\hat{\mathbf{x}}}$ and $\hat{\sigma}^2\mathbf{Q}_{\hat{\mathbf{x}}}$.



	Volume apriori	*Volume aposteriori*
ELEV	*0.0002*	*32.96*
$[\alpha,\nu]_0$	*0.0082*	*44.38*
$[\alpha,\nu]=[0.005,1]$	*1.12*	*1.08e5*
$[\alpha,\nu]=[0.005,1]$ Noise factor 0.3	*0.44*	*10.21*
ID	*2.86e-06*	*0.02*

Fig. 4 Impact of varying the VCM on the Point error ellipsoid of the estimate (left) and (right) on the volume of the corresponding apriori and aposteriori ellipsoids

From Fig. 4 (right) and (left) the well-known underestimation of the apriori precision using diagonal VCM is highlighted, the volume of the ellipsoid being up to 40 times smaller than for the fully populated case with $[\alpha, \nu]_0$. Using a low $\alpha = 0.005$ leads to an overestimation of the precision with a volume 100 times higher than for $[\alpha, \nu]_0$. This effect is related to the inverse of the covariance matrices and can be easily proven if an AR(1) process is considered (Appendix, see also Rao and Toutenburg 1999). It is thus important when using apriori VCM to control the validity of the solution by using exemplarily an overall model test.

We note that the volume of the apriori ellipsoid can be artificially decreased by adding a noise matrix [Fig. 4 (left) right bottom] by a factor two compared with $[\alpha, \nu] = [0.005, 1]$, no noise. Thus the non-unicity of the parameters is highlighted as different sets may lead to the same results for this quantity. This effect does not mean that one set should not be preferred as mentioned in Sect. 4.2, the whitening of the observations being strongly impacted by $[\alpha, \nu]$ as well as the noise factor. The orientations in space of all the ellipsoids are similar by $< 1^\circ$ as long as the same cosine variance model is used. Taking an identity model changes the orientation of the axis. However, the volume of the error ellipsoid is underestimated by more than a factor 1000 in both apriori and aposteriori case.

From Fig. 4 (right), the aposteriori ellipsoid with the ELEV model seems to be a better estimation of the precision and nearly corresponding to the reference value of 44.38. However, this higher volume has to be linked with a higher $\hat{\sigma}$ (Sect. 4.4) and the corresponding solution should be excluded with an overall model test. Thus the least-squares results should always be critically considered to avoid a model misspecification and generally if correlations are present, they should not be neglected for coherence and reliability of the solution.

4.3.2 Impact on the aposteriori variance factor and the estimates

A poorly estimated apriori VCM leads to a biased least-squares solution. The biases of the aposteriori variance factor and the estimates can be expressed literally, see Xu (2013) and Kutterer (1999). However, these formulas necessitate the knowledge of the true VCM and are essentially useful for simulations purpose (Kermarrec and Schön 2017). In real case, the global validity of the solution can be checked thanks to the overall model test. In our case moreover, the true coordinates are known. Thus, we consider that the more accurate solution corresponds to the smallest 3Drms.

Following the methodology of 3.3.1., Table 1 presents the results obtained by using different Mátern parameter sets.

The first line of Table 1 highlights the effect of a small α (i.e. a too long correlation length). It can be seen that both the 3Drms and the aposteriori variance factor will be impacted by such a misspecification. For instance, the 3Drms increases from 35% with $[\alpha, \nu] = [0.001, 1]$ compared with the reference set. At the same time, $E\left(\hat{\sigma}_{\hat{\mathbf{W}}}\right)$ is over the apriori value decreasing the number of batches available for computing the solution. Thus, we point out the importance of avoiding an underestimation of the parameter α as mentioned in 3.3.1. The effect of increasing α is not presented for sake of shortness. It leads however to a solution which becomes closer to the one given with the ELEV model.

In the second line of Table 1, the smoothness is decreased to ½ corresponding to an exponential model. A negligible increase of the 3Drms can be seen when $\alpha = \alpha_0$ by at the same time an increase of $E\left(\hat{\sigma}_{\hat{\mathbf{W}}}\right)$ by 0.35 mm (10%). When $\alpha = 0.005$, the results are

Table 1 $E\left(\hat{\sigma}_{\hat{\mathbf{W}}}\right)$ and the mean of the 3Drms are computed with VCM corresponding to different parameter sets $[\alpha, v]$. The results obtained for diagonal VCM (ELEV and ID) are given additionally. An overall model test was applied

	$[\alpha, v]_0 = [0.012, 1.1]$	$[\alpha, v] = [0.005, 1]$	$[\alpha, v] = [0.001, 1]$
$E(3Drms)$ [mm]	***55.95***	63.42	75.00
$E\left(\hat{\sigma}_{\hat{\mathbf{W}}}\right)$ [mm]	***3.43***	3.91	4.52

	$[\alpha, v] = [0.01, 1/2]$	$[\alpha, v] = [0.005, 1/2]$	ELEV	ID
$E(3Drms)$ [mm]	56.07	55.90	*57.07 (adapted σ_0)*	*82.71 (adapted σ_0)*
$E\left(\hat{\sigma}_{\hat{\mathbf{W}}}\right)$ [mm]	3.79	3.38	*9.73*	*22.79*

Bold corresponds to the Matern parameter set found by MLE whereas italic means that anadapted σ_0 was used to compute the solution

similar with the one of $[\alpha, v]_0$, pointing out the non-unicity of the best solution. Thus, the effect of taking a smoothness of ½ can be compensated by decreasing α. We highlight however that the whitening of the observations will not be similar, particularly at high frequencies (Sect. 4.1).

If the correlations are disregarded, we note a strong increase of $E\left(\hat{\sigma}_{\hat{\mathbf{W}}}\right)$ to 9.5 mm with the ELEV model and 22 mm using the ID model. Thus without adapting the apriori variance factor, no solution could be computed. However, if we increase artificially σ_0 to reach the $E\left(\hat{\sigma}_{\hat{\mathbf{W}}}\right)$, the 3Drms with the ELEV model gets similar to the value found with fully populated VCM. This solution remains statistically incorrect and fully populated models should be preferred due to the realistic $E\left(\hat{\sigma}_{\hat{\mathbf{W}}}\right)$. For the ID model, a model misspecification can be guessed as $E\left(\hat{\sigma}_{\hat{\mathbf{W}}}\right)$ reaches 22 mm and the 3Drms is 27 mm over the ELEV value.

4.4 L3 dataset and short baselines

In order to assess to which extend fully populated VCM would impact observations from the ionosphere-free linear combination, the Fourier amplitude spectrum of the double differenced time series with a satellite at 25° elevation is analysed in Fig. 5. L1 measurements exhibit stronger low frequency between 0.0005 and 0.004 Hz than L3 observations (i.e. between 2 and 10 times higher for L1 than for L3) due to the remaining unmodelled ionospheric effects. The high frequencies content is however similar for L1 and L3.

Consecutively, the impact of fully populated VCM on the whitening of L3 observations will be less important as a ML estimation of $[\alpha, v] \approx [0.028, 1]$ highlights. Due to the low correlation level, the effect on the least-squares results will not be as important as for L1 observations. Indeed, using the same data set and methodology as previously, the aposteriori variance factor was found to be 2.8 mm for the fully populated case and 3.3 mm for the ELEV case. Correspondingly, the 3Drms showed an improvement at the submm level: from 2.17 mm with the diagonal to 2.15 mm with the correlation model.

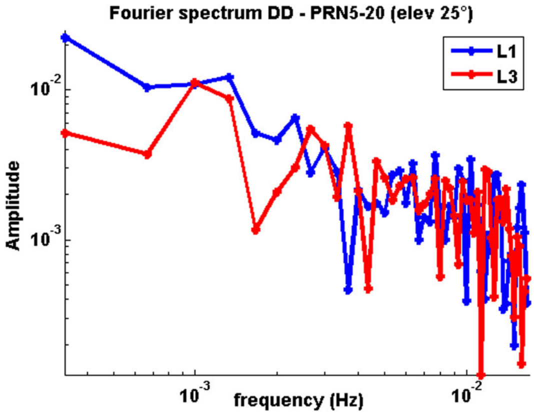


Fig. 5 Fourier amplitude spectrum for L1 and L3 double differenced observations. A satellite at 25° elevation was chosen

These variations can be considered as negligible. Thus, the filtering effect of fully populated VCM is less important than for the corresponding L1 observations. Results with short baselines are similar, the frequency content being, expect in extreme cases when for instance multipath is present, homogeneous and close to a white noise. However, this does not mean that correlations should be neglected. In case of multipath, the more realistic aposteriori variance factor obtained with fully populated VCM can influence positively the 3Drms when the overall model test is applied, particularly when batches of less than 1 h are computed.

5 Conclusions: on taking correlations into account

Because it remains easier to use diagonal covariance matrices in a least-squares adjustment for relative positioning, correlations of GPS phase observations are generally neglected. In this contribution, an innovative model based on the flexible and easy to use Mátern covariance family adapted to GPS observations was presented. Both the smoothness and the correlation length are allowed to vary in a physically plausible range. These parameters are determined by MLE either for all satellites independently or only for the reference satellite. Fully populated variance covariance matrices can be built and integrated in the least-squares adjustment.

Thanks to a particular case study corresponding to a 80 km long baseline, the model was shown to give a better whitening of the observations. If the smoothness impacts more strongly the high frequency content, the correlation parameter allows a down weighting of the low frequencies that come from unmodelled effects. Moreover, a more realistic precision together with reliable results from test statistics such as the overall model test could be obtained, the impact at the estimates level being negligible compared with results found with the cosine variance model that disregards correlations. The risks of underestimating the correlation parameter were pointed out as leading to a higher aposteriori variance factor and voluminous error ellipsoid, whereas a low smoothness gave results close to the diagonal model.

Thanks to the parametric formulation, the proposed model can be adapted to various datasets without having to use LS-VCE procedures which are computational demanding. Although in our experience a smoothness of 1 is preferable, the 3Drms under an exponential model is comparable as long as the correlation parameter is decreased accordingly by approximately 0.005. In that case, the VCM becomes invertible thanks to a close formula. As a consequence, the equivalent diagonal model proposed by Kermarrec and Schön (2016) can be easily use, allowing more reliable least-squares results by less computational burden.

Acknowledgements The authors gratefully acknowledge the EPN network and corresponding agencies for providing freely the data. Anonymous reviewers are warmly thanks for their valuable comments which helped improve the original manuscript.

Appendix: On the inverse of the covariance matrix

Studying the inverse of the fully populated covariance matrix is interesting to gain a better insight on the way correlations act on the previous results, particularly on the apriori cofactor matrix of the estimates and by extension of the ambiguities.

However, except in particular cases, VCM based on Mátern model do not have an explicit formulation of their inverse. We will therefore first present the particular case of an AR(1) process and in a second step take an example of the matrices used in this article.

AR(1) and Mátern matrices

In this case, the inverse of the covariance matrix reads

$$\mathbf{W}_{AR(1)}^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho & 0 & \dots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \ddots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}$$

where ρ is the autocorrelation function.

Thus, from a fully populated $\mathbf{W}_{AR(1)}$, only a sparse matrix remains which lines have 2 different values (1) the diagonal elements and (2) a negative value on the off-diagonal. A scaling factor $\frac{1}{1-\rho^2}$ depending on the correlation length can be further identified. This structure can be extended for Mátern covariance matrices, i.e. a factor depending on the Mátern parameter set chosen is mainly responsible for the scaling of apriori cofactor matrices of the estimates as shown in Sect. 4. The error ellipsoid may thus be artificially too voluminous if a wrong correlation structure is taken into account.

Taking noise matrix into account

In order to see how noise matrices are impacting the inverse of the covariance matrices, a particular case is chosen to compute the covariance function as defined in Eq. (9) corresponding to a satellite at 45° elevation. Only the first 100 epochs of the block diagonal matrix are analysed.

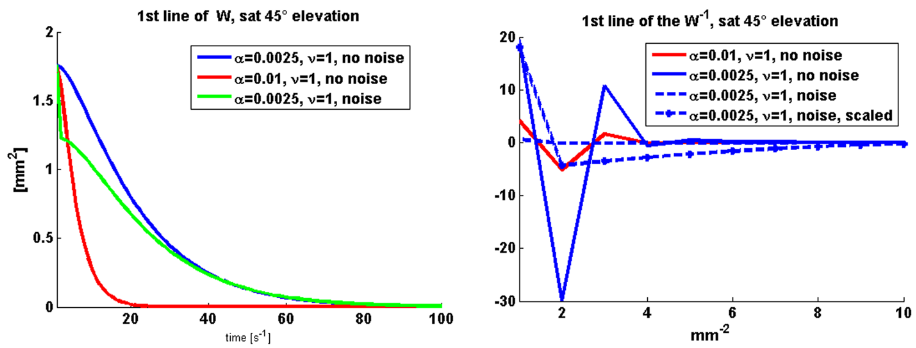


Fig. 6 Left: exemplarily line of the non-double differenced VCM for different parameter sets with and without a noise VCM and right the corresponding inverse. The satellite taken is having a 45° elevation

Fig. 6 (left) shows the corresponding line of the covariance matrix for different correlation structures. For $[\alpha, \nu] = [0.0025, 1]$ and due to the smoothness of 1, the first values of the covariance decreases slowly with time

The corresponding line of the inverse of the covariance matrix for the different cases are depicted in Fig. 6 (right). The signature of the inverse of a fully populated VCM is clearly seen when no noise matrix is taken in consideration, i.e. small oscillations around the 0–value. The amplitude of the variations increases as α decreases. If a noise matrix is added only the first oscillation remains, the other ones being replaced by a ramp, thus damping the effect of the correlations. This was highlighted by studying the whitening of the residuals and this is the reason why the results given with or without noise matrix are similar up to a given value of α and ν where the equivalence between the smoothed and the original curve is getting too high.

References

Abramowitz M, Segun IA (1972) Handbook of mathematical functions. Dover, New York

Amiri-Simkooei AR, Teunissen PJG, Tiberius C (2009) Application of least-squares variance component estimation to GPS observables. J Surv Eng 135(4):149–160

Amiri-Simkooei AR, Jazaeri S, Zangeneh-Nejad F, Asgari J (2016) Role of stochastic model on GPS integer ambiguity resolution success rate. GPS Solut 20(1):51–61

Bischoff W, Heck B, Howind J, Teusch A (2005) A procedure for testing the assumption of homoscedasticity in least squares residuals: a case study of GPS carrier-phase observations. J Geodesy 78(2005):379–404

Bona P (2000) Precision, cross correlation, and time correlation of GPS phase and code observations. GPS Solut 4(2):3–13

Bruyninx C, Habrich H, Söhne W, Kenyeres A, Stangl G, Völksen C (2012) Enhancement of the EUREF permanent network services and products. Geodesy Planet Earth IAG Symp Ser 136:27–35

El-Rabbany A (1994) The effect of physical correlations on the ambiguity resolution and accuracy estimation in GPS differential positioning. Ph.D. thesis, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada

Gelfand AR, Diggle PJ, Fuentes M, Guttorp P (2010) Handbook of spatial statistics. Chapman & Hall/CRC Handbooks of Modern Statistical Methods, London

Grafarend EW, Awange J (2012) Applications of linear and nonlinear models. Springer, Berlin

Greene WH (2003) Econometric analysis, 5th edn. Prentice Hall, Upper Saddle River

Handcock MS, Wallis JR (1994) An approach to statistical spatial-temporal modeling of meteorological fields. J Am Stat Assoc 89(426):368–378

Howind J, Kutterer H, Heck B (1999) Impact of temporal correlations on GPS-derived relative point positions. J Geodesy 73(5):246–258
Jin SG, Luo O, Ren C (2010) Effects of physical correlations on long-distance GPS positioning and zenith tropospheric delay estimates. Adv Space Res 46:190–195
Journel AG, Huifbregts CJ (1978) Mining geostatistics. Academic Press, New York
Kermarrec G, Schön S (2014) On the Mátern covariance family: a proposal for modeling temporal correlations based on turbulence theory. J Geodesy 88(11):1061–1079
Kermarrec G, Schön S (2016) Taking correlations into account with a diagonal covariance matrix. J Geodesy 90(9):793–805
Kermarrec G, Schön S (2017) Apriori fully populated covariance matrices in least-squares adjustment—case study: GPS relative positioning. J Geodesy 91(5):465–484
Koch KR (1999) Parameter estimation and hypothesis testing in linear models. Springer, Berlin
Kutterer H (1999) On the sensitivity of the results of least-squares adjustments concerning the stochastic model. J Geodesy 73:350–361
Leandro R, Santos M, Cove K (2005) An empirical approach for the estimation of GPS covariance matrix of observations. In: proceeding ION 61st annual meeting, The MITRE Corporation & Draper Laboratory, 27–29 June 2005. Cambridge
Li B (2016) Stochastic modeling of triple-frequency BeiDou signals: estimation, assessment and impact analysis. J Geodesy 90(7):593–610
Li B, Shen Y, Lou L (2008) Assessment of stochastic models for GPS measurements with different types of receivers. Chin Sci Bull 53(20):3219–3225
Li B, Lou L, Shen Y (2016) GNSS elevation-dependent stochastic modelling and its impacts on the statistic testing. J Surv Eng 142(2):04015012
Luo X (2012) Extending the GPS stochastic model y means of signal quality measures and ARMA processes. Ph.d. Karlsruhe Institute of Technology
Luo X, Mayer M, Heck B (2011) On the probability distribution of GNSS carrier phase observations GPS Solution 15(4):369–379
Mátern B (1960) Spatial variation-Stochastic models and their application to some problems in forest surveys and other sampling investigation. PhD Meddelanden Från Statens Skogsforskningsinstitut Band 49, Nr 5
Meier S (1981) Planar geodetic covariance functions. Rev Geophys Space Phys 19(4):673–686
Misra P, Enge P (2012) Global positioning system. Revised Second Edition, Ganga-Jamuna Press, Lincoln, Massachusetts
Moschas F, Stiros S (2013) Noise characteristics of high-frequency, short-duration GPS records from analysis of identical, collocated instruments. Measurement 46(2013):1488–1506
Radovanovic RS (2001) Variance-covariance modeling of carrier phase errors for rigorous adjustment of local area networks. In: IAG 2001 scientific assembly. Budapest, Sep 2–7, 2001
Rao C, Toutenburg H (1999) Linear models, least-squares and alternatives. Springer, New York
Rasmussen CE, Williams C (2006) Gaussian processes for machine learning. MIT Press, Cambridge
Santos MC, Vanicek P, Langley RB (1997) Effect of mathematical correlation on GPS network computation. J Surv Eng 123(3):101–112
Satirapod C, Wang J, Rizos C (2003) Comparing different GPS data processing techniques for modelling residual systematic errors. J Surv Eng 129(4):129–135
Schön S, Brunner FK (2008) A proposal for modeling physical correlations of GPS phase observations. J Geodesy 82(10):601–612
Seeber G (2003) Satellite geodesy. Walter de Gruyter, Berlin
Stein ML (1999) Interpolation of spatial data. Some theory for kriging. Springer, New York
Teunissen PJG (2000) Testing theory and introduction. Series on Mathematical Geodesy and Positioning. Delft University Press, Dordrecht
Teunissen PJG, Amiri-Simkooei AR (2008) Least-squares variance component estimation. J Geodesy 2008(82):65–82
Tiberius C, Kenselaar F (2000) Estimation of the stochastic model for GPS code and phase observables. Surv Rev 35(277):441–454
Tikhonov AN, Goncharsky AV, Stepanov VV, Yagola AG (1995) Numerical methods for the solution of Ill-posed problems. Kluwer Academic Publishers, Dordrecht
Wang J, Satirapod C, Rizos C (2002) Stochastic assessment of GPS carrier phase measurements for precise static relative positioning. J Geodesy 76(2):95–104
Wheelon AD (2001) Electromagnetic scintillation part I GeomETRICAL OPTICS. Cambridge University Press, Cambridge
Whittle P (1954) On stationary processes in the plane. Biometrika 41:434–449

Wild U, Beutler G, Gurtner W, Rothacher M (1989) Estimating the ionosphere using one or more dual frequency GPS receivers. In: Proceedings of the 5th international geodetic symposium on satellite positioning. Las Cruces, pp 724–736

Xu P (2013) The effect of incorrect weights on estimating the variance of unit weigth. Stud Geophys Geodesy 57:339–352

Yaglom AM (1987) Correlation theory of stationary and related random functions I basic results. Springer, New York

Zastavnyi VP (1993) Positive definite functions depending on a norm. Russ Acad Sci Doklady Math 46:112–114

GPS Solut
DOI 10.1007/s10291-017-0665-y

ORIGINAL ARTICLE

Taking correlations into account: a diagonal correlation model

Gaël Kermarrec[1] · **Steffen Schön**[1]

Received: 9 June 2017 / Accepted: 29 August 2017

Abstract The true covariance matrix of the GPS phase observations is unknown and has to be assumed or estimated. The variance of the least-squares residuals was empirically shown to have an elevation dependency and is often expressed as a sum of a constant and an exponential function. Disregarding correlations that are for instance due to atmospheric effects, the variance covariance matrices are diagonal. This simplification leads to errors in the estimates, including the float ambiguity vector, as well as to an overoptimistic precision. Thus, results of test statistics such as the outlier or the overall model test are impacted. For the particular case of GPS positioning, an innovative proposal was made to take correlations into account easily, condensed in an equivalent diagonal matrix. However, the a posteriori variance factor obtained with this simplification is strongly underestimated and in most cases the inversion of fully populated matrices has anyway to be carried out. In this contribution, we propose an alternative diagonal correlation model based on a simple exponential function to approximate the developed equivalent model. This way, correlations can be included in a diagonal variance covariance matrix without computation burden. A case study with an 80-km baseline where the ambiguities are estimated together with the coordinates in the least-squares adjustment demonstrates the potential of the model. It leads to a proposal based on the autocorrelation coefficient for fixing its parameters.

✉ Gaël Kermarrec
gael.kermarrec@web.de

Steffen Schön
schoen@ife.uni-hannover.de

[1] Institut für Erdmessung (IfE), Leibniz Universität Hannover, Schneiderberg 50, 30167 Hannover, Germany

Keywords Variance model · Weighting · GPS · Least-squares · Exponential model

Introduction

As soon as a high-precision positioning is required, an improved stochastic model that better reflects the true (co)variance of the observations is necessary. As taking correlations into account in least-squares adjustments leads to fully populated variance covariance matrices (VCM), they are often neglected and improving the stochastic model is principally synonymous with modeling the variance more accurately.

The most commonly used variance model is the elevation-dependent weighting based on the cosecant function called in the following ELEV. It was introduced by Vermeer (1997) or Collins and Langley (1999) and is implemented in many software packages, e.g., Dach et al. (2007). This model is simple and only based on the elevation of the satellites as it represents an obliquity factor. However, by giving an excessive low weight to satellites below 15°, elevation information that may be useful for better parameter estimation such as the up component is taken away. Alternatively, elevation-dependent exponential functions were proposed to fit the residuals variance, starting with Euler and Goad (1991). Later, Gerdan (1995) compared this exponential weighting with other strategies for the case of double differenced observations, whereas Han (1997) focused on its impact on ambiguity resolution. More recently, Li et al. (2016) used the exponential weighting function to analyze its impact on test statistics. Until now, the way to estimate the parameters has mainly been empirical, i.e., mostly by least-squares fitting of the variance. Other similar variance models are using SNR

Published online: 06 September 2017

(Talbot 1988, Luo et al. 2014) or CN0 values (Wieser and Brunner 2000) accounting for receiver, antenna, as well as signal quality. Exemplarily, the sigma-epsilon variance model as introduced by Hartinger and Brunner (1999) depends on the receiver antenna combination, while the sigma-delta model (Brunner et al. 1999) corrects for additional signal distortion. The goal of including station dependencies and data quality indicators is to fit or model the variance more accurately. Improved ambiguity resolution could be exemplarily achieved with a simplified version of the SNR weighting (Luo et al. 2014). Comparable results were obtained by using least-squares variance–covariance estimation (LS-VCE, Amiri-Simkooei et al. 2016) or MINQUE procedures (Wang et al. 2002). Moreover, results from test statistics such as outlier detection or overall model tests could be considered as more accurate (Li et al. 2016).

Temporal correlations between measurements are mostly empirically considered by fitting an exponential function to the autocorrelation of the least-squares residuals (Howind et al. 1999, El-Rabbany and Kleusberg 2003). Different correlation lengths were found depending on the satellites geometry or frequencies used. Alternatively, LS-VCE can take correlations into account (Teunissen and Amiri-Simkooei 2008). However, because of associated computational burden, correlations that may not only be due to atmospheric propagation but also depend on the receiver in case of high sampling rates are mainly disregarded. As a consequence, an overestimated precision of the results and potentially wrong test statistics and ambiguity resolution are risked. In their contribution, Kermarrec and Schön (2016) proposed a way to take correlations into account thanks to an equivalent model (EQUI) that resumes fully populated VCM called FULLY, in a diagonal matrix. However, this powerful proposal underestimates the a posteriori variance factor. As it plays an important role in testing the validity of the least-squares solution (Teunissen and Kleusberg 1998), an alternative should be found to cope with this issue. Besides allowing to take correlations easily into account and improving correspondingly the least-squares solution, comparisons with results from some purely diagonal variance models would be made possible.

In this contribution, a “diagonal correlation model” (DCM) for GNSS phase observations is proposed to face this challenge. It allows to account for correlations in the least-squares adjustment without computing and inverting fully populated matrices. To derive the EQUI or FULLY VCM, we will make use of the physically relevant proposal of Kermarrec and Schön (2017) to model the covariance of GPS phase measurements and briefly presented in “Appendix”.

We first summarize the principal concepts of least-squares estimations as well as ambiguity resolution. The equivalence model called EQUI and the proposed DCM are presented. Results of simulations used both to validate the model and to understand its link with other variance models are described in a second part where a sensitivity analysis by varying the two DCM parameters is carried out. In the last section, results from a detailed data analysis are presented by varying the session lengths for correlated observations from a medium baseline. We conclude with a discussion and a proposal for fixing the DCM parameters to account for correlations.

Least-squares adjustment

Due to the redundant observations, positioning with global navigation satellite system (GNSS) is done thanks to a least-squares adjustment. The corresponding linearized functional model reads

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon} \tag{1}$$

We call $\mathbf{y}$ the $n \times 1$ observation vector, $\mathbf{A}$ the deterministic $n \times u$ design matrix with full column rank ($rk(\mathbf{A}) = u$). In a relative positioning case, the number of double differences is n and the number of double differenced ambiguities is n_{amb}. The $u \times 1$ parameter vector to be estimated is $\mathbf{x}$ and $\boldsymbol{\varepsilon}$ the $n \times 1$ observation error vector. When float ambiguities and positions are estimated, $\mathbf{A}$ and $\mathbf{x}$ are partitioned, i.e., $\mathbf{A} = [\mathbf{A_C} \quad \mathbf{A_A}]$ and $\mathbf{x} = [\mathbf{x_C} \quad \mathbf{x_{A,float}}]^{\mathbf{T}}$. The $n \times 3$ matrix $\mathbf{A_C}$ and the $n \times n_{amb}$ matrix $\mathbf{A_A}$ are the design matrices of the coordinates and ambiguities, respectively. $\mathbf{x_C}$ corresponds to the vector of coordinates and $\mathbf{x_{A,float}}$ to the float ambiguity vector. Additionally, a tropospheric parameter can be estimated and the design matrix has to be adapted accordingly. We assume further that $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathbf{T}}) = \sigma_0^2\mathbf{W}_0$, where $\mathbf{W}_0$ is a $n \times n$ positive definite fully populated cofactor matrix of the residuals, σ_0^2 the a priori variance factor. The mathematical expectation is denoted by $E(.)$.

From Koch (1999), the solution of (1) reads $\hat{\mathbf{x}}_0 = \left(\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{-1}\mathbf{y}$. $\hat{\mathbf{x}}$ is the solution computed with the a priori guessed VCM $\hat{\mathbf{W}}$, i.e., $\hat{\mathbf{x}} = \left(\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{y}$. The slope matrix K which projects the observations into the parameter space is defined as $\hat{\mathbf{x}} = \mathrm{K}y$ with

$$\mathrm{K} = \left(\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1} \tag{2}$$

We note $\hat{\mathbf{W}} = \mathbf{W}_0 + \Delta\mathbf{W}$ where $\Delta\mathbf{W}$ is the difference between the assumed and the true cofactor matrix $\mathbf{W}_0$. The variance factor is unbiased under $\mathbf{W}_0$ and given by $\hat{\sigma}^2_{\mathbf{W_0}} = \frac{(\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}_0)^{\mathbf{T}}\mathbf{W}_0^{-1}(\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}_0)}{n-u}$, whereas using the approximated VCM, we obtain $\hat{\sigma}^2_{\hat{\mathbf{W}}} = \frac{(\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}^{\mathbf{T}})\hat{\mathbf{W}}^{-1}(\mathbf{y}-\mathbf{A}\hat{\mathbf{x}})}{n-u}$. Although the least-squares estimator remains unbiased under the use of $\hat{\mathbf{W}}$, an incorrect specification of the VCM leads to errors. We shortly review the corresponding effects on characteristic quantities, following Kermarrec and Schön (2017):

- Impact on the a posteriori variance factor
  From Xu (2013) or Koch (1999), $E\left(\hat{\sigma}^2_{\hat{\mathbf{W}}}\right)$ can be expressed as

$$E\left(\hat{\sigma}^2_{\hat{\mathbf{W}}}\right) = \hat{\sigma}^2_{\mathbf{W_0}} + tr\left\{\left(\mathbf{I} - \hat{\mathbf{W}}^{-1}\mathbf{A}\left(\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathbf{T}}\right)\hat{\mathbf{W}}^{-1}\Delta\mathbf{W}\right\}\frac{\hat{\sigma}^2_{\mathbf{W_0}}}{n-u} \quad (3)$$

  The trace term tr represents the bias due to the estimated VCM.
- Loss of efficiency
  The following ratio can be used to analyze the loss of efficiency when using $\hat{\mathbf{W}}$ instead of $\mathbf{W}_0$ (Stein 1999):

$$R_{\mathrm{MSE}} = \frac{\mathrm{MSE}_{\hat{\mathbf{x}}-\mathbf{x}}}{\mathrm{MSE}_{\mathbf{x}-\hat{\mathbf{x}}_0}} = \frac{\mathrm{tr}\left(\left[\hat{\mathbf{W}}^{-1}\mathbf{A}\left(\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-2}\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\right]\mathbf{W}_0\right)}{\mathrm{tr}\left(\left(\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{-1}\mathbf{A}\right)^{-1}\right)} - 1 \quad (4)$$

- Mean square error (MSE) and root-mean-square (RMS)
  The value $\mathrm{MSE}_{\hat{\mathbf{x}}-\mathbf{x}_0} = \left(\mathrm{tr}\left(\left[\hat{\mathbf{W}}^{-1}\mathbf{A}\left(\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-2}\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\right]\mathbf{W}_0\right) - \mathrm{tr}\left(\left(\mathbf{A}^{\mathbf{T}}\mathbf{W}_0^{-1}\mathbf{A}\right)^{-1}\right)\right)^{1/2}$ allows getting an order of magnitude in (mm) of the theoretical differences between correct and incorrect estimates for a particular design matrix (Strand 1974). In the case of data analysis, the true VCM is unknown, and the root-mean-square of the estimates called 3DRMS is used, i.e., $3\mathrm{DRMS} = \sqrt{\frac{tr\left(\hat{\mathbf{x}}^{\mathbf{T}}\hat{\mathbf{x}}\right)}{3}}$.

Please note that through this contribution, we understand as "coordinate precision" the sum of the three first values of the diagonal of the cofactor matrix of the estimates $\left(\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}$ corresponding to the coordinate's part.

Ambiguities fixing and validation strategy

We use the Lambda method (Teunissen 1995) to fix the float ambiguity to an integer vector. The results are two vectors of integer candidates $\hat{\mathbf{x}}^{\mathbf{1}}_{\mathbf{A,fix}}, \hat{\mathbf{x}}^{\mathbf{2}}_{\mathbf{A,fix}}$, which correspond to the two smallest values of the distance between the float and the fixed ambiguity vector, i.e.,

$$\left\|\hat{\mathbf{x}}^{\mathbf{1}}_{\mathbf{A,fix}} - \hat{\mathbf{x}}_{\mathbf{A,float}}\right\|_{\mathbf{Q_A}} \sqrt{\left(\hat{\mathbf{x}}^{\mathbf{1}}_{\mathbf{A,fix}} - \hat{\mathbf{x}}_{\mathbf{A,float}}\right)^T \mathbf{Q}_{\mathbf{A}}^{-1}\left(\hat{\mathbf{x}}^{\mathbf{1}}_{\mathbf{A,fix}} - \hat{\mathbf{x}}_{\mathbf{A,float}}\right)} = d_1$$
$$\left\|\hat{\mathbf{x}}^{2}_{\mathbf{A,fix}} - \hat{\mathbf{x}}_{\mathbf{A,float}}\right\|_{\mathbf{Q_A}} = \sqrt{\left(\hat{\mathbf{x}}^{2}_{\mathbf{A,fix}} - \hat{\mathbf{x}}_{\mathbf{A,float}}\right)^T \mathbf{Q}_{\mathbf{A}}^{-1}\left(\hat{\mathbf{x}}^{2}_{\mathbf{A,fix}} - \hat{\mathbf{x}}_{\mathbf{A,float}}\right)} = d_2 \quad (5)$$

We have $d_1 \leq d_2$. For the sake of comparison of the results given with different VCM, we introduce additionally the Euclidian distance between the fixed and the float ambiguity vector as

$$d_1^* = \sqrt{\left(\hat{\mathbf{x}}^{\mathbf{1}}_{\mathbf{A,fix}} - \hat{\mathbf{x}}_{\mathbf{A,float}}\right)^T \left(\hat{\mathbf{x}}^{\mathbf{1}}_{\mathbf{A,fix}} - \hat{\mathbf{x}}_{\mathbf{A,float}}\right)} \Big/ n_{\mathrm{amb}} \quad (6)$$

To avoid a fixing to an incorrect integer, different strategies or tests were proposed. In this contribution, and because of its popularity, the discriminant test called the ratio test or R-ratio (Euler and Schaffrin 1991) is retained. The R-ratio ${}^{d_1}/_{d_2}$ is compared to an empirical threshold value which is often assumed to be $\mu = 0.5$ (Wei and Schwarz 1995) or 1/3 (Leick et al. 2015). If the R-ratio is smaller than μ, the float solution is taken into consideration instead of a possibly wrongly fixed ambiguity vector. If the ambiguities can be fixed to integers, the corrected estimate $\hat{\mathbf{x}}_{\mathbf{C,fixed}}$ reads $\hat{\mathbf{x}}_{\mathbf{C,fixed}} = \hat{\mathbf{x}} - \mathbf{Q_{CA}}\mathbf{Q}_{\mathbf{A}}^{-1}\left(\hat{\mathbf{x}}_{\mathbf{A,float}} - \hat{\mathbf{x}}_{\mathbf{A,fixed}}\right)$. Interested readers may refer to Wang et al. (2000) for a comparison of the different validation procedures.

Stochastic model

The stochastic model plays an important role in obtaining an optimal solution of the least-squares problem (1), particularly when the functional relationship is not completely known. Exemplarily, the mean behavior of atmospheric refraction can be well modeled, e.g., by estimating an additional tropospheric zenith wet delay. However, only insufficient knowledge about turbulent refractivity variations is available. Thus, these parts can be adequately taken into account in the stochastic model. In the following section, we present the concept of the equivalent diagonal or EQUI model (Kermarrec and Schön 2016) as well as the diagonal correlation variance model that overcomes the weakness of the EQUI model. Through this contribution, we make use of the covariance function as shortly presented in "Appendix."

The equivalent diagonal model

Based on the work of Luati and Proietti (2011), the equivalence between diagonally weighted least-squares

(DWLS) and the generalized least-squares (GLS) estimator could be empirically applied to GNSS adjustments. One condition for the equivalence to hold states that each element of the diagonal matrix $\hat{\mathbf{W}}_{\mathrm{EQUI}}^{-1}$ is the sum of the row elements of the inverse of the fully populated cofactor matrix $\hat{\mathbf{W}}_{\mathrm{FULLY}}^{-1}$.

To further understand the concept behind the EQUI model, we adopt an AR(1) model to describe the correlation structure of GPS phase measurements. This model will be further used to derive the DCM and corresponds to a smoothness of ½ in the proposal of "Appendix." It is a crude model simplification, particularly when correlations are due to atmospheric effects (Kermarrec and Schön 2014, see also Luo et al. 2012 for ARIMA process). It is here introduced for the didactic purpose. Indeed, the inverse of the corresponding VCM can be exactly expressed thanks to the autocorrelation coefficient ρ (Rao and Toutenburg 1999). Assuming without loss of generality low variations of the satellite elevation, the elevation-dependent factor of the cofactor matrix can be factorized and the elements of the corresponding equivalent diagonal VCM for one satellite are:

- $$t = t_{\mathrm{first},i},\ t = t_{\mathrm{last},i} : \frac{1}{\sin(\mathrm{El}_i(t))^2}\left(\frac{1-\rho^2}{1-\rho}\right) = \frac{1}{\sin(\mathrm{El}_i(t))^2}(1+\rho)$$
- $$\text{otherwise} : \frac{1}{\sin(\mathrm{El}_i(t))^2}\left(\frac{1-\rho^2}{(1-\rho)^2}\right) = \frac{1}{\sin(\mathrm{El}_i(t))^2}\left(\frac{1+\rho}{1-\rho}\right) \quad (7)$$

We call $t_{\mathrm{first},i}$ and $t_{\mathrm{last},i}$ the first and last epoch at which the satellite i is present. The effect of correlations can be interpreted from (7) as giving a strong weight to the observations of the first and last epochs at which a satellite is present. A smaller weight, proportional to the ELEV variance model, is associated with all other values inside the batch. As for long batches of observations ($n \gg 2$ epochs), the impact of the extreme values on the results becomes negligible, Fig. 1 (top) presents only the central diagonal elements when the ρ is varied from 0 (no correlation) to 1 (full correlations). In this figure, we simulated, moreover, the addition of an identity noise matrix. We define $\hat{\mathbf{W}}_{\mathrm{EQUI,noise}} = \delta_{\mathrm{EQUI}}\hat{\mathbf{W}}_{\mathrm{EQUI}} + (1-\delta_{\mathrm{EQUI}})\mathbf{I}$ where δ_{EQUI} is varied from 0 ($\hat{\mathbf{W}}_{\mathrm{EQUI,noise}} = \mathbf{I}$ the identity matrix) to 1 ($\hat{\mathbf{W}}_{\mathrm{EQUI,noise}} = \hat{\mathbf{W}}_{\mathrm{EQUI}}$). The identity matrix acts both as VCM accounting for white noise added to the matrix of correlations and as a regularization matrix when the FULLY model is used. From Fig. 1 (top), we see that adding a noise matrix is similar to decreasing the effect of the equivalent matrix, i.e., as decreasing ρ (dotted green lines). Thus, different parameter combinations will lead nearly to the same equivalent variance, which we call the non-unicity of $[\rho, \delta_{\mathrm{EQUI}}]$.

A diagonal correlation model

The equivalence holds true for the estimates and their a priori cofactor matrices. Unfortunately, $\hat{\sigma}_{\hat{\mathbf{W}}\mathrm{EQUI}}$ is systematically underestimated (i.e., too small) since the equivalent variance is larger than 1 for the variance of satellites at 90° elevation (Fig. 1 top, blue lines) so that no overall model test can be applied to validate the least-squares solution (Teunissen 2000). In order to overcome this weakness, an easy-to-use alternative diagonal correlation model (DCM) is proposed. Based on this consideration, the proposal was chosen to have following properties:

- Approximate the behavior of the equivalent model, i.e., a downweighting of the variance model for all elevations
- Be based on a known and approved function to model the variance, following Euler and Goad (1991)
- Avoid the underestimation of $\hat{\sigma}_{\hat{\mathbf{W}}\mathrm{EQUI}}$. A simple empirical scaling to fit $\hat{\sigma}_{\hat{\mathbf{W}}\mathrm{FULLY}}$ is nearly impossible to obtain because of the infinity of cases to treat depending on the correlation structure
- Have a cofactor of 1 for satellite at 90° elevation to make comparisons possible

To summarize, this model should allow to take correlations into account in an "hidden way" in order to improve the reliability of the least-squares solution. From these criteria, an exponential function was selected for this contribution. The so-called DCM is thus defined as:

$$\sigma_{\mathrm{DCM}}^2(t) = \sigma_0^2\left((1-\delta_{\mathrm{DCM}}) + \delta_{\mathrm{DCM}}\frac{\exp^{-\mathrm{El}_i(t)/\gamma}}{\exp^{-90/\gamma}}\right) \quad (8)$$

We call the positive factor $\gamma\,(^\circ)$ the exponential factor. The elevation of the satellite i at time t is denoted El_i. The additional parameter δ_{DCM} models a scaled identity noise matrix as in the previous section. This proposal provides thus a great flexibility through the two parameters that can be varied to model different equivalent correlation structures. Although a comparison can be tempting, the DCM is *not* a usual variance model such as proposed by Euler and Goad (1991), Li et al. (2016) or Luo et al. (2014). It should be thought as a "diagonal correlation model" and is

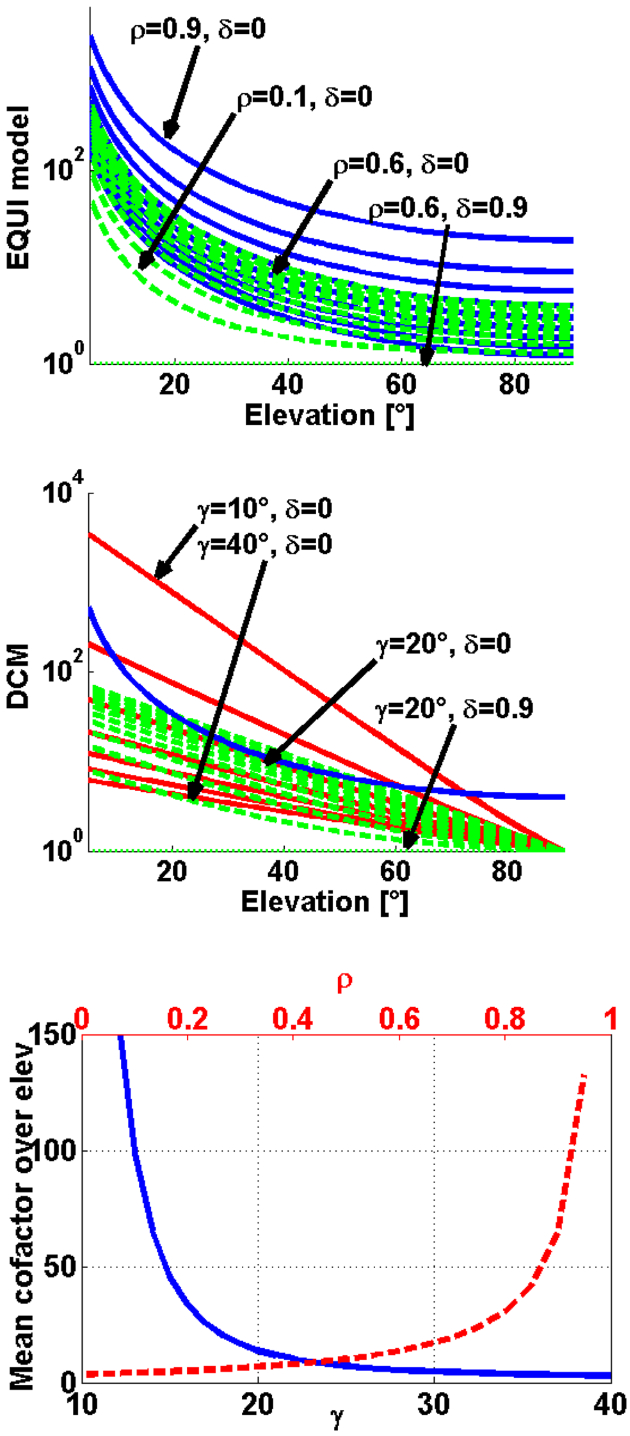


Fig. 1 Comparisons of the cofactor from the EQUI model and DCM for different set of parameters (no unit). *Top* the values of the equivalent diagonal model are plotted by varying ρ from 0 (*lowest blue line*) to 0.9 (*highest blue line*) by steps of 0.1 (logarithm scale). The *green dotted lines* are corresponding to the addition of an identity noise matrix for $\rho = 0.6$. *Middle* the DCM with γ varied by steps of 5° from 10° to 40° for $\delta_{\mathrm{DCM}} = 0$ (*red lines*). The *green dotted lines* correspond to $\gamma = 20°$ and different values of δ_{DCM}. The EQUI model with $\rho = 0.6$ corresponds to the *blue line*. *Bottom* the mean of the variance over all elevations is computed by varying γ and ρ with $\delta_{\mathrm{DCM}} = 0.3$

intentionally not called "variance model." A traditional fitting of the variance residuals will not give this structure.

In Fig. 1 (middle), the cofactor $\frac{\sigma^2_{\mathrm{DCM}}(t)}{\sigma_0^2}$ is plotted for different values of δ_{DCM} and γ where the elevation is varied from 5° to 90°. The blue line corresponds to the EQUI model for $[\rho, \delta_{\mathrm{EQUI}}] = [0.6, 0]$, the red lines to the DCM with $[\gamma, \delta_{\mathrm{DCM}}] = [10° - 40°, 0]$, and the green lines to $[\gamma, \delta_{\mathrm{DCM}}] = [20°, 0 - 0.9]$. The differences between the two models become here more visible as the DCM model is not based on a cosine function. Similar to Fig. 1 (top), an increase of δ_{DCM} corresponds approximately to a decrease of γ. It is thus possible to find nearly the same variance shape for different parameter sets, i.e., our "non-unicity." We note here additionally that the signature of the correlations inside a batch for a given satellite, i.e., beginning and end values differing from the one of the rest of the batch, gets lost with the DCM model, affecting possibly the estimates for very short batches.

Nonetheless, the DCM does not intend to fit exactly the equivalent correlation function but to approach it in mean over all elevations. Figure 1 (bottom) shows that it is always possible to find a $[\gamma, \delta_{\mathrm{DCM}}]$ corresponding to the AR(1) equivalent model defined by $[\rho, \delta_{\mathrm{DCM}} = \delta_{\mathrm{EQUI}}]$. The value $\rho = 0.85$ corresponds exemplarily to $\gamma = 15°$, whereas $\rho = 0.4$ corresponds to $\gamma = 22°$ with $\delta_{\mathrm{DCM}} = \delta_{\mathrm{EQUI}} = 0.3$. Thus, different correlation lengths can be modeled by varying the parameter γ. From (8), common exponential variance models can be derived by varying the parameters $[\gamma, \delta_{\mathrm{DCM}}]$ accordingly. Interested readers can verify that this corresponds to increasing both δ_{DCM} and γ for the proposal of Luo et al. (2014).

Matrix comparisons

To make our proposal more tangible for a GPS case, a visual comparison of the diagonal values of the VCM computed with different models is proposed. To that end, a representative constellation of 9 satellites corresponding to the skyplot drawn in Fig. 2 (bottom) is considered. The VCM are built using a batch approach, i.e., one submatrix is computed for each satellite observed during 100 epochs. In Fig. 2 (top), the DCM is compared with the EQUI and ELEV models. An AR(1) is assumed with $\rho = 0.85$ for the EQUI model and $\delta_{\mathrm{DCM}} = \delta_{\mathrm{EQUI}} = 0.3$. The similarities (see Fig. 1, bottom) between the DCM with $\gamma = 15°$ and the equivalent model are visually confirmed. Interestingly, the variances of high elevation satellites (PRN13, 28, 30) differ which was expected from the results of Fig. 1 (middle). As a consequence, a deeper study of the sensitivity of the least-squares results to the variables of the DCM seems necessary. It should assess the extent to which

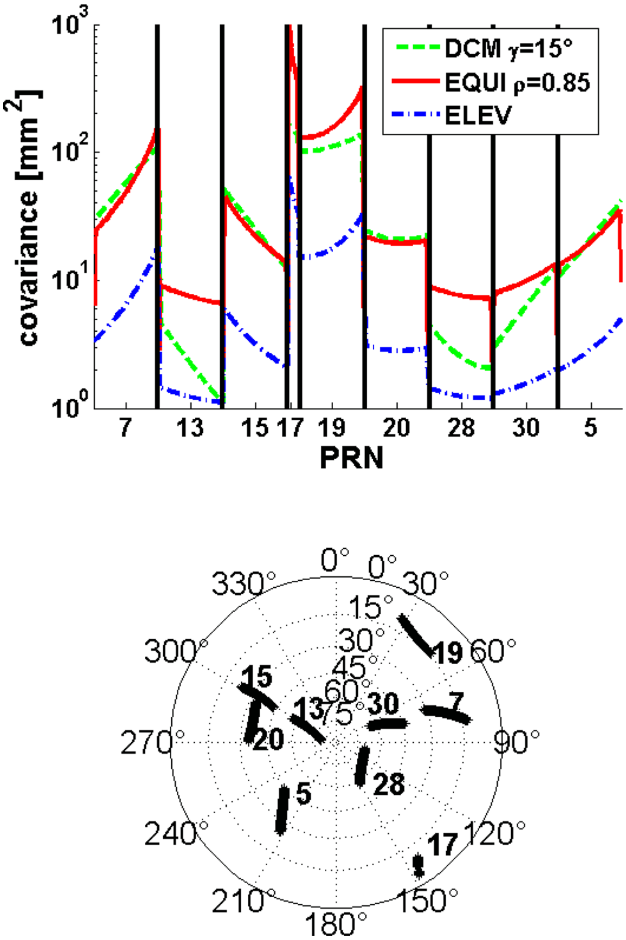


Fig. 2 *Top* comparison in log scale of the variances (ELEV) or equivalent variance (DCM, EQUI). One batch corresponds to 100 epochs, data rate 30 s. For the DCM (*green line*), $\gamma = 15°$ and for EQUI (*red line*) $\rho = 0.85$ (AR(1) model). The *blue line* corresponds to ELEV, $\delta = 0.3$. *Bottom* corresponding skyplot

a misspecification impacts the solution and validates the DCM as a model that accounts for correlations. Two cases were correspondingly identified: (1) the true VCM is corresponding to a DCM and (2) the true VCM is taking atmospheric-like correlations in consideration. We intentionally focus on modeling such correlations that are known to be the most important for GPS phase observations. An extension to other kind of correlations is proposed in the last section of this contribution. A relative positioning is up to now taken into consideration.

Sensitivity analysis of the DCM model

In a first step, simulations are carried out by varying the parameters set of the estimated DCM $\hat{\mathbf{W}}_{\mathrm{DCM}}$ around the reference values $[\gamma, \delta_{\mathrm{DCM}}]_0 = [20, 0.3]$ used to build $\mathbf{W}_{0\mathrm{DCM}}$. For the sake of simplicity, we skip the subscript noise. Time series with the reference structure are computed by means of Cholesky decomposition. A known integer vector is added to simulate ambiguities. The cutoff was taken to 5°, the sampling rate to 30 s. The satellite sky distribution of the two stations separated by 80 km corresponds approximately to Fig. 2 (bottom). We quantify the effects of misspecifications on quantities such as the loss of efficiency (4), the bias of $\hat{\sigma}_{\hat{\mathbf{W}}}$ (3). Monte Carlo simulations were carried out for the distance float-fix (6).

Results of simulations

The results of the sensitivity analysis on the quantities defined previously are summarized in Fig. 3. As expected, a minimum is reached when $\hat{\mathbf{W}}_{\mathrm{DCM}}$ corresponds to the structure of reference. However, we note that all presented quantities increase strongly when the γ is smaller than γ_0. Thus, underestimation of the parameter should be avoided, particularly $\gamma \leq 12°$. This is further associated with a wrong fixing of the ambiguity to integer due to the error in the float ambiguity vector. This error can be non-optimally corrected by increasing artificially δ_{DCM}, the loss of efficiency reaching 0.15 (Fig. 1 top, green line). Thus, as previously mentioned and similarly to the correlation structure, solutions that approximate the correct one are non-unique (Kermarrec and Schön 2017) but some of them are still preferable, particularly in terms of loss of efficiency and precision.

As γ increases, which corresponds to neglecting more and more correlations, R_{MSE} remains small with respect to the reference case, compared with the same decrease of γ. For $[\gamma, \delta_{DCM}] = [\gamma_0 + 20°, 0.3]$, we note a negligible increase of 0.01 for the loss of efficiency. However, for $[\gamma, \delta_{\mathrm{DCM}}] = [\gamma_0 - 10°, 0.3]$, the R_{MSE} is twice higher than with γ_0. For small variations around the reference, i.e., $\gamma = \gamma_0 {}^{+}/_{-} 5°$, a difference of maximum a few submm at the parameter level is expected from the MSE and d_1^*. This result holds also when $\delta_{\mathrm{DCM}} = 0$. However, if $\gamma = \gamma_0$ but $\delta_{\mathrm{DCM}} = 0.6$ is taken into consideration, a positive bias of $\hat{\sigma}_{\hat{\mathbf{W}}\mathrm{DCM}}$ by up to 2 mm with respect to the case $\delta_{\mathrm{DCM}} = 0$ is obtained, R_{MSE} remaining minimal. In real case, a rejection of the solution with the overall model test would be possible. Thus, neglecting correlations (Fig. 1 middle) by increasing both $[\gamma, \delta_{\mathrm{DCM}}]$ can be considered as less risky than overestimating the correlation length, i.e., decreasing $[\gamma, \delta_{\mathrm{DCM}}]$.

In a second step, the reference VCM is changed and assumed to be of type $\mathbf{W}_{0\mathrm{FULLY}}$ with $[\alpha, \nu]_0 = [0.01, 1]$. As mentioned, these values were chosen to model approximately the structure of atmospheric correlations (Kermarrec and Schön 2014). The smoothness of 1 can be shown to

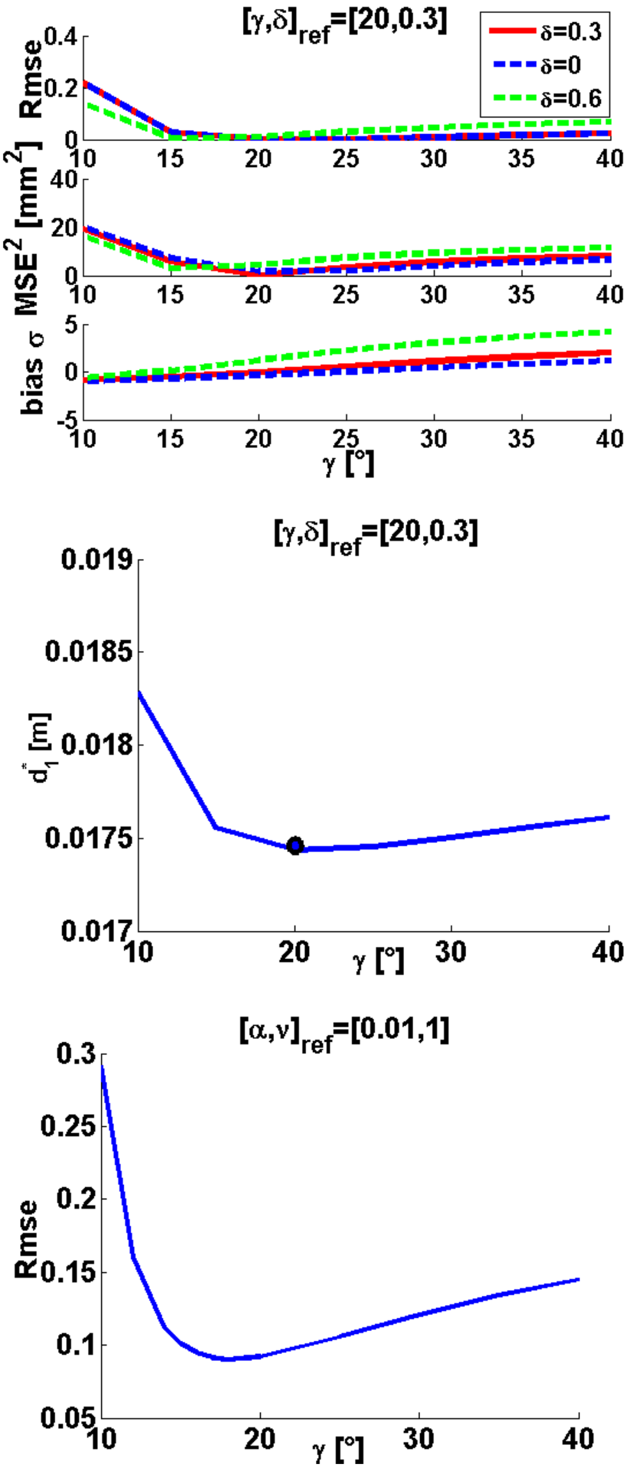


Fig. 3 Results from simulations. *Top* R_{MSE}, MSE and bias of $\hat{\sigma}_{\hat{\mathbf{W}}\mathrm{DCM}}$. The structure of the estimated DCM VCM is varied around the reference set $[\gamma, \delta_{\mathrm{DCM}}]_0 = [20, 0.3]$. *Middle* d_1^* is given in (m). *Bottom* the R_{MSE} by varying the γ for a reference matrix of FULLY type with $[\alpha, \nu]_0 = [0.01, 1]$, $\delta_{\mathrm{DCM}} = \delta_{\mathrm{FULLY}} = 0.3$

be physically relevant (Kermarrec et al. 2017), and the correlation length can be linked to $\rho = 0.85$ for an AR(1) model. As in this contribution we are neither willing to approximate exactly the equivalent model nor wanting to describe exactly the correlation structure, we allow ourselves to disregard the impact of the smoothness on the results by using an AR(1) structure. A noise matrix with $\delta_{\mathrm{FULLY}} = \delta_{\mathrm{DCM}} = 0.3$ is taken into consideration and the same methodology as previously is used. From the shape of the loss of efficiency (Fig. 3 bottom), we point out once more the risk of taking $\gamma < 12°$. It is essential to mention the clear minimum of R_{MSE} for $15° < \gamma < 20°$. This minimum is evidently not 0 as for the previous simulations where the reference matrix was corresponding to $\mathbf{W}_{0\mathrm{DCM}}$. However, its existence gives weight to the possibility of modeling elevation-dependent correlations of FULLY type thanks to the DCM. Moreover, making a parallel with the results of Fig. 1 (bottom), the range of values of γ for the minimum corresponds to the approximated correlated structure with an AR(1) model and $\rho = 0.85$. We note additionally that increasing γ and thus neglecting more and more correlations lead to a higher R_{MSE} than the minimum. This highlights the importance of taking correlations into account for a more reliable precision. We will come back to this result in the last section when making a proposal to fix the parameters of the DCM.

K matrix

In order to further compare the DCM with respect to the ELEV model, a short study of K (2) is proposed. The FULLY or EQUI cases are not depicted, i.e., the first and last values of the batches being much higher that the middle values. Figure 4 presents exemplarily the third column of K sorted per PRN (Up component). The behavior is similar to the two other columns (north and east components). The previous example, corresponding to the skyplot of Fig. 2 (bottom), is carried on.

For the value $\delta_{\mathrm{DCM}} = 0.3$ (Fig. 4 top), we notice that the slopes of **K** are getting similar for the DCM and the ELEV model as γ increases (i.e., $\gamma = 40°$) for satellites from 30° elevation (PRN 13, 15, 20 28, 30). This behavior validates our intuition, i.e., such high values of γ correspond to neglecting correlations. The non-unicity of the parameter combination is further highlighted in Fig. 4 (bottom) since the slopes for $[\gamma, \delta_{\mathrm{DCM}}] = [15, 0]$ and $[\gamma, \delta_{\mathrm{DCM}}] = [10, 0.3]$ have strong similarities (PRN13). Moreover, the **K** values for $\delta_{\mathrm{DCM}} = 0.6$ and $\delta_{\mathrm{DCM}} = 0.3$ are nearly identical. Thus, a certain variability is allowed, provided that the overall model test does not fail as mentioned previously (Fig. 3 top). Following the results of the previous simulations, we propose to fix a limit for γ to $\gamma_{\min} = 12°$ due to the hardly explainable slopes for $\gamma = 10°$ (PRN13). An increase of δ_{DCM} for $\gamma_{\min} < 12°$ in that particular case did not seem optimal in terms of loss of efficiency (Fig. 3 top).

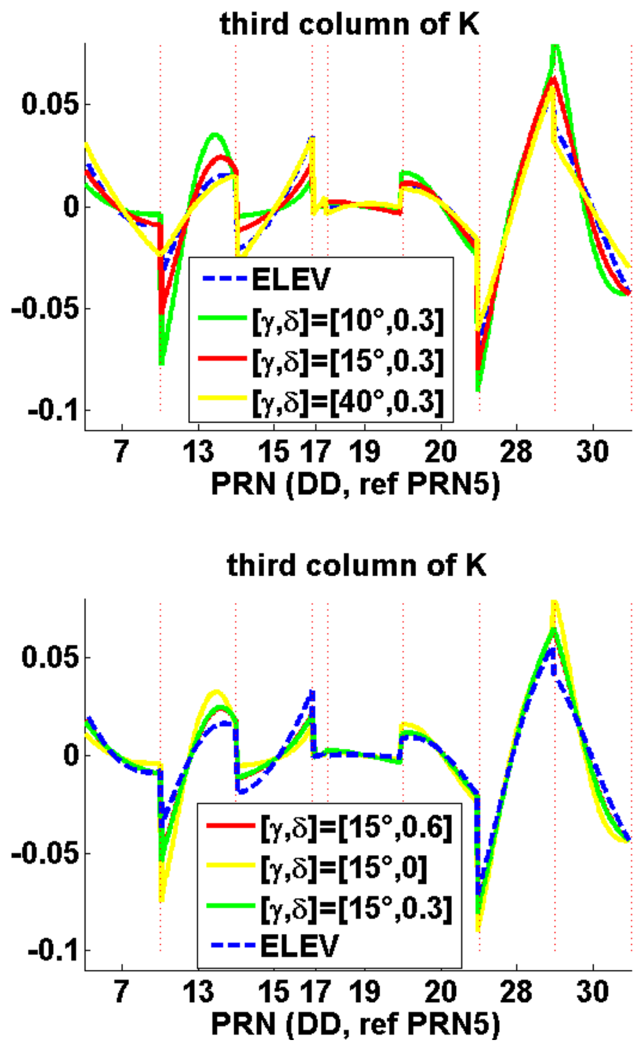


Fig. 4 Third column of the **K** matrix for different variance models (ELEV and DCM) is shown. *Top*: $\delta_{\mathrm{DCM}} = 0.3$ and γ is varied, *bottom* $\gamma = 15°$ for different values of δ_{DCM}. The batch length is 100 epochs with data rate 30 s. The label "PRN" corresponds to the PRN-reference satellite (PRN5) double differences

Conclusions of the simulations

The simulations have confirmed the similarities between the FULLY model and DCM for a given reference structure. We dissuaded using values of γ under 12° and pointed out the non-unicity of the best parameter combinations under this restriction. For the sake of stability, δ_{DCM} should not be taken to 0, a value of 0.3 being a good compromise. The importance of validating the solution by using for instance the overall model test was highlighted. As a conclusion from the analysis of **K**, we advise using the DCM when the geometry is enough averaged in order to compensate for a hardly explainable weighting from particular satellites. The similarities between the EQUI model [AR(1)] and DCM were indeed shown to hold in mean over all elevations (Fig. 1 bottom).

The conclusions of the simulations were derived for given reference structures but could be easily generalized for other sets, i.e., also for those corresponding to non-correlated cases by shifting the curves. The design matrix corresponds here to a realistic positioning case and was shown not to influence the previous conclusions on the DCM.

Case study

On the light of the previous simulations, a case study is proposed where the impact of the diagonal correlation model on quantities such as the 3DRMS and $\hat{\sigma}_{\hat{\mathbf{W}}}$ is studied. The factor γ is varied in a range from 10° to 40° for $\delta_{\mathrm{DCM}} = 0.3$. Although this value seems empirical, it was chosen based on the results of the simulations, making use of the non-unicity of the best set. Thus, for the sake of coherence and numerical stability for short batches, this plausible value was kept through the whole contribution.

Real case study

A relative positioning situation with a baseline of 80 km is chosen. L1GPS data having a 30-s data rate from two stations from the European Permanent Network (EPN, Bruyninx et al. 2012) KRAW and ZYWI were used (Kermarrec and Schön 2017). The corresponding double differences observations were shown to be correlated with previous contributions of the authors. The correlation structure could be mostly related to atmospheric effects ($[\alpha, \nu] \approx [0.01, 1]$). The reference values are the long-term station coordinates from EPN solution so that in mean a lower 3DRMS is corresponding to a more accurate and reliable positioning. The 3DRMS is chosen as a global indicator. In order to allow general conclusions on the impact of the stochastic model, the north, east, or up components are not given as they strongly depend on the underlying geometry and are thus less representative than the 3DRMS.

The GPS days DOY 201 of 2015 was chosen arbitrarily as well as the starting time GPS-SOD 6000 s. We aim in this section to validate under which conditions the FULLY or EQUI model can be replaced by the DCM for a given (and approximately known) correlation structure by studying the behavior of the least-squares results when γ is varied. Thus, other starting times would lead to similar shapes and conclusions due to the well averaged GPS constellation. No change in the reference satellite occurs during the session, and outliers were pre-eliminated. A cutoff of 5° was taken. The ambiguities were estimated as float together with the position. We used the ratio test with a threshold of 0.5 as validation test to fix the ambiguities to an integer using the Lambda method. Otherwise, the ambiguities were let float. The observations session is divided into batches of different lengths: 10 batches of 200 epochs, 20 batches of 100 epochs, 40 batches of 50 epochs, and 100 batches of 20 epochs, the same satellites being visible. Using this strategy, the impact of the less erroneous float or fix ambiguity vector can be assessed (Fig. 3 middle).

The mean values over the number of batches were computed. σ_0 was assessed to 3 mm with regard to the data quality and processing strategy. For batches shorter than 50 epochs, the residuals may not be normal distributed anymore. Thus, the overall model test may not correspond to the F-distribution (Williams et al. 2013) and was not applied for the 20-epoch case. The results given with the FULLY and ELEV models are presented. The EQUI model corresponds exactly to the FULLY case except for $\hat{\sigma}_{\hat{\mathrm{W}}\mathrm{EQUI}}$ is not included. Following the simulations, no tropospheric parameter was estimated. For batches shorter than 1 h, the risk of estimating an additional parameter is to be unable to separate the up component from the tropospheric parameter. For the sake of homogeneity, we decided not to estimate a tropospheric parameter for long batches. Nevertheless, estimating an additional parameter or not would not have changed the general conclusions of this contribution which are based on comparing results from different stochastic models by varying their parameters.

This assumption is from Fig. 1 (bottom), we expect a value of $\gamma = 15°$ to be optimal as soon as the geometry is enough averaged.

Results

From Fig. 5 (top left, 20 epochs case), a high variability of the 3DRMS by varying γ can be seen, which we interpret as the effect of a non-averaged geometry (Fig. 4). For very short batches (real-time kinematic like processing), the DCM model performs "between" the FULLY and ELEV models. As the ambiguities could not be fixed, the 3DRMS under the FULLY model outperforms by more than 6 cm the results from the DCM and ELEV model, the mean of $\hat{\sigma}_{\hat{\mathrm{W}}\mathrm{FULLY}}$ being 3.1 mm. The FULLY solution can be thus considered as the most correct one statistically.

As the batch length increases, the similarities between the FULLY model and the DCM for $\gamma = 15°$ become apparent for both the mean of $\hat{\sigma}_{\hat{\mathrm{W}}}$ and the 3DRMS. The corresponding results may not obligatory lead to the lowest attainable 3DRMS (case 50 and 200 epochs) which is reached for a lower γ of 10°. However, the slightly underestimated $\hat{\sigma}_{\hat{\mathrm{W}}\mathrm{DCM}}$ of less than 2.2 mm makes us take this optimistic result with care. Although the value stays in an accepted range, it corresponds, following the simulations, to a strong loss of efficiency (underestimated precision) and should be excluded. On the contrary, as γ increases (200-epochs case, Fig. 6 bottom right), the results given under the ELEV and the DCM models are getting similar, i.e., only a submm difference for an 80-km baseline is found. This has to be linked with the shape of the weightings and the linear combination of the observations (Figs. 1, 4) and is coherent with the results found in Luo et al. (2014).

The fixing of ambiguities was improved by approximately 5% for both the FULLY model and DCM with $\gamma = 15°$ for the 4 batch lengths. Increasing γ tends to decrease this percentage as the model is getting closer to the ELEV weighting. Obviously, the impact of the less erroneous float ambiguity vector is stronger for short batches and decreases with increasing batch lengths.

Test statistics and precision remain in all cases more reliable which can be seen by comparing the a priori precision Q given by the sum of the diagonal values of the a priori cofactor matrix of the coordinates for the different models. Exemplarily, and without loss of generality, the 50-epoch case was chosen. Whereas Q reaches 441 and 352 for the FULLY model and the DCM with $\gamma = 15°$, the a priori precision with the ELEV model is strongly underestimated ($Q = 64$) and corresponds to the values found for the DCM with $\gamma = 40°$. This short analysis confirms thus our previous conclusions.

Short baselines

For medium and long baselines, atmospheric effects are not canceled out by double differencing and correlate the observations. In such cases, the DCM with $\gamma = 15°$ is a good alternative to the FULLY or equivalently EQUI models. However, this model should be avoided for shorter batches and the FULLY or EQUI models preferred.

For short baselines, the impact of taking correlations into account is correspondingly smaller. On the one hand, the correlation length is shorter, i.e., atmospheric effects cancel out in most cases. On the other hand, the data quality and the ambiguity fixing are often better than for longer baselines. Thus, the results given under the models FULLY and by extension EQUI, DCM, and ELEV are nearly identical. However, less biased test statistics can still impact the positioning positively by excluding with a higher reliability batches with the overall model test, particularly when multipath is present. A decrease in the false alarm for more reliable outlier detection is also expected (Li et al. 2016). As a consequence, the DCM can still be used to model correlations for shorter baselines by adapting the parameters accordingly as proposed in the next section.

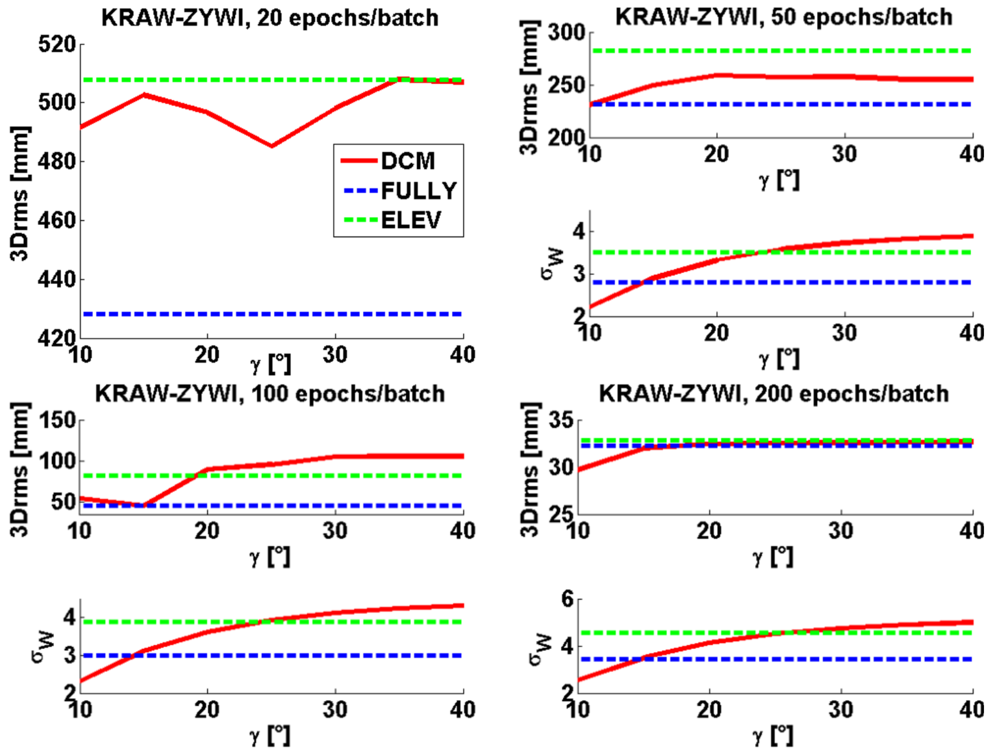


Fig. 5 Comparison of the impact of stochastic models on the mean of the 3DRMS (mm) and $\hat{\sigma}_{\hat{W}}$ (mm) by varying γ. The batch lengths are: *top left* 20 epochs, *top right* 50 epochs, *bottom left* 100 epochs, *bottom right* 200 epochs. *Red*, *blue*, and *green lines* correspond to the DCM, FULLY, and ELEV model, respectively

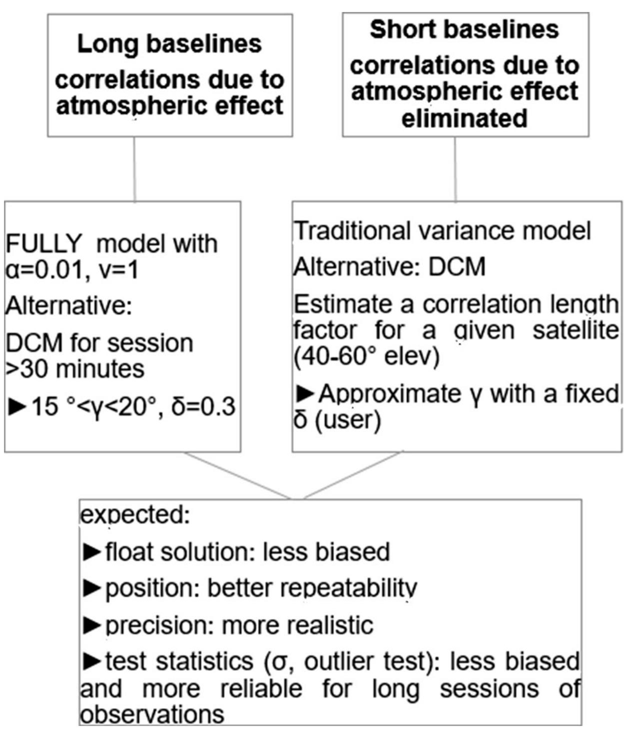


Fig. 6 How to use the DCM model: a summary

A proposal to fix the DCM parameters

Following the previous conclusions, Fig. 6 describes a simple strategy to fix the parameters of the DCM. The main reasons why and when this model should be preferred are summarized in the box called "expected." We here only shortly present the methodology.

In a first step, the correlation coefficient ρ of the double differences is estimated for a given satellite at mean elevation (40°–60°) using a simple MATLAB routine. A noise factor is either empirically fixed or estimated from the observations (Li et al. 2016). We propose for the sake of simplicity to use the empirical value of 0.3. Following Fig. 1 (bottom), the corresponding γ can be derived. Used for long baselines, this procedure may lead numerically to $\rho > 0.95$ due to the AR(1) approximation and corresponds to $\gamma < \gamma_{\min}$. As a consequence, and because correlations in that case are mainly due to atmospheric effects, we advise to fix it to 0.85. It should be clear that an accurate description of correlations should depend on time and location. Interested readers are advised to consult results in Kermarrec and Schön (2017) of a sensitivity analysis of the parameters of the correlation function presented in "Appendix". This discussion is, however, beyond the scope of the paper, a stochastic model being in essence always an approximation of the unknown truth.

Independent of the baseline length, the DCM should be used with care for short batches when the geometry is not enough averaged. However, considered as a hidden correlation model based on an exponential function, it can replace the commonly used variance models in all other cases.

Conclusions

In this contribution, a proposal to approximate the equivalent diagonal model EQUI that summarizes fully populated VCM built thanks to the model proposed by Kermarrec and Schön (2017) into diagonal matrices for GNSS positioning was presented. This model called diagonal correlation model (DCM) is based on a simple exponential function where the modeling of an additional white noise is let to a scaled identity VCM. It is thought and designed as an alternative to the EQUI model. As no inversion of a fully populated matrix is needed, one of the main weakness of the EQUI model is overcome. Moreover, the diagonal value reaches 1 for satellites at 90° of elevation so that no underestimation of the a posteriori variance factor occurs, making comparison obtained with traditional variance models possible.

Simulations allowed to assess the sensibility of some chosen least-squares quantities to variations of the parameters of the DCM VCM. Whereas an underestimation of the exponential factor is riskier in terms of loss of efficiency than an overestimation, the noise factor impacts less strongly the results. A link in mean over all elevations between the DCM parameters and the correlation structure was further highlighted. A real case study using observations from a medium baseline where unmodeled or atmospheric effects were not eliminated validates the conclusions of the simulations. For batches shorter than 50 epochs, the 3DRMS with the DCM was sensitive to variations of the exponential factor due to a reduced averaging of the geometry and challenging data quality. However, from 50 epochs, the a posteriori variance factor under the DCM model with $\gamma = 15^\circ$ was similar to the FULLY model in terms of 3DRMS, the differences getting smaller as the batch lengths increased. Thus, more reliable test statistics, improved ambiguity resolution and precision are in all cases expected from the DCM, FULLY or EQUI models with respect to the ELEV model, independently of the baseline length or processing strategy. The corresponding VCM remains, however, diagonal and does not replace FULLY matrices which can whiten the observations and should be preferred for short sessions.

Acknowledgements The authors gratefully acknowledge the EPN and corresponding agencies for providing the data freely. Three reviewers are warmly acknowledged for their constructive comments.

Appendix: a model for GNSS phase correlations

The FULLY model is shortly presented (Kermarrec and Schön 2017). The covariance function is derived from a model for phase covariance due to turbulent tropospheric refractivity variations (Schön and Brunner 2008). The covariance function C between 2 observations of satellites i and j at the time t and $t+\tau$ reads:

$$C_{it}^{jt+\tau} = \frac{\beta\delta}{\sin(\mathrm{El}_i(t))\sin\left(\mathrm{El}_j(t+\tau)\right)}(\alpha\tau)^{\nu}K_{\nu}(\alpha\tau)$$

El_i and El_j are the elevations of the satellite i and j, respectively. δ is a scaling parameter so that the variance equals 1 for satellites at 90° elevation using small argument approximations (Schön and Brunner 2008). K_ν is the modified Bessel function of order ν. β is a weighting factor modeling the covariance between different satellites. ν is called the smoothness; α describes the decay of the covariance function with time.

References

Amiri-Simkooei AR, Jazaeri S, Zangeneh-Nejad F, Asgari J (2016) Role of stochastic model on GPS integer ambiguity resolution success rate. GPS Solut 20(1):51–61

Brunner FK, Hartinger H, Troyer L (1999) GPS signal diffraction modelling: the stochastic SIGMA-delta model. J Geod 73(5):259–267

Bruyninx C, Habrich H, Söhne W, Kenyeres A, Stangl G, Völksen C (2015) Enhancement of the EUREF permanent network services and products. In: Kenyon S, Pacino M, Marti U (eds) Geodesy for planet earth. IAG symposia series, vol 136. Springer, Berlin, pp 27–35

Collins JP, Langley RB (1999) Possible weighting schemes for GPS carrier phase observations in the presence of multipath. U. S. Army Corps Eng Topogr Eng Cent 1999:1–40

Dach R, Hugentobler U, Fridez P, Meindl M (eds) (2007) Bernese GPS Software Version 5.0. Astronomical Institute, University of Bern, Switzerland

El-Rabbany AE-S, Kleusberg A (2003) The effect of physical correlations on the ambiguity resolution and accuracy estimation in GPS differential positioning. J Surv Eng 129(1):28

Euler HJ, Goad CC (1991) On optimal filtering of GPS dual frequency observations without using orbit information. Bull Geod 65(2):130–143

Euler HJ, Schaffrin B (1991) On a measure for the discernibility between different ambiguity solutions in the static kinematic GPS mode. In: Schwarz KP, Lachapelle G (eds) Kinematic systems in geodesy. IAG symposia, vol 107. Springer, Berlin, pp 285–295

Gerdan GP (1995) A comparison of four methods of weighting double difference pseudo range measurements. Aust Surv 40(4):60–66

Han S (1997) Quality-control issues relating to instantaneous ambiguity resolution for real-time GPS kinematic positioning. J Geod 71(6):351–361

Hartinger H, Brunner FK (1999) Variances of GPS phase observations: the SIGMA-model. GPS Solut 2(4):35–43

Howind J, Kutterer H, Heck B (1999) Impact of temporal correlations on GPS-derived relative point positions. J Geod 73(5):246–258

Kermarrec G, Schön S (2014) On the Mátern covariance family: a proposal for modelling temporal correlations based on turbulence theory. J Geod 88(11):1061–1079

Kermarrec G, Schön S (2016) Taking correlations in GPS least-squares adjustments into account with a diagonal covariance matrix. J Geod 90(9):793–805

Kermarrec G, Schön S (2017) A priori fully populated covariance matrices in least-squares adjustment—case study: GPS relative positioning. J Geod 91(5):465–484
Kermarrec G, Schön S, Kreinovich V (2017) Possible explanation of empirical values of the matern smoothness parameter for the temporal covariance of GPS measurements. Appl Math Sci 11(35):1733–1737
Koch KR (1999) Parameter estimation and hypothesis testing in linear models. Springer, Berlin
Leick A, Rapoport L, Tatarnikov D (2015) GPS satellite surveying, 4th edn. Wiley, New York
Li B, Lou L, Shen Y (2016) GNSS elevation-dependent stochastic modelling and its impacts on the statistic testing. J Surv Eng 142(2):04015012
Luati A, Proietti T (2011) On the equivalence of the weighted least-squares and the generalised least-squares estimators, with applications to kernel smoothing. Ann Inst Stat Math 63(4):673–686
Luo X, Mayer M, Heck B (2012) Analysing time series of GNSS residuals by means of AR(I)MA processes. In: Sneeuw N, Novák P, Crespi M, Sansò F (eds) VII Hotine-Marussi Symposium on Mathematical Geodesy. International Association of Geodesy Symposia, vol 137. Springer, Berlin, pp 129–134
Luo X, Mayer M, Heck B, Awange JL (2014) A realistic and easy-to-implement weighting model for GNSS phase observations. IEEE Trans Geosci Remote Sens 52(10):6110–6118
Rao C, Toutenburg H (1999) Linear models, least-squares and alternatives, 2nd edn. Springer, New York
Schön S, Brunner FK (2008) A proposal for modelling physical correlations of GPS phase observations. J Geod 82(10):601–612
Stein ML (1999) Interpolation of spatial data. Some theory for kriging. Springer, New York
Strand ON (1974) Coefficient errors caused by using the wrong covariance matrix in the general linear model. Ann Stat 2(5):935–949
Talbot N (1988). Optimal weighting of GPS carrier phase observations based on the signal-to-noise ratio. In: Proceedings of international symposia, global positioning systems. Gold Coast, Queensland, pp 4.1–4.17, 17–19 Oct
Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. J Geod 70:65–82
Teunissen PJG (2000) Testing theory and introduction. Series on mathematical geodesy and positioning. Delft University Press, Delft
Teunissen PJG, Amiri-Simkooei AR (2008) Least-squares variance component estimation. J Geod 2008(82):65–82
Teunissen PJG, Kleusberg A (1998) GPS for Geodesy, 2nd edn. Springer-Verlag, Berlin
Vermeer M (1997) The precision of geodetic GPS and one way of improving it. J Geod 71(4):240–245
Wang J, Stewart P, Tsakiri M (2000) A comparative study of integer ambiguity validation procedures. Earth Planet Space 52:813–817
Wang J, Satirapod C, Rizos C (2002) Stochastic assessment of GPS carrier phase measurements for precise static relative positioning. J Geod 76(2):95–104
Wei M, Schwarz KP (1995) Fast ambiguity resolution using an integer nonlinear programming method. In: ION GPS-1995, Institute of Navigation, Palm Spring, CA, USA, pp 1101–1110, Sep 1995
Wieser A, Brunner FK (2000) An extended weight model for GPS phase observations. Earth Planet Space 52:777–782
Williams MN, Gomez Grajales CA, Kurkiewicz D (2013) Assumptions of multiple regression: correcting two missconceptions. Pract Assess Res Eval 18(11):2
Xu P (2013) The effect of incorrect weights on estimating the variance of unit weight. Stud Geophys Geod 57:339–352

Gaël Kermarrec is currently an external Ph.D. candidate at the Geodetic Institute, Karlsruhe University, Germany. She obtained her M. Sc degrees at the Ecole des Sciences Geographiques de l'Etat (ENSG) in Paris in 2000. Her current research focuses on modeling correlations of GNSS observations.

Prof. Dr. Steffen Schön is university professor of satellite navigation at the Leibniz University in Hanover.

Research Article Open Access

G. Kermarrec* and S. Schön

Fully populated VCM or the hidden parameter

Received June 28, 2017; accepted November 2, 2017

Abstract: Least-squares estimates are trustworthy with minimal variance if the correct stochastic model is used. Due to computational burden, diagonal models that neglect correlations are preferred to describe the elevation dependency of the variance of GPS observations. In this contribution, an improved stochastic model based on a parametric function to take correlations between GPS phase observations into account is presented. Built on an adapted and flexible Mátern function accounting for spatiotemporal variabilities, its parameters can be fixed thanks to Maximum Likelihood Estimation or chosen apriori to model turbulent tropospheric refractivity fluctuations. In this contribution, we will show in which cases and under which conditions corresponding fully populated variance covariance matrices (VCM) replace the estimation of a tropospheric parameter. For this equivalence "augmented functional versus augmented stochastic model" to hold, the VCM should be made sufficiently large which corresponds to computing small batches of observations. A case study with observations from a medium baseline of 80 km divided into batches of 600 shows improvement of up to 100 mm for the 3Drms when fully populated VCM are used compared with an elevation dependent diagonal model. It confirms the strong potential of such matrices to improve the least-squares solution, particularly when ambiguities are let float.

Keywords: correlations, equivalence stochastic functional model, GNSS phase measurement, hidden tropospheric parameter, least-squares, Mátern covariance function, stochastic model

1 Introduction

Because of an overdetermined system of equations with more observations than unknowns, GNSS measurements are often processed with least-squares estimation methods. The functional model which describes the relationship between the observations and the parameters to be estimated is well-known (Hofmann-Wellenhof et al. 2001); the same cannot be said for the stochastic model. However, the correct modelling of non-deterministic effects can be considered as a prerequisite in order to reach a minimum variance of the estimates. Heteroscedasticity of GPS residuals (Bischoff et al. 2005) is widely assumed and the elevation dependency of the variance of GNSS observations is described thanks to cosine, exponential or CNO/SNR based functions, see exemplarily Vermeer (1997), Wang et al. (1998) or Luo et al. (2014). Even if many factors act on correlating the observations, such as the atmosphere (Schön and Brunner 2008), or the receiver itself (Bona 2000, Amiri-Simkooei and Tiberius 2007), correlations remain mostly disregarded. Besides computational demanding iterative procedures on the residuals (Koch 1999, Teunissen and Amiri-Simkooei 2008), empirical models for correlations between GNSS measurements have been concretely used (El-Rabbany 1994, Howind et al. 1999). However, due to a lack of an accurate and plausible description, correlations are often neglected. Additionally, diagonal variance covariance matrices (VCM) are less difficult to handle than fully populated VCM accounting for correlations. Since the least-squares solution remains unbiased even with approximated stochastic models as long as the residuals are zero-mean, no main differences are expected at the estimates level in ideal cases. This was confirmed for example by Radovanovic (2001). However, when correlations are neglected, the least-squares estimator is less efficient and significance tests biased (Williams et al. 2003). The consequences are for example an overoptimistic precision, a worthier ambiguity resolution or outlier detection (Kermarrec and Schön 2017, Amiri-Simkooei et al. 2016, Li et al. 2017 for Beidou). The development of a better and realistic stochastic model is a way to face this issue (Tralli and Lichten 1990).

Based on a Mátern covariance function and physical considerations, Kermarrec and Schön (2017) proposed a new approach to describe elevation dependent correlations in an understandable manner. This function has two main parameters: the smoothness and a correlation parameter and thus allows a greater flexibility with respect to simpler non elevation dependent functions, such as the

***Corresponding Author: G. Kermarrec:** Leibniz University Hannover, Germany, E-mail: gael.kermarrec@web.de
S. Schön: Leibniz University Hannover, Germany

first order Gauss Markov model (AR(1)) proposed by El-Rabbany (1994). To model atmospheric effects, the parameters can be fixed to given values following Kermarrec and Schön (2014).

In this contribution, we mathematically derive how integrating fully populated VCM built with this function in the least-squares adjustment can impact the least-squares solution. Indeed, in some particular cases, not only the test statistics become more trustworthy and less biased under such an improved stochastic model but also the estimates themselves can be impacted. Thus, the structure of the paper is as follows: in the first part, we will describe shortly the proposed correlation model. The second part explains the concept of the "hidden elevation dependent parameter" to present when such a function can replace a non-estimable tropospheric parameter. In the third part and thanks to an example, we will more concretely highlight the impact on the solution of non-diagonal VCM build with the proposal. The appendix deals with the problem of precision and ambiguity resolution.

2 Stochastic model: a proposal for correlations

2.1 Mathematical background

The point positioning problem is usually solved by first linearizing the observation equations w.r.t. the unknown parameters. Based on approximate parameter values, the so-called linearized functional model is obtained that describes the mathematical relationship between the estimates and the observations. After rearranging, the Observed Minus Computed (OMC) term can be computed which is the difference between actual observations and modelled observations. The corresponding equation is:

$$\mathbf{y} = \mathbf{A}\Delta\mathbf{x} + \varepsilon \tag{1}$$

In this contribution, we assume a relative positioning scenario with GNSS phase observations. We call $\mathbf{y}$ the $n \times 1$ vector of Observed-Minus-Computed (OMC) double differences, ε the $n \times 1$ error vector. We assume that the error term has zero mean and a normal distribution, $E\left(\varepsilon\varepsilon^{\mathbf{T}}\right) = \sigma^2\mathbf{W}_0$ where $\mathbf{W}_0$ is the positive definite and fully populated cofactor matrix of the double differences and E the mathematical expectation. σ^2 is the apriori variance factor. Dealing with phase measurements is inherently ambiguous, the ambiguities are estimated in a first step as float, i.e. part of the functional model. The design matrix for GNSS positioning can be thus partitioned as $\mathbf{A} = [\ \mathbf{A_C} \quad \mathbf{A_A}\ ]$. The $(n, 3)$ matrix $\mathbf{A_C}$ and the (n, n_{amb}) matrix $\mathbf{A_A}$ describe the design matrices of the coordinates and ambiguities, respectively, where n and n_{amb} are the number of double differences and the number of double differenced ambiguities, respectively. If a tropospheric parameter has to be estimated, the design matrix is extended accordingly, as described for example in Kermarrec and Schön (2016). Similarly to the design matrix, the correction vector for the unknown parameters $\Delta\mathbf{x} = [\ \Delta\mathbf{x_C} \quad \mathbf{x_A}\ ]$ is divided into a correction on the estimated coordinates and the float ambiguity.

The Generalized Least Squares Estimator (GLSE) reads $\Delta\hat{\mathbf{x}}_0 = \left(\mathbf{A^T}\mathbf{W}_0{}^{-1}\mathbf{A}\right)^{-1}\mathbf{A^T}\mathbf{W}_0{}^{-1}\mathbf{y}$ (Koch 1999). In practise, $\mathbf{W}_0$ is unknown and replaced by an assumption or apriori variance covariance matrix (VCM) which we call $\hat{\mathbf{W}}$. As a consequence, the feasible generalized least-squares solution (FGLSE) is given by:

$$\Delta\hat{\mathbf{x}} = \left(\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{y} \tag{2}$$

The apriori cofactor matrix of the estimated parameter $\Delta\hat{\mathbf{x}}$ is $\mathbf{Q}_{\hat{\mathbf{x}}} = \left(\mathbf{A^T}\hat{\mathbf{W}}^{-1}\mathbf{A}\right)^{-1}$, partitioned as follows into an ambiguity and coordinates part:

$$\mathbf{Q}_{\hat{\mathbf{x}}} = \begin{bmatrix} \mathbf{Q_C} & \mathbf{Q_{CA}} \\ \mathbf{Q_{CA}} & \mathbf{Q_A} \end{bmatrix} \tag{3}$$

Calling $\mathbf{v}$ the vector of residuals and $n - u$ the degree of freedom, the aposteriori variance factor for the FGLSE is given by

$$\hat{\sigma}^2_{\mathbf{W}} = \frac{\left(\mathbf{y} - \mathbf{A}\Delta\hat{\mathbf{x}}\right)^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\left(\mathbf{y} - \mathbf{A}\Delta\hat{\mathbf{x}}\right)}{n-u} = \frac{\mathbf{v^T}\hat{\mathbf{W}}^{-1}\mathbf{v}}{n-u} \tag{4}$$

The least-squares estimator is unbiased, consistent and efficient if the least-squares assumptions are not violated, particularly if the residuals are 0-mean and the correct stochastic model is used (Williams et al. 2003). In case of GNSS positioning, heteroscedasticity should be taken into account in the modelling as well as correlations between measurements, when needed. It is thus of central importance for a trustworthy positioning to avoid misspecifications of the stochastic model and describe the temporal relationship between observations.

Fixing the ambiguities to integer

For a high accuracy of the solution, the float ambiguity vector should be fixed to integer. Various strategies can be used from a simple rounding to more complicated methods such as the FARA (Erickson 1992) or the Lambda method (Teunissen 1995). To prevent from a wrong fixing to integer, the fixed ambiguity vector has to be validated. This can be done for example thanks to discriminant tests such as the ratio test (Verhagen and Teunissen 2013). Eventually, a Fixed Failure-rate Ratio Test (Wang and Feng 2013) or look-up tables (Teunissen and Verhagen 2009) can be used. When not otherwise mentioned, we made use of the Lambda method to fix the ambiguity and use a simple ratio test with a threshold of 0.5 (Wei and Schwarz 1995). As will be shown, the results of this contribution are not impacted by the fixing or validation method. For the sake of completeness however, a short analysis of the impact of correlations on the ratio test is proposed in the appendix.

2.1.1 A proposal to model temporal correlations

An adapted version of the model developed by Kermarrec and Schön (2014) is chosen to describe temporal elevation dependent correlations of GNSS phase measurements. The reader is referred to Kermarrec and Schön (2017) for more details on the choice of this function as well as a comparison with existing strategies such as the model from El-Rabbany (1994) or empirical ARMA processes (Luo et al. 2012).

The covariance C between two observations of satellites with PRN i and j at time t and $t+\tau$ reads:

$$C_{it}^{jt+\tau} = \frac{\rho_{weight}\delta}{\sin\left(El_i\left(t\right)\right)\sin\left(El_j\left(t+\tau\right)\right)}\left(\alpha\tau\right)^{\nu}K_{\nu}\left(\alpha\tau\right) \quad (5)$$

El_i and El_j are the elevations of the satellite with PRN i and j respectively, ρ_{weight} is a weighting factor modelling the covariance between different satellites. In this contribution, we choose to fix $\rho_{weight} = 1$ for correlations between observations from one satellite with itself (i.e. *i=j*) and disregard correlations between different satellites, i.e. $\rho_{weight} = 0$ if $i \neq j$. Although our model allows to account for cross-correlations based on physical considerations (Kermarrec and Schön 2014), it is unnecessary to account for them for the following derivation about the augmented stochastic model. δ is a scaling parameter so that the variance equals 1 for satellites at 90° elevation. α is called the correlation parameter $\left[s^{-1}\right]$ and ν the smoothness. The modified Bessel function of order ν (Abramowitz and Stegun 1972) is denoted by K_ν. Through this contribution, we will refer to the set $[\alpha,\nu]$ as the "Mátern parameters set". This covariance function is derived from a rational spectral density function (Kermarrec and Schön 2014) and thus the corresponding VCM $\hat{\mathbf{W}}_{\mathbf{UD},corr}$ of undifferenced phase observations are positive definite (Mátern 1960).

The spectral density of the Mátern covariance function is given by:

$$S\left(\omega\right) = \frac{2^{\nu-1}\gamma\Gamma\left(\nu+d/2\right)\alpha^{2\nu}}{\pi^{d/2}\left(\omega^2+\alpha^2\right)^{\nu+d/2}} \quad (6)$$

where $\omega^2 = \omega_1^2+\omega_2^2+\ldots+\omega_d^2$ is the angular frequency, Γ the Gamma function (Abramowitz and Segun 1972). The dimension of the field d is 1 in case of time series of observations. From Eq. (6), it can be seen that the behaviour of $S\left(\omega\right)$by letting $\omega \to 0$ is both influenced by the smoothness ν and the correlation parameters α, whereas ν plays a more important role in filtering high frequencies (i.e. as $\omega \to \infty$).

Since the Mátern covariance function in Eq. (5) is weighted by an elevation dependent factor, the covariance is different for each satellite or satellite pairs. The Mátern parameters can be computed by Maximum Likelihood Estimation (Kermarrec and Schön 2017) and are thus depending on the observations (L1, L2, eventually P1 or P2) or alternatively fixed. The value $\nu = 1/2$ corresponds for instance to a first order Gauss Markov process, i.e. an exponential function as proposed by El-Rabbany (1994). The values of $[\alpha,\nu] = [0.008, 5/6]$ following Kermarrec and Schön (2014) and Wheelon (2001) were shown to model tropospheric correlations due to the turbulent variations of the refractivity index.

Through this contribution, we will make use of the set $[\alpha,\nu] = [0.01, 1.05]$ to model elevation dependent correlations due to atmospheric effects. The reasons of this particular choice are shortly highlighted:

- The mean-square differentiability of the field is ensured (Stein 1999) which is for physical reasons an interesting property of the covariance function. Indeed, seeing a GPS unit as a combination of resistors, capacitors and inductors, the differentiability of the current intensity on time and so the measured quantity has to be given. The voltage of the inductor is for instance proportional to the time derivative of the current which may thus be finite (Kermarrec et al. 2017).
- By taking a slightly higher smoothness than $5/6$ (i.e. the "tropospheric" value), the correlation parameter α has to be reduced making use of the non-orthogonality property of the Mátern covariance function (Gelfand et al. 2011). This result was confirmed by Kermarrec and Schön (2017).

Both for the sake of numerical stability when inverting fully populated matrices and for modeling additional white noise, the undifferenced VCM $\hat{\mathbf{W}}_{\mathbf{UD},fully}$ are built as a linear combination of $\hat{\mathbf{W}}_{\mathbf{UD},corr}$ and the identity matrix $\mathbf{I}$ modelling white noise as follows:

$$\hat{\mathbf{W}}_{\mathbf{UD},fully} = (1-\beta)\hat{\mathbf{W}}_{\mathbf{UD},corr} - \beta\mathbf{I} \quad (7)$$

β is a positive noise factor between 0 and 1 which can be estimated from the OMC or fixed apriori. This proposal corresponds to an elementary model as proposed by

Schwieger (2007). Undifferenced matrices $\hat{\mathbf{W}}_{\mathbf{UD},fully}$ can be built for each satellite with Eq. (7) for a chosen number of epochs η_{epoch} where the satellite is visible, i.e. the batch length. The whole matrix is referred to as the FULLY VCM. The corresponding diagonal VCM $\hat{\mathbf{W}}_{\mathbf{UD},elev}$ where only heteroscedasticity is taken into account is called the ELEV model. Its elements are corresponding to the diagonal of $\hat{\mathbf{W}}_{\mathbf{UD},corr}$, i.e. the commonly used cosine model. Subsequently, the cofactor matrix for a relative positioning scenario with double differences reads $\hat{\mathbf{W}} = \mathbf{M}^{\mathbf{T}}\hat{\mathbf{W}}_{\mathbf{UD}}\mathbf{M}$, where $\mathbf{M}$ is the matrix operator of double differencing.

3 The hidden parameter

Dealing with OMC, we assume that the ionosphere and the troposphere are firstly modelled with enough accuracy in the pre-processing step (Hoffmann and Wellenhof 1999). In some cases, e.g. for medium-long baselines from approximately 20 km length, tropospheric effects do not cancel out by double differencing. Thus a differential tropospheric parameter is estimated as part of the functional model. Due to its small variations between epochs, one value is computed per satellite for session from 1 hour of observations, i.e. one batch of observations. As the temporal resolution is restricted as a consequence of the lack of separability between parameters, particularly with the Up component, usually no additional parameter is estimated for shorter sessions. Unfortunately, the effect of the troposphere still impacts the coordinates for sessions shorter than one hour.

In this section, we will show how a fully populated VCM computed with Eq. (7) can replace the estimation of a tropospheric parameter for sessions shorter than one hour in an elegant way. The mathematical derivation proposed by Blewitt (1998) is presented and extended to the particular case of GPS.

3.1 Augmented functional model versus augmented stochastic model

Augmented functional model

In order to improve the solution of Eq. (1), an additional parameter $\Delta\mathbf{z}$ can be taken into account. For the GPS case, we can consider $\Delta\mathbf{z}$ to be a differential tropospheric parameter. In that case the augmented model reads:

$$\mathbf{y} = \mathbf{A}\Delta\mathbf{x} + \mathbf{B}\Delta\mathbf{z} + \boldsymbol{\epsilon} \tag{8}$$

$\mathbf{B}$ is the design matrix with dimension $\eta_{sat} \times \eta_{epoch}$ corresponding to $\Delta\mathbf{z}$ where η_{sat} is the number of visible satellites. If Eq. (8) is written in terms of partitioned matrices, it can be shown by applying the lemma on matrix inversions for symmetric matrices that the solution $\Delta\hat{\mathbf{x}}$ is given by $\Delta\hat{\mathbf{x}} = (\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{P}\mathbf{A})^{-1}\hat{\mathbf{W}}^{-1}\mathbf{P}\mathbf{y}$ with $\mathbf{P} = \mathbf{I} - \mathbf{B}(\mathbf{B}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{B})^{-1}\mathbf{B}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}$ being a projection operator.

We can thus define a reduced weight matrix as

$$\hat{\mathbf{W}}_{\mathbf{red}}^{-1} = \hat{\mathbf{W}}^{-1}\mathbf{P} = \hat{\mathbf{W}}^{-1} - \hat{\mathbf{W}}^{-1}\mathbf{B}(\mathbf{B}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{B})^{-1}\mathbf{B}^{\mathbf{T}}\hat{\mathbf{W}}^{-1} \tag{9}$$

If the estimates are expressed as $\Delta\hat{\mathbf{x}} = (\mathbf{A}^{\mathbf{T}}\hat{\mathbf{W}}_{\mathbf{red}}^{-1}\mathbf{A})^{-1}\hat{\mathbf{W}}_{\mathbf{red}}^{-1}\mathbf{y}$, a parallel with Eq. (2) can be drawn. With the knowledge of $\hat{\mathbf{W}}_{\mathbf{red}}$, it is thus possible to compute $\Delta\hat{\mathbf{x}}$ without having to compute $\Delta\hat{\mathbf{z}}$. This is exactly what we aim to achieve in the GPS case for short sessions, due to the lack of separability between the tropospheric and Up parameters. Unfortunately, the reduced weight matrix $\hat{\mathbf{W}}_{\mathbf{red}}^{-1}$ is singular. As a consequence, assessing the stochastic model which would lead to such a VCM and allows for the direct computation of $\hat{\mathbf{W}}_{\mathbf{red}}$ is impossible.

Augmented stochastic model

This difficulty can be overcome by seeing the augmented parameter $\Delta\mathbf{z}$ as a source of noise, i.e. a "process noise", similarly to what is done in Kalman filtering. Concretely, we define $\boldsymbol{\epsilon}_{red}$ as an augmented noise, i.e. $\boldsymbol{\epsilon}_{red} = \mathbf{B}\Delta\mathbf{z} + \boldsymbol{\epsilon}$. As a consequence, the augmented stochastic model reads

$$\begin{aligned}\hat{\mathbf{W}}_{\mathbf{red}}^{\star} &= E(\boldsymbol{\epsilon}_{red}\boldsymbol{\epsilon}_{red}^{T})\\ &= E((\mathbf{B}\Delta\mathbf{z} + \boldsymbol{\epsilon})(\mathbf{B}\Delta\mathbf{z} + \boldsymbol{\epsilon})^{\mathbf{T}})\\ &= E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{T}) + \mathbf{B}\mathbf{E}(\Delta\mathbf{z}\Delta\mathbf{z}^{\mathbf{T}})\mathbf{B}^{\mathbf{T}}\\ &= \hat{\mathbf{W}} + \mathbf{B}\hat{\mathbf{W}}_{\mathbf{z}}\mathbf{B}^{\mathbf{T}}\end{aligned} \tag{10}$$

where $\hat{\mathbf{W}}_{\mathbf{z}}$ is the apriori covariance matrix of the additional parameter. To make a parallel with Eq. (9), $\hat{\mathbf{W}}_{\mathbf{red}}^{\star}$ can be inverted so that

$$\begin{aligned}\hat{\mathbf{W}}_{\mathbf{red}}^{\star-1} &= (\hat{\mathbf{W}} + \mathbf{B}\hat{\mathbf{W}}_{\mathbf{z}}\mathbf{B}^{\mathbf{T}})^{-1}\\ &= (\hat{\mathbf{W}}^{-1} - \hat{\mathbf{W}}^{-1}\mathbf{B}(\mathbf{B}^{\mathbf{T}}\hat{\mathbf{W}}^{-1}\mathbf{B} + \hat{\mathbf{W}}_{\mathbf{z}}^{-1})^{-1}\mathbf{B}^{\mathbf{T}}\hat{\mathbf{W}}^{-1})\end{aligned} \tag{11}$$

It can be seen from Eq. (11) that the equivalence between $\hat{\mathbf{W}}_{\mathbf{red}}^{\star-1}$ and $\hat{\mathbf{W}}_{\mathbf{red}}^{-1}$ holds only if $\hat{\mathbf{W}}_{\mathbf{z}}^{-1}$ is made small or alternatively if $\hat{\mathbf{W}}_{\mathbf{z}}$ dominates in Eq. (10). Moreover, $\hat{\mathbf{W}}_{\mathbf{z}}$ should not contain apriori information on the variance of the process noise (Blewitt 1998) which is already taken into account in $\mathbf{B}$.

The "hidden" tropospheric parameter

In this section, we aim to present didactically how matrices built with Eq. (7) are corresponding to an augmented stochastic model, i.e. a "hidden" estimation of a tropospheric parameter. This highlights how taking correlations into account for short batches can replace the estimation of this additional "non-estimable" parameter.

To this end, we first note that the matrix **B** is filled with the squared root of the elements of $\hat{\mathbf{W}}_{\mathbf{UD,elev}}$ (Kermarrec and Schön 2016). Furthermore, in Eq. (10), we make the assumption that $E\left(\varepsilon\varepsilon^T\right) = \mathbf{I}$ and assume homoscedasticity of the errors defined in Eq. (1). We build $E(\Delta\mathbf{z}\Delta\mathbf{z}^\mathbf{T})$ based on a simplified version of Eq. (5), i.e. $C_t^{t+\tau} = \delta(\alpha\tau)^\nu K_\nu(\alpha\tau)$ and choose the Mátern parameters corresponding to a tropospheric modelling with $[\alpha, nu] = [0.01, 1.05]$, thus identical for all satellites (Kermarrec and Schön 2014). We intentionally disregard the elevation dependency. Therefore $\hat{\mathbf{W}}_\mathbf{z}$ is fulfilled under the aforementioned condition to account for correlations introduced by the data's dependence on the process noise with "no prior information on the variance of the process noise" (Blewitt 1999).

Returning shortly to section 2, we notice that the elevation dependent factor in Eq. (5) is based on a cosine function whose square root is also used to fill **B**. Therefore we can write $\hat{\mathbf{W}}_{\mathbf{UD},fully} = \mathbf{B}\hat{\mathbf{W}}_\mathbf{z}\mathbf{B}^T$ and express the VCM of the augmented noise as

$$\hat{\mathbf{W}}^\star_{\mathbf{red}} = \mathbf{I} + \hat{\mathbf{W}}_{\mathbf{UD,fully}} \tag{12}$$

Condition for the equivalence

We have seen that for the equivalence to hold, **hatW**$_\mathbf{z}$ should be built to account for correlations, so that the process noise dominates in Eq. (11). This can be seen starting for example from the equivalent diagonal model presented the appendix, where correlations appear to act similarly to a large weighting factor of the corresponding diagonal matrix, in this case the identity matrix. If correlations are neglected, $\hat{\mathbf{W}}_\mathbf{z} = \mathbf{I}$. Thus the equivalence is much weaker, besides the fact that it does not correspond anymore to a covariance matrix for the tropospheric parameter. Similarly, if the correlation length is much smaller than the batch length, the corresponding fully populated VCM are sparse and nearly correspond to a diagonal VCM, i.e. the 0-value of the covariance is rapidly reached with respect to the batch length.

Using the proposed Mátern parameter set $[\alpha, \nu] = [0.01, 1.05]$, the corresponding correlation length is approximately 600 s. As a consequence, we propose to define a batch-size limit for the equivalence to hold fixed to 3600 s (1 hour of observations). This is also the often assumed condition whether to estimate a tropospheric parameter, independently of the data rate. Eventually, it is possible to decrease α or increase ν to fill the matrix more strongly. Besides the fact that it deviates strongly from a tropospheric correlation model, it has the disadvantage of impacting also the aposteriori variance factor (Kermarrec and Schön 2017) and can thus only be used under the control that no overestimation occurs which will correspond to an underestimation of the precision (Appendix).

From the reduced matrix to a VCM for GPS phase measurements

We note that in Eq. (12) $\hat{\mathbf{W}}^\star_{\mathbf{red}}$ is not corresponding to a cofactor matrix for GPS, i.e. a value of 1 for the variance for a satellite at 90° is not given anymore. Hence, although the estimates will not be influenced by the scaling (Kutterer 1999), the results of statistical tests such as the overall model test cannot be compared anymore with the usually used ELEV model. Thus we use instead a scaled matrix so that the reduced matrix reads $\hat{\mathbf{W}}^\star_{\mathbf{red}} = \beta_{\mathbf{red}}\mathbf{I} + (\mathbf{1} - \beta_{\mathbf{red}})\hat{\mathbf{W}}_{\mathbf{UD,fully}}$, β_{red} being a noise parameter between 0 and 1. By doing so, we slightly weaken the equivalence by decreasing the impact of $\hat{\mathbf{W}}_{\mathbf{UD,fully}}$ (Kermarrec and Schön 2017). This is unproblematic using the proposed Mátern parameter set and mentioned batch length limit. Eventually the weakening could be compensated by decreasing α from 0.005, using the non-orthogonality of $[\alpha, \nu]$ (Stein 1999).

The circle is now complete as the same expression is obtained as in Eq. (7). As a consequence, when correlations are taken into account with the proposed model of Eq. (5), we account for a tropospheric parameter without estimating it explicitly, a "hidden" parameter.

Note that we could have taken $E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T) = \hat{\mathbf{W}}_{\mathbf{UD,elev}}$ in Eq. (10), which would have corresponded to an elevation dependent noise following Radovanovic (2001). This choice is left to the reader. The authors have a preference for an identity noise matrix to make a parallel with the Tikhonov regularization.

3.2 An additional interpretation of fully populated VCM

In the previous section, we have explained how using fully populated VCM can replace the estimation of a tropospheric parameter, the equivalence being valid as long as the VCM is made sufficiently large, i.e. for short batches.

It is worth additionally mentioning that in case of short batches in GPS positioning, the ideal assumption for the least-squares estimator to be unbiased are often not reached (Rao and Toutenburg 1999, Koch 1999). For example, non-normal errors of the residuals may significate that F-distributions cannot be assumed for the aposteriori variance factor but either student distribution (Williams et al. 2013). Moreover, the condition that the residuals are zero-mean may not be fulfilled, particularly for long baselines when observations have drifts due to unmodelled remaining effects. Fortunately, when fully populated matrices build with Eq. (5) are taken into account in the least-squares adjustment, a filtering of such unwanted effects is obtained. This can be seen thanks to Eq. (6), e.g. the smoothness and the correlation parameter impact the frequency content of the observations. As a consequence, using FGLSE with the FULLY model instead of the purely diagonal ELEV model, a decrease of the bias of the least-squares solution is obtained corresponding to a lower loss of efficiency. This leads to a more trustworthy position with an associated non overoptimistic precision and better test statistics such as overall model, outlier detection tests or ambiguity validation tests. (see appendix for more details).

3.3 Ambiguity fixed

Through the development of the equivalence, we have considered a global model and assumed that the ambiguity is estimated as float together with the position and not fixed in advance (Eq. (1)). If the integer ambiguities are known in advance, the equivalence still holds. As it is not used for the less biased float ambiguity under a more correct stochastic model particularly for short batches, the solution (i.e. coordinates) obtained with different VCM will be less different.

4 A case study

The concept of the hidden parameter is not straightforward to validate. Indeed as its name indicates, it corresponds to cases where no parameter can be estimated. In order to overcome this issue, a methodology is proposed based on decreasing the batch length and comparing the solution found under fully populated VCM with respect to a diagonal VCM in cases where the true position is known.

4.1 Observations

Data from the European Permanent Network EPN (Bruyninx et al. 2012) from two stations KRAW and ZYWI are chosen as example for a medium baseline (80km) positioning scenario. OMC observations are computed with 30s rate observations and a cut-off of 3°. The ionospheric and tropospheric delays are partially estimated in a preprocessing step with the Klobuchar and Hopfield models, respectively. A relative positioning scenario is considered and the North East Up (NEU) coordinates are estimated at GPS day DOY220, year 2015. The starting time is GPS-SOD 6000s and was taken arbitrarily. It was shown not to impact the conclusions, i.e. the geometry playing a minor role in the results of our comparison (Kermarrec and Schön 2017, Appendix 2). The reference values are the long term station coordinates from the EPN solution.

4.2 Methodology

We compute the least-squares results given by the FULLY VCM described in section 3 and the diagonal ELEV matrices. We place ourselves in a case where it is assumed that no tropospheric parameter can be estimated so that batches have a length of maximum 100 epochs at 30 s. In case of longer batches, an additional tropospheric parameter should be taken into consideration as the equivalence does not hold anymore, i.e. the FULLY model does not replace the tropospheric parameter. Five batch lengths were selected to show the influence of the fully populated VCM on the float solution when no tropospheric parameter is estimated:

1. 20 batches with 100 epochs (100-epochs-case, 60000 s)
2. 25 batches with 80 epochs (80-epochs-case, 60000 s)
3. 33 batches with 60 epochs (60-epochs-case, 59400 s)
4. 50 batches with 40 epochs (40-epochs-case, 60000 s)
5. 100 batches with 20 epochs (20-epochs-case, 60000 s)

As previously mentioned, a batch approach is retained, i.e. one solution is computed for each batch. The aim of this methodology is to show how decreasing the batch length, i.e. strengthening the equivalence augmented stochastic versus functional model, will impact the positioning.

To this end, a global estimator of the least-squares solution is retained. The reference being in our case the 0 vector since the position was known exactly, the 3Drms is computed for each batch and averaged over all batches for both stochastic models of consideration. The 3Drms difference between the ELEV and FULLY is then formed, i.e. $\sum_{i=1}^{m}(3Drms_{FULLY}(i))-\sum_{i=1}^{m}(3Drms_{ELEV}(i))$ where m is the number of batches corresponding to case 1-5. As the estimation of a tropospheric parameter mainly influences the height component, we similarly compute the rms difference for the height component only. For short batches, the F-distribution of the ratio $\frac{\hat{\sigma}_{\hat{W}}^2}{\sigma_0^2}$ may not be given anymore (Williams et al. 2003, Kermarrec et al. 2017). Thus we only compute the mean of the aposteriori variance factor over all batches and compare it with the assumed apriori value to assess roughly the trustworthiness of the solution. We took σ_0 = 4 mm, i.e. a relevant and plausible value for double differences observations.

We choose to let the ambiguities float in order to have a "global" functional model and make use of the better estimated float ambiguities when improving the stochastic model. Moreover, a comparison of the results with different stochastic models is easier to follow as the fixing to the correct ambiguities strongly improve the final solution. Fixing the ambiguities in advance in a preprocessing step leads to less high differences between ELEV and FULLY model following the results of Kermarrec and Schön (2017). Nevertheless, using fully populated VCM, more batches can be fixed with respect to the ELEV model as described in the appendix. As a consequence, the conclusions of the case study will not be impacted by this choice.

For the sake of completeness and although unrealistic, we add the results given when an additional tropospheric parameter is estimated with the ELEV model for the 40-epochs case.

4.3 Results

The results of the case study are presented in Table 1.

Impact of decreasing the batch length

The least-squares estimator being unbiased, the impact of the stochastic model on the positioning decreases for longer batches. For the 100-epochs-case for example, a 3Drms difference of 0.1 mm is obtained which grows to 106 mm for the 20-epochs-case, highlighting the strong impact of the FULLY populated VCM. If the difference increases, $\sum_{i=1}^{m}(3Drms_{FULLY}(i))$ decreases and becomes closer to the 0 value. As mentioned in section 3, this result gives weight to the equivalence augmented stochastic – functional model as soon as the FULLY VCM are "full". Additionally and for case 1 for example, the value of $E\left(\hat{\sigma}_{WFULLY}\right) = 4.1$ mm is close to the chosen $\sigma_0 = 4$ mm so that the solution can be considered as trustworthy. This is not the case for the ELEV VCM where $E\left(\hat{\sigma}_{WELEV}\right) = 6.5$ mm highlights a model misspecification due to the biased aposteriori variance factor. The same conclusions hold true for the other cases, although the differences between ELEV and FULLY decreased as expected. Improving the stochastic model by means of correlations is thus of main importance to obtain both less biased test statistics and a better positioning.

Using the equivalence and without weakening the data strength, the Up component is strongly improved by up to 37 mm for the 20-epochs-case. This highlights the main importance of using fully populated matrices for short batches. This difference decreases to 10 mm for the 60-epochs-case and is nearly 0 for the 100-epochs case, i.e. for longer batches the use of fully populated model do not replace the estimation of an additional parameter.

Estimating a tropospheric parameter for short batches

In case an additional tropospheric parameter is nevertheless estimated - for case 4 for example -, as done for longer batches, we note that $E(\hat{\sigma}_{WELEV}) = 7.1$ mm. Thus a model misspecification is guessed using the ELEV VCM which is confirmed by the difference between the 3Drms FULLY-ELEV which is up to 90 mm higher than without estimating a parameter. As a consequence, the FULLY model is without a doubt a better alternative than the ELEV model.

Fixing the ambiguities to integer

If the ambiguities are fixed in advance in a preprocessing step, the differences between the models decrease. For the case 1 for example, a 3Drms difference of only 7 mm is obtained. Thus the effect of the FULLY model still impacts the solution but at a lower level. If the ambiguities are estimated as float and fixed for each batch using the ratio test with a threshold of 0.5 (Wei and Schwarz 1995), the fixing to integer can be improved by 5-10% following the results of the simulation presented in the appendix. As a consequence, improving the stochastic model will have a "snowball effect" on the 3Drms, the results of test statistics (ambiguity, outlier detection test, overall model test) being less biased as shown in the appendix for the ambiguity validation test (see also Li et al. 2016). Thus we definitively advice using such models, independently of the strategy used and particularly for short batches when the troposphere is expected to influence the results.

5 Conclusion

In this contribution, we made use of a weighted Mátern covariance function to describe the elevation dependent correlations of GNSS phase observations. For correlations due to turbulent tropospheric variations of the index of refractivity, the Mátern parameters (smoothness and correlation length) can be fixed apriori based on physical considerations. This function was mathematically shown to correspond to taking an additional tropospheric parameter into account without having to estimate it separately. This equivalence augmented stochastic model-functional model can be used as soon as the separability between the tropospheric parameter and the Up component is not ensured in the least-squares adjustment. It is thus particularly interesting for estimating the Up component with a higher trustworthiness in case of short batches of observations.

Table 1: 3Drms differences (FULLY-ELEV) model and rms differences for the Up component. Five different batch lengths are computed. The ambiguities are let float.

No trop.	Case 1	Case 2	Case 3	Case 4	Case 5
3Drms difference FULLY-ELEV [mm]	0.1	16.9	23.0	32.9	106.0
rms difference Up component FULLY-ELEV [mm]	−0.4	5	23.1	14.4	37.4
trop.	Case 4				
3Drms difference FULLY-ELEV [mm]	113.1				
rms difference Up component FULLY-ELEV [mm]	69.6				

In a case study using double differenced observations from a 80 km baseline, this equivalence leads to an improvement of up to 10 cm for observations divided in batches of 20 epochs at 30 s with respect to an elevation dependent diagonal VCM when using the float ambiguities. Taking correlations into account lead thus in a noticeable way to an improvement of the positioning solution for short batches, particularly when the ambiguities cannot be fixed to integer with enough reliability and let float. The impact decreases for longer batches and if the ambiguities are fixed. However, less biased test statistics and a less overoptimistic precision is still obtained with respect to the purely diagonal model. The equivalence holds as soon as the covariance is made sufficiently large. This condition was translated for the GPS case and shown to be plausible for batches up to 3600 s length.

Acknowledgement: The authors gratefully acknowledge the EPN network and corresponding agencies for providing the data free of cost. The Lambda toolbox is provided freely by TU Delft.

A Appendix 1

The equivalent diagonal model

In this appendix, some insights on how correlations act on the apriori cofactor of the estimates (called the precision) and the ratio test are given. For didactic purposes, we use an AR(1) model for GPS phase correlations which corresponds to a smoothness of 1/2 in our proposal. In that particular case, the inverse of the corresponding VCM can be exactly expressed thanks to the known or estimated autocorrelation coefficient ρ_{AR} (Rao and Toutenburg 1999). In Kermarrec and Schön (2016), it is explained how correlations can be taken into account thanks to a reduced diagonal VCM.

The inverse of the equivalent VCM for the VCM from an AR(1) process reads:

$$\mathbf{W}^{-1}_{\mathbf{AR(1)\_EQUI}} = \frac{1}{1-\rho_{AR}^2}\begin{bmatrix} 1-\rho_{AR} & 0 & 0 & \cdots & 0 & 0 \\ 0 & (1-\rho_{AR})^2 & 0 & \ddots & 0 & 0 \\ 0 & 0 & (1-\rho_{AR})^2 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & (1-\rho_{AR})^2 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1-\rho_{AR} \end{bmatrix}$$

To derive the inverse of the FULLY VCM, we assume low variations of the satellite elevation. Thus the elevation dependent factor of the covariance matrix derived thanks to the proposed model can be factorized. As a consequence, the elements of the corresponding equivalent diagonal VCM sorted per epochs for one satellite are:

- First and last diagonal values: $\frac{1}{\sin(El_i(t))^2}(1+\rho_{AR})$
- All other diagonal values: $\frac{1}{\sin(El_i(t))^2}\left(\frac{1+\rho_{AR}}{1-\rho_{AR}}\right))$

As highlighted in Luati and Proietti (2011), the equivalent VCM thus has two diagonal values -at the beginning and the end of a batch of observations- that are lower than the middle diagonal values, all elements being simply proportional to the ELEV model, i.e. corresponding to a higher weighting as $\left(\frac{1+\rho_{AR}}{1-\rho_{AR}}\right) > 1$.

Precision of the least-squares results using FULLY

This particularity of the equivalent VCM (or its inverse) has the consequence that when correlations are taken into account, the impact of the extreme values on the results is getting negligible for long batches. Thus the matrix $\mathbf{Q_{\hat{x}FULLY}} = (\mathbf{A^T\hat{W}^{-1}_{FULLY}A})^{-1}$ and $\mathbf{Q_{\hat{x}ELEV}} = (\mathbf{A^T\hat{W}^{-1}_{ELEV}A})^{-1}$ are only linked by a scaling factor depending on the correlation length. This result is derived for an AR(1) model and may be slightly different for higher smoothness where more diagonal entries than only the 2 first values are different than the middles values. Luatti and Proietti (2011) show for example that for an AR(3) model 3 first values were different. Thus, even for short batches, a scaling factor can link with a good approximation $\mathbf{Q}_{\hat{\mathbf{x}}\mathbf{FULLY}}$ and $\mathbf{Q}_{\hat{\mathbf{x}}\mathbf{ELEV}}$ when our proposed model is used. As a consequence, the error ellipsoids will have slightly the same orientation in space and the precision with a FULLY model will be more realistic, i.e. no overestimation as for diagonal VCM will occur. The least-squares solution is therefore more trustworthy.

Impact on ambiguity resolution of FULLY

The second consequence of this result can be shown at the ambiguity fixing level. Indeed, when using the Fixed Failure Rate Ratio Test (FFRT) with a FULLY model to estimate an accurate threshold (Wang and Feng 2013), it is expected that the same value as with an ELEV VCM will be found.

Independently of the chosen threshold, the impact of misspecifying the stochastic model up to neglecting correlations on the ratio test defined as $R = \frac{||\hat{\mathbf{x}}^1_{\mathbf{A,fix}} - \hat{\mathbf{x}}_{\mathbf{A,float}}||_{\mathbf{Q_A}}}{||\hat{\mathbf{x}}^2_{\mathbf{A,fix}} - \hat{\mathbf{x}}_{\mathbf{A,float}}||_{\mathbf{Q_A}}} = \frac{d_1}{d_2} \leq \mu_R$ (Euler and Schaffrin 1991) can be assessed. We call $\hat{\mathbf{x}}^1_{\mathbf{A,fix}}$, $\hat{\mathbf{x}}^2_{\mathbf{A,fix}}$ the two vectors of integer candidates that are corresponding to the two smallest values of the distance between the float and two fixed ambiguity vectors in the metric of the covariance matrix.

To assess the impact of the FULLY model on the ambiguity fixing, we make use of Monte Carlo simulations where time series corresponding to a true VCM with $[\alpha, nu] = [0.01, 1]$ are computed. In order to assess the sensitivity of the model, the parameters $[\alpha, \nu]$ are varied around the true set where it can be shown from Eq. (6) that increasing corresponds to neglecting correlations. A constellation of 8 satellites observed during 3000s was taken in consideration and a relative positioning strategy used. To the 10000 simulated time series corresponding to the correlation structure of reference, the same but arbitrary ambiguity vector was added. The following results are independent of the constellations or the batch length (Kermarrec and Schön 2017 Appendix 2).

From Fig. 1, it can be clearly seen that neglecting correlations corresponds to a small increase of the ratio test by 0.1 and thus to a slight decrease of the probability to fix the ambiguities for a given similar threshold. This fact may be amplified in real cases when the least-squares assumption are slightly violated. This effect is emphasized when the correlation parameter is smaller than the reference, highlighting the importance of non-underestimating . In the ideal case of simulations, the ambiguities were fixed correctly with the Lambda method whether correlations are correctly taken in consideration or neglected. This may not be the case for real cases, particularly for small batches and thus correlations when present should not be disregarded as developed previously. As a consequence, it is expected that taking correlations into account leads to less biased ambiguity validation tests and thus allows an increase of the ambiguity success rate with respect to using a diagonal VCM for an assumed fix threshold.

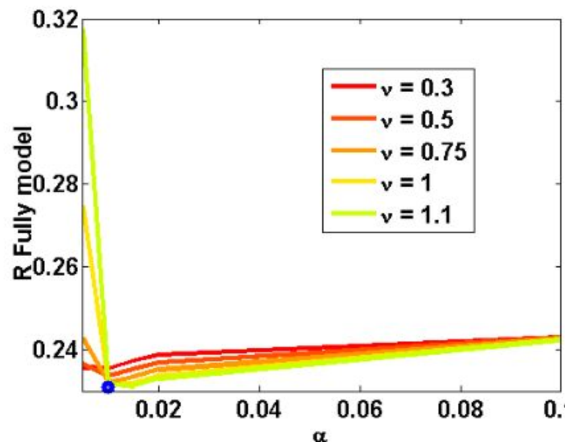


Figure 1: Results of the Monte Carlo simulations study with 10000 iterations per Mátern parameters set. Simulated observations are correlated with $[\alpha, \nu]_0 = [0.01, 1]$ (blue point), the Mátern parameters of the estimated VCM are varied. The mean value of the ratio R is presented.

References

Abramowitz M, Segun IA (1972) Handbook of mathematical functions. Dover, New York edition

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

Amiri-Simkooei A.R., Jazaeri S., Zangeneh-Nejad F. and Asgari J., 2016, Role of stochastic model on GPS integer ambiguity resolution success rate, GPS Solut. 20, 1, 51-61

Bischoff W., Heck B., Howind J. and Teusch A., 2005, A procedure for testing the assumption of homoscedasticity in least squares residuals: a case study of GPS carrier-phase observations. J Geod., 78, 7-8, 379-404.

Blewitt G., 1998, GPS Data Processing Methodology: From Theory to Applications. In Teunissen PJG and Kleusberg A (Eds.) GPS for Geodesy (pp231-270). 2nd ed. Springer-Verlag Berlin Heidelberg New York

Bona P., 2000, Precision, cross correlation, and time correlation of GPS phase and code observations. GPS Solut., 4, 2, 3-13

Bruyninx C., Habrich H., Söhne W., Kenyeres A., Stangl G. and Völksen C., 2012, Enhancement of the EUREF Permanent Network Services and Products. Geodesy for Planet Earth, IAG Symposia Series, 136:27-35

Erickson C., 1992, Investigation of CA code and carrier measurements and techniques for rapid static GPS surveys. Report of the Department of Geomatics Engineering of the University of Calgary, vol 20044

El-Rabbany A., 1994, The effect of physical correlations on the ambiguity resolution and accuracy estimation in GPS differential positioning. PhD thesis, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada

Euler H.J. and Schaffrin B., 1991, On a Measure for the Discernability between Different Ambiguity Solutions in the Static Kinematic GPS mode. IAG Symposia No 107, Kinematic Systems in Geodesy. Springer, Berlin Heidelberg New York 285-295

Gelfand AE, Diggle P, Fuentes M, Guttorp P (2010) Handbook of Spatial Statistics. Boca Raton: Chapman & Hall/CRL

Hofmann-Wellenhof B., Lichtenegger H. and Collins J., 2001, GPS Theory and Practice, Fifth, revised edition. Springer Wien New York

Howind J., Kutterer H. and Heck B., 1999, Impact of temporal correlations on GPS-derived relative point positions, J Geod., 73, 5, 246-258

Kermarrec G. and Schön S., 2014, On the Mátern covariance family: a proposal for modeling temporal correlations based on turbulence theory, J Geod., 88, 11, 1061-1079

Kermarrec G and Schön S., 2016, Taking correlations into account with a diagonal covariance matrix, J Geod., 90, 9, 793-805

Kermarrec G. and Schön S., 2017, Apriori fully populated covariance matrices in least-squares adjustment – case study: GPS relative positioning. J Geod., 91, 5, 465-484

Kermarrec G., Schön S. and Kreinovich V., 2017, Possible explanation of empirical values of the Màtern smoothness parameter for the temporal covariance of GPS measurements. Appl. Math. Sci., 11, 35, 1733-1737.

Koch K.R., 1999, Parameter Estimation and Hypothesis Testing in Linear Models, Springer Berlin.

Kutterer H., 1999, On the sensitivity of the results of Least-Squares adjustments concerning the stochastic model, J Geod., 73, 7, 350-361.

Li B., Lou L. and Shen Y., 2016, GNSS Elevation-Dependent Stochastic Modelling and Its Impacts on the Statistic Testing. J Surv Eng., 142, 2, 04015012

Li B., Zhang L. and Verhagen S., 2017, Impacts of BeiDou stochastic model on reliability: overall test, *w*-test and minimal detectable bias, GPS Solut. 21,3, 1095-1112

Luati A. and Proietti T., 2011, On the equivalence of the weighted least squares and the generalised least squares estimators, with applications to kernel smoothing, Ann. Inst. Stat. Math., 63, 4, 673-686

Luo X., Mayer M. and Heck B., 2012, Analysing Time Series of GNSS Residuals by Means of ARIMA Processes. International Association of Geodesy Symposia, Vol. 137, pp 129-134

Luo X., Mayer M., Heck B., Awange J.L., 2014, A realistic and easy-to-implement weighting model for GNSS phase observations, IEEE Transactions on Geoscience and Remote Sensing 52, 10, 6110-6118

Mátern B., 1960, Spatial variation-Stochastic models and their application to some problems in forest surveys and other sampling investigation. Medd. Statens Skogsforskningsinstitut 49, 5

Radovanovic R.S., 2001, Variance-Covariance Modeling of Carrier Phase Errors for Rigorous Adjustment of Local Area Networks. IAG 2001 Scientific Assembly. Budapest, Hungary, September 2-7, 2001

Rao C. and Toutenburg H., 1999, Linear Models, Least-Squares and Alternatives. Springer-Verlag New York Second Edition

Schön S. and Brunner F.K., 2008, A proposal for modeling physical correlations of GPS phase observations, J Geod. 82, 10, 601-612

Schwieger V., 2007, Determination of Synthetic Covariance Matrices – an Application to GPS Monitoring Measurements. Proceedings on 15^{th} European Signal Processing Conference, Poznan, Polen, 03-07.09.2007

Stein M. L., 1999, Interpolation of spatial data. Some theory for kriging. Springer New York.

Teunissen P.J.G., 1995, The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation, J Geod, 70, 1-2, 65-82

Teunissen P.J.G. and Amiri-Simkooei A.R., 2008, Least-squares variance component estimation, J Geod 82, 2, 65–82

Teunissen P.J.G. and Verhagen S., 2009, The GNSS Ambiguity Ratio-Test revisited: a better way of using it, Survey Review 41:138-151

Tralli D.M. and Lichten S.M., 1990, Stochastic estimation of tropospheric path delays in global positioning system geodetic measurements. Bull. Geodesique 64, 2, 127-159

Verhagen S. and Teunissen P.J.G., 2013, The ratio test for future GNSS ambiguity resolution, GPS Solut 17, 4, 535-548

Vermeer M., 1997, The precision of geodetic GPS and one way of improving it, J Geod 71:240-245

Wang J., Stewart M.P. and Tsakiri M., 1998, Stochastic modelling for GPS static baseline data processing, J Surv Eng., 124, 4, 171-181

Wang L. and Feng Y., 2013, Fixed Failure Rate Ambiguity Validation Methods for GPS and Compass. In: Sun J., Jiao W., Wu H., Shi C. (eds) China Satellite Navigation Conference (CSNC) 2013 Proceedings. Lecture Notes in Electrical Engineering, vol 244. Springer, Berlin, Heidelberg

Wei M. and Schwarz K.P., 1995, Fast ambiguity resolution using an integer nonlinear programming method, In: Proceedings of the ION GPS-1995, Palm Springs CA, pp1101-1110

Wheelon A.D., 2001, Electromagnetic Scintillation part I Geometrical optics. Cambridge University Press, Cambridge

Williams M.N., Gomez Grajales C.A. and Kurkiewicz D., 2013, Assumptions of Multiple Regression: Correcting Two Missconceptions. Practical Assessment. Research and Evaluation 18, 11