Estimation of the minimal detectable horizontal acceleration of GNSS CORS

RENAN RODRIGUES TOLEDO COSTA[1], IVANDRO KLEIN[1,2], ELIEL JESSÉ MORAIS DE JESUS JUNIOR[1], CHRISTIAN GONZALO PILAPANTA AMAGUA[1] AND PAULO SERGIO DE OLIVEIRA JUNIOR[1]

1 Graduate Program in Geodetic Sciences, Federal University of Parana, Curitiba 81531-990, Brazil (renanrtc@gmail.com)

2 Land Surveying Program, Federal Institute of Santa Catarina, Florianópolis SC 88020-300, Brazil

Received: October 26, 2023; Revised: May 27, 2024; Accepted: June 27, 2024

ABSTRACT

Earth's surface velocities are routinely extracted from Global Navigation Satellite System (GNSS) position time series. In addition to velocity estimates, acceleration may be a crucial parameter for modeling non-linear motion. Typically, a statistical hypothesis test is employed to evaluate the significance of the involved parameters and guide the selection of the appropriate model. In this contribution, we formulate a statistical test procedure from the generalized likelihood ratio test to analyze the significance of the acceleration in the model. The proposed procedure is compared with results obtained using the Akaike Information Criterion and Bayesian Information Criterion. Additionally, Minimal Detectable Horizontal Acceleration is provided as an indicator of the sensitivity of the acceleration detection. The GNSS time series of position estimates from the Nevada Geodetic Laboratory were used for this study. The experiments demonstrated a good agreement between the statistical test proposed and the information criteria approach. Therefore, the proposed statistical test may be another criterion to help the user in the important task of model selection.

Keywords: Global Navigation Satellite System, statistical test, acceleration, AIC, BIC, *MDHA*, time series

1. INTRODUCTION

Geodesy has made significant contributions to geodynamic studies, mainly due to the maturation and development of Global Navigation Satellite System (GNSS). The rise of GNSS systems and long-term accumulated GNSS observations from global reference stations have provided valuable data for geodesy and geodynamics studies since the 1990s. The GNSS Continuously Operating Reference Station (CORS) serves as a static geodetic reference, and its long-term observation series provide (geo)kinematical information (*Kenyeres and Bruyninx, 2004*). Thus, GNSS CORS spread worldwide are important for determining Earth's deformation processes, such as the tectonic plate movements and

vertical deformation of the crust (*Ren et al., 2021*). The so-called lithospheric plates are constantly moving in different directions and with varying magnitudes of displacement. Therefore, the position of a point located on a lithospheric plate, such as a GNSS CORS, also varies over time.

Geodetic coordinates on the Earth's surface are determined by a specific geodetic reference system, which is defined and realized by spatial geodetic techniques such as GNSS and others. Conventionally, the reference used is the International Terrestrial Reference System (ITRS), with its realization being the International Terrestrial Reference Frame (ITRF). The ITRF essentially consists of a set of positions for ITRF stations, such as GNSS CORS, associated with a specific reference epoch and their respective velocities. The station coordinates used in the realization of the reference system, located on the Earth's surface, require periodic updates. As a result, several realizations of the ITRF are available. For instance, ITRF2020 accurately models the nonlinear station motions for seasonal signals (annual and semi-annual) found in the GNSS CORS time series. These updates are crucial for maintaining accurate and up-to-date reference systems (*Altamimi et al. 2016, 2023*).

In recent years, the use of GNSS CORS time series has become more frequent in several research fields, such as geophysics, geodynamics and geodesy (*Kreemer et al., 2014*; *Bogusz et al., 2016a,b*; *Klos et al., 2018*; *Montillet and Bos, 2020*; *Ren et al., 2021)*. Obtaining daily solutions for the coordinate time series of GNSS CORS, encompassing over 10 years of data, provides a substantial dataset for analyzing station components. Such solutions are valuable for the assessment of spatial and temporal evolution.

When we analyze the GNSS CORS time series, it is usually assumed that each horizontal component is described by the sum of the linear rate and various periodic terms such as the annual and semi-annual seasonal signals (*Van Dam et al., 2001*; *Dong et al., 2002*; *Bevis and Brown, 2014*). These effects represent variations that may be associated with events or cycles occurring on an annual timescale, as well as variations taking place at six-month intervals and are influenced by factors such as seasonal changes, the Earth's tilt, and climatic phenomena. Jumps or offsets can also occur, due, e.g., to replacement in the antenna or receiver of the GNSS CORS (*Montillet and Bos, 2020*).

Furthermore, frequently, the parameters of interest in the time series of GNSS coordinates are the velocities and their associated uncertainties, which must be determined with higher reliability. However, it is well-known that not all GNSS CORS time series follow a simple linear behavior. According to *Bogusz et al. (2016a)*, since the time series do not always follow a linear behavior, an acceleration term in the form of a quadratic polynomial function is added to the model to better describe the non-linear movement. This non-linear motion may be a response to purely geophysical processes. Moreover, other models can be included, such as post-seismic relaxation functions (*Langbein et al., 2006*). As demonstrated by *Bos et al. (2020)*, it can be said that coordinate time series can be interpreted as trajectories. Therefore, the kinematic models that geodesists and geophysicists use to describe these time series are trajectory models. Figure 1 illustrates an integration of all these effects into a unified trajectory model. Here, we will restrict to the horizontal (east and north) components, being the vertical component outside the scope of this paper (see, e.g., *Van Dam et al., 2001*).

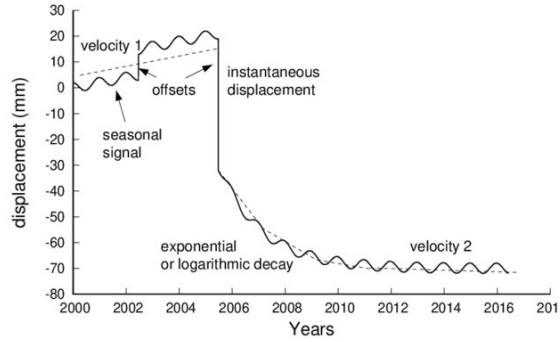

Fig. 1. Sketch of a trajectory model containing common phenomena (*Bos et al., 2020*, © 2020 Springer Nature Switzerland AG).

We assume that the horizontal displacement is the combination of a deterministic model and stochastic noise. The observed motion of each site in each horizontal component can be given as follows (*Bos et al., 2020*):

$$\Delta_{0-k} = \Delta t_{0-k} v_i + \frac{\Delta t_{0-k}^2}{2} \nabla_i + \sum_{j=1}^{m_j} c_j H\left(t_k - t_j\right) + \sum_{p=1}^{2} a_p \sin\left(\omega_p t_k\right) + b_p \cos\left(\omega_p t_k\right), \quad (1)$$

where Δ_{0-k} is the displacement from the reference epoch t_0 to epoch t_k (in m), Δt_{0-k} is the time difference from the reference epoch t_0 to the considered epoch t_k (in years), v_i is the velocity of the horizontal component (in m year^{-1}), ∇_i is the acceleration of the horizontal component (in m year^{-2}), the subscript i distinguishes between the east and north components, H is the Heaviside or unit step function, c_j describes the direction and magnitude of the jump which occurs at the time t_j (in m), m_j is the number of jumps, a_p and b_p are the amplitude coefficients for the sine and cosine terms of the periodic signal (in m), and ω_p is the angular velocity (in rads). Two pairs of periodic signals were considered for the annual and semiannual components, as is standard in the literature (*Bevis and Brown, 2014*; *Klos et al., 2018*).

In recent decades, numerous studies in the literature have focused on estimating the horizontal movement of GNSS CORS (see, e.g., *Bevis and Brown, 2014*; *Kreemer et al., 2014*; *Bogusz et al., 2016b*). The improved modeling of non-linear long-term trends has recently been incorporated into the ITRF2014 and ITRF2020 realizations. However, this is restricted to post-seismic deformation for GNSS CORS located at sites impacted by major earthquakes (*Altamimi et al., 2016, 2023*).

Therefore, the question of whether the non-linear long-term trend must be considered in the horizontal components of GNSS CORS time series remains open. This study aims to contribute to this research field by introducing the application of the generalized likelihood ratio test (GLRT), as presented by *Teunissen (2006)*, to assess the statistical significance of the estimated acceleration in Eq. (1). Additionally, the reliability theory proposed by

Baarda (1968) will be applied here in a novel context, aiming to determine the Minimal Detectable Horizontal Acceleration (*MDHA*).

Statistical tests and reliability measures are routinely applied in other research fields, such as quality control and deformation analysis of geodetic networks (*Rofatto et al., 2017, 2020*; *Prószyński and Lapiński, 2021*). It should be noted that acceleration estimation has already been investigated in GNSS networks for structural monitoring (*Durdag et al., 2018*). The GLRT has also been previously applied in the deformation analysis of geodetic networks (see, e.g., *Zaminpardaz et al., 2020*). However, to the best of our knowledge, there are no similar studies in the literature related to the horizontal components of GNSS CORS time series.

2. STATISTICAL PROCEDURE FOR ACCELERATION DETECTION

Here we briefly describe the GLRT proposed for the significance test of acceleration in the horizontal components of the GNSS CORS time series. Details are found, e.g., *in Teunissen (2006)* or *Rofatto et al. (2020)*. The sample space, which is the set of all possible outcomes for the experiment, is divided into two parts: 1) the rejection region, also known as the critical region and 2) the acceptance region, also known as the non-rejection region, of the null test hypothesis H_0 (*Teunissen, 2006*). In our case, the null hypothesis H_0 is expressed assuming null acceleration ($\nabla_i = 0$) in Eq. (1). Therefore, the null hypothesis H_0 is given as follows:

$$H_0: \quad \boldsymbol{y} = \mathbf{A}\boldsymbol{x}:$$

$$\Delta_{0-k} = \Delta t_{0-k} v_i + \sum_{j=1}^{m_j} c_j H\left(t_k - t_j\right) + \sum_{p=1}^{2} a_p \sin\left(\omega_p t_k\right) + b_p \cos\left(\omega_p t_k\right), \tag{2}$$

$$\mathbf{D}(\boldsymbol{y}) = \mathbf{C}.$$

Here $\boldsymbol{y}$ is the vector of n observations, $\mathbf{A}$ is the design or jacobian matrix, Δ_{0-k} is the vector of u unknown model parameters (in this case v_i, c_j, a_p, b_p and ω_p), $\mathbf{C}$ is the covariance matrix of $\boldsymbol{y}$, $\mathbf{D}$ is the dispersion operator. In contrast to H_0, the acceleration is significant ($\nabla_i \neq 0$) on the alternative hypothesis H_A given as follows:

$$H_A: \quad \boldsymbol{y} = \mathbf{A}\boldsymbol{x} + c_i \nabla_i:$$

$$\Delta_{0-k} = \Delta t_{0-k} v_i + \frac{\Delta t_{0-k}^2}{2} \nabla_i + \sum_{j=1}^{m_j} c_j H\left(t_k - t_j\right) + \sum_{p=1}^{2} a_p \sin\left(\omega_p t_k\right) + b_p \cos\left(\omega_p t_k\right), \tag{3}$$

$$D(\boldsymbol{y}) = \mathbf{C}.$$

In other words, we intend to test if the $q = 1$ additional parameter (∇_i) in the model of H_A is statistically significant or not according to the GLRT.

Regarding the decision about the hypotheses, two types of errors can be made. Type I error is about rejecting H_0 when it is true, that is, the occurrence of a "false positive", with a probability of occurrence designated by α (the significance level of the test). On the other

hand, Type II error is about accepting H_0 when H_0 is false, that is, when H_A is true, whose probability of occurrence is designated by β. In other words, β corresponds to the probability of a "false negative" (*Teunissen, 2006*; *Rofato et al., 2017, 2020*).

The test statistic (T_q) is given by the difference between the weighted sum of squared residuals (w.s.s.r) under the null hypothesis ($\hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{W}\hat{\boldsymbol{v}}$), where $\hat{\boldsymbol{v}}$ is the residual vector, and the w.s.s.r under the alternative hypothesis ($\hat{\boldsymbol{v}}_A^{\mathrm{T}}\mathbf{W}\hat{\boldsymbol{v}}_A$), as demonstrated in *Teunissen (2006)*, where $\mathbf{W}$ is the weight matrix, $\mathbf{W} = \mathbf{C}^{-1}$. With an unknown variance factor (σ_0^2), the test statistic T_q in H_0 follows the F distribution with $q = 1$ degrees of freedom in the numerator and $n - k$ degrees of freedom in the denominator, with $k = u + 1$ (u is the number of parameters of the model, but it also estimates the posterior variance factor). Otherwise, the test statistic T_q in H_0 follows the χ^2 distribution with $q = 1$ degrees of freedom. Thus, the test decision is given by:

$$\begin{aligned} &\sigma_0^2 \text{ unknown:} \quad T_q = \hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{W}\hat{\boldsymbol{v}} - \hat{\boldsymbol{v}}_A^{\mathrm{T}}\mathbf{W}\hat{\boldsymbol{v}}_A, \quad \text{do not reject } H_0 \text{ if } T_q \le F_{q=1,n-k,\alpha}, \\ &\sigma_0^2 \text{ known:} \quad\;\; T_q = \hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{W}\hat{\boldsymbol{v}} - \hat{\boldsymbol{v}}_A^{\mathrm{T}}\mathbf{W}\hat{\boldsymbol{v}}_A, \quad \text{do not reject } H_0 \text{ if } T_q \le \chi^2_{q=1,\alpha}. \end{aligned} \tag{4}$$

In the case of modeling time series by Hector software (HS), it does not provide the w.s.s.r in each model H_0 and H_A (see, e.g., *Bos, 2022*). Therefore, considering the relation between the estimated (a posteriori) variance factor $\hat{\sigma}^2$ and the w.s.s.r (see, e.g., *Lehmann and Lösler, 2016*), we can calculate $\left(\hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{W}\hat{\boldsymbol{v}}\right)_{H_i}$ for cases with acceleration ($H_i = H_A$) and without acceleration ($H_i = H_0$), based on the respective σ_{H_i} (standard deviation of the driving noise) obtained in the HS output report, as follows:

$$\left(\hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{W}\hat{\boldsymbol{v}}\right)_{H_i} = \sigma^2_{H_i}\left(n - k_{H_i}\right). \tag{5}$$

Here n is the number of observations and k_{H_i} is the number of model parameters in H_i for H_0 or H_A: $k_{H_A} = k_{H_0} + 1$. In the alternative hypothesis, there is one additional parameter, which is the acceleration. See *Bos (2022)* for details about the HS output report.

It should be noted that other model parameters in Eq. (1), such as jumps and periodic signals, can also be tested for significance. However, our analysis focuses on the long-period trend of time series, and thus the significance test of other model parameters rather than acceleration is beyond the scope of this study. The non-linear modeling of the long-period trend was included only in the last two ITRF realizations (*Altamimi et al., 2016, 2023*) and only for stations with post-seismic events. Therefore, it remains a current research topic with many possibilities for further investigation.

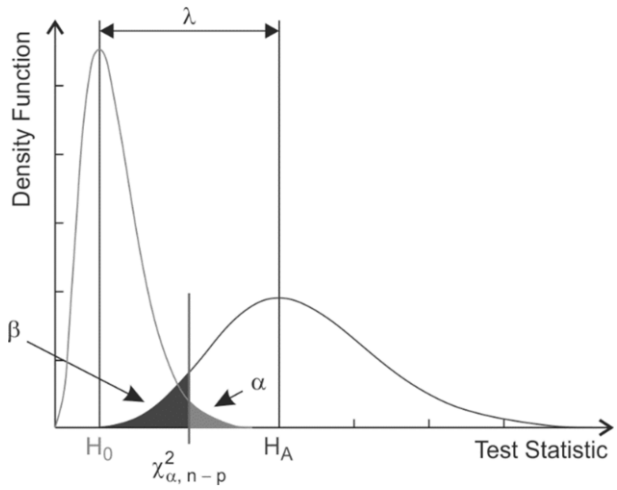

Fig. 2. The central and non-central density functions (after *Kuusniemi and Lachapelle, 2004*, © 2004 The Institute of Navigation).

3. MINIMAL DETECTABLE HORIZONTAL ACCELERATION

The conventional reliability theory proposed by *Baarda (1968)* is related to the detection and identification of bias (non-random errors) in observations through statistical tests. Its reliability theory employs appropriate measures to quantify the *MDB* (Minimal Detectable Bias) of each observation. The *Baarda's (1968)* reliability theory seeks to test the significance of the additional parameter ∇_i of H_A. In that case, the possible bias in the *i*-th observation, and determine the threshold value at which this additional parameter becomes significant or detected by the statistical test.

Here, we present a novel application for the original reliability theory proposed by *Baarda (1968)*. In this sense, our additional parameter ∇_i in H_A is not a bias in observation, but the acceleration of a horizontal component of a GNSS CORS. Thus, our objective will be to determine *MDHA*.

The separation between the mean vector of H_0 and H_A is driven by the so-called non-centrality parameter of the model (λ) as illustrated in Fig. 2 and given as follows:

$$\lambda = \nabla_q^{\mathrm{T}} \mathbf{C}_q^{\mathrm{T}} \mathbf{W} \boldsymbol{\Sigma}_{\hat{v}} \mathbf{W} \mathbf{C}_q \nabla_q \,, \tag{6}$$

where $\mathbf{C}_q \nabla_q$ is the model error with q additional parameters ∇_q in H_A, $\mathbf{C}_q$ is a known (pre-stipulated) coefficient matrix and $\boldsymbol{\Sigma}_{\hat{v}}$ is the covariance matrix of estimated residuals in H_0. If the variance factor is unknown, then the non-centrality parameter is expressed as (*Prószyński and Lapiński, 2021*):

$$\lambda = \frac{\nabla_q^{\mathrm{T}} \mathbf{C}_q^{\mathrm{T}} \mathbf{W} \boldsymbol{\Sigma}_{\hat{v}} \mathbf{W} \mathbf{C}_q \nabla_q}{\hat{\sigma}^2} . \tag{7}$$

In our case, we have $q = 1$ (acceleration term) in H_A, and thus the non-centrality parameter is given by

$$\lambda = \nabla_i \mathbf{c}_i^{\mathrm{T}} \mathbf{W} \boldsymbol{\Sigma}_{\hat{v}} \mathbf{W} \mathbf{c}_i \nabla_i , \tag{8}$$

where

$$\mathbf{c}_i = \left[\frac{\Delta t_{o-1}^2}{2} \quad \frac{\Delta t_{o-2}^2}{2} \quad \cdots \quad \frac{\Delta t_{o-k}^2}{2} \right]^{\mathrm{T}} .$$

The reader is referred to *Teunissen (2006)* for details.

Isolating the term ∇_i results in

$$\nabla_i^2 = \lambda \left(\mathbf{c}_i^{\mathrm{T}} \mathbf{W} \boldsymbol{\Sigma}_{\hat{v}} \mathbf{W} \mathbf{c}_i \right)^{-1} \Rightarrow \left| \nabla_i \right| = \sqrt{\frac{\lambda}{\mathbf{c}_i^{\mathrm{T}} \mathbf{W} \boldsymbol{\Sigma}_{\hat{v}} \mathbf{W} \mathbf{c}_i}} . \tag{9}$$

However, the value of λ is unknown. Therefore, considering a pre-stipulated value for $\lambda = \lambda_0$ (see *Baarda, 1968*; *Teunissen, 2006*), we have the *MDHA* absolute value as follows:

$$\left| MDHA \right| = \left| \nabla_i^0 \right| = \sqrt{\frac{\lambda_0}{\mathbf{c}_i^{\mathrm{T}} \mathbf{W} \boldsymbol{\Sigma}_{\hat{v}} \mathbf{W} \mathbf{c}_i}} . \tag{10}$$

Note that Eq. (10) is identical to that used for the *MDB* of an observation. However, *Baarda's (1968)* reliability theory predates the advent of GNSS CORS time series. Thus, the reliability measure *MDHA* is a new application of the original contribution by *Baarda (1968)*, extending beyond the scope of quality control and deformation analysis of geodetic networks.

The reference non-centrality parameter of the model λ_0 is obtained as a function of the stipulated probability levels in F distribution: $\lambda_0 = f\left(q = 1, n-k, \alpha_0, \beta_0\right)$ or χ^2 distribution: $\lambda_0 = f\left(q = 1, \alpha_0, \beta_0\right)$, see, e.g., *Prószyński and Łapiński (2021)* for details. For example, if $q = 1$, $\alpha_0 = 0.05$ (5%), $\beta_0 = 0.2$ (20%), $n - k = 18$, the non-centrality parameter of the model in F distribution is given by $\lambda_0 = 8.7774$.

4. DESIGN OF EXPERIMENTS

The GNSS data was provided by the Nevada Geodetic Laboratory (NGL) at the University of Nevada, which conducts research in the field of space geodesy to study scientific problems that have both regional and global significance (*Blewitt et al., 2018*). Such data consists of files containing calculated coordinates and their respective

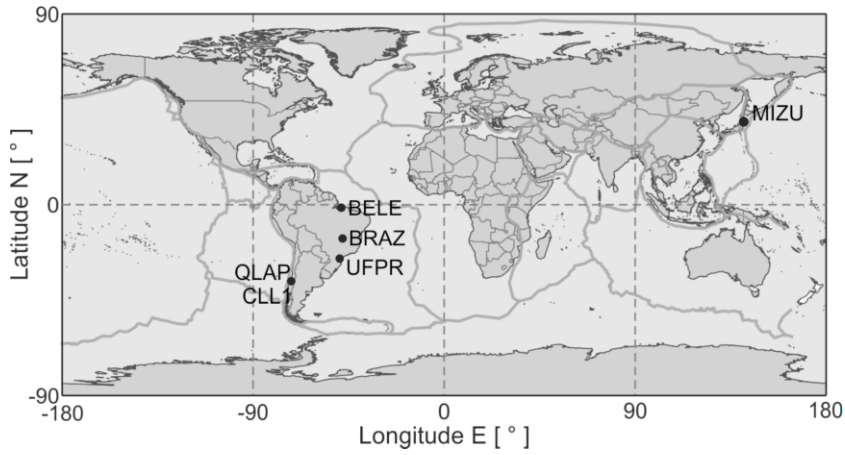

Fig. 3. Locations of the GNSS Continuously Operating Reference Stations (CORS) used in this study. Mercator projection, WGS84 datum. The thick grey curve depicts tectonic plate boundaries.

uncertainties. The NGL routinely processes the observations by using GipsyX version 1.0 software (*Bertiger et al., 2020*). The processing method used by NGL is Precise Point Positioning (PPP) and the data is tied to the IGS14 (ITRF2014). This results in the provision of daily GNSS solutions called NGL14.

The products used in the processing include (*Blewitt et al., 2018*):

- GPS satellite orbit position/velocity estimates;
- GPS satellite clock estimates;
- GPS satellite attitude parameters;
- WLPB estimates (wide lane and phase biases);
- Daily transformation parameters from NNR to IGS14;
- Time-pole parameter estimates;
- GPS satellite eclipse shadow times;;
- Name of IGS antenna calibration files

Locations of the GNSS CORS, used in this study, are shown in Fig. 3.

Figure 4a presents the displacement of the east component of station CLL1, while Fig. 4b presents the same for the north component. The east and north components of the QLAP station are shown in Fig. 5a,b, respectively.

The MIZU station exhibits different behavior due to the high incidence of seismic events. Therefore, data from 2012 onwards are used (Fig. 6), after the Tohoku earthquake with a magnitude of 9.1 in 2011. If the complete time series were to be employed, it would require post-seismic deformation modeling, which is out of scope of this study.

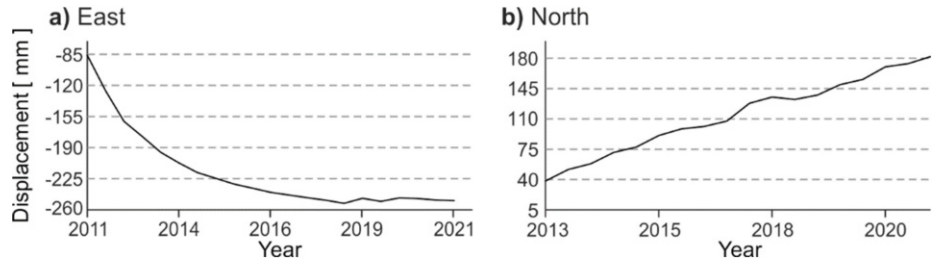

Fig. 5. The same as in Fig. 4, but for the QLAP station.

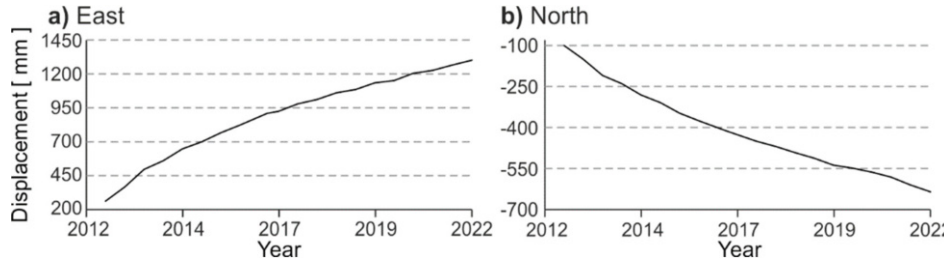

Fig. 6. The same as in Fig. 4, but for the MIZU station.

4.1. Analysis of time series by Hector software

Hector software is an open-source software developed in C++ and serves as a powerful tool for analyzing time series data obtained from GNSS stations (*Bos et al., 2013*). Based on the Maximum Likelihood Estimation (MLE), HS utilizes a robust framework that jointly estimates deterministic and stochastic models. This approach enhances the accuracy and reliability of the obtained results. Moreover, the methodology employed by HS was previously used by other software packages such as CATS (Create and Analyze Time Series) and east_noise, further attesting to its effectiveness. The obtained parameters that describe the position, seasonal signals, velocity, acceleration, and the combination of GGM (Generalised Gauss Markov) and white noise model were estimated using MLE (*Bos et al., 2013*; *Bos, 2022*).

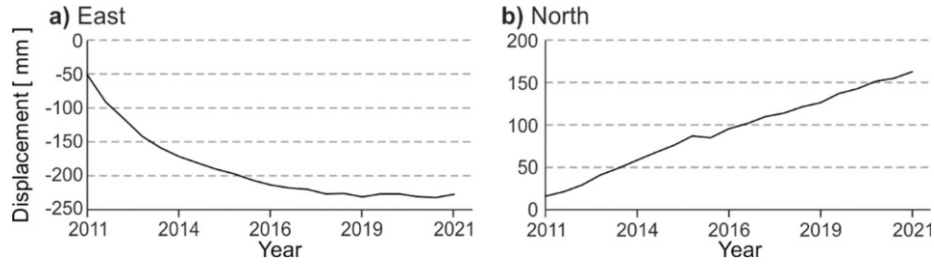

Fig. 4. Original data of **a)** east, and **b)** north component of displacement from the CLL1 station.

In general, the HS incorporates robust noise analysis methods to treat the characteristics of the noise signals. The total variance of the noise (a posteriori variance factor) is set by $\hat{\sigma}^2$, which is usually called the 'driving' noise (*Bos et al., 2013*; *He et al., 2018*; *Bos, 2022*).

For the model selection (e.g. with or without acceleration), the definition of the log-likelihood is expressed as follows:

$$\ln(L) = -\frac{1}{2}\left(n\ln(2\pi) + \ln\det(\mathbf{C}) + \hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{C}^{-1}\hat{\boldsymbol{v}} \right), \tag{11}$$

where n is the actual number of observations (gaps do not count), $\hat{\boldsymbol{v}}$ is the vector of residuals and $\mathbf{C} = \sigma^2\overline{\mathbf{C}}$ is the covariance matrix, which is a function of the sum of various noise models and the estimated (a posteriori) variance factor $\hat{\sigma}^2$, $\overline{\mathbf{C}}$ is the cofactor matrix. Details can be found in *Bos (2022)*. The number of parameters is denoted by k, which is the sum of the parameters present in the design matrix and the noise models and also the variance $\hat{\sigma}^2$ of the driving white noise process.

The HS employs two analysis models: 1) MLE supplemented with model selection using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), and 2) spectral analysis. In this research, our focus will be on the first approach, which utilizes MLE in conjunction with the AIC and the BIC. The AIC and BIC are employed to select the most suitable model from a set of competing models. The AIC was initially proposed by *Akaike (1974)* and the BIC was introduced by *Schwarz (1978)*. The methodology begins with MLE and subsequently incorporates penalties to the model to prevent overfitting.

The rule for selecting a model based on the AIC is that a smaller *AIC* value indicates a better-fitting model. Similarly, for the BIC, a smaller *BIC* value indicates a better model fit. In addition to AIC and BIC, HS offers $\mathrm{BIC_{tp}}$ and $\mathrm{BIC_c}$ as options for model selection.

The formulas for model selection using *AIC* and *BIC* are as follows (*Bos, 2022*; *Ventura et al., 2019*):

$$AIC = 2k + 2\ln(L), \tag{12}$$

$$BIC = k\ln(n) + 2\ln(L), \tag{13}$$

$$BIC_{tp} = k\ln\left(\frac{n}{2\pi}\right) + 2\ln(L), \tag{14}$$

$$BIC_c = k\ln(n)\frac{k}{n-k-2} - 2\ln(L). \tag{15}$$

The model selection process in HS involves choosing the model with the minimum value of either the *AIC* or the *BIC*. It is important to understand that these criteria are relative measures for comparing different model choices rather than absolute criteria. For details, see *Lehmann and Lösler (2016)*, *He et al. (2018)* and *Ventura et al. (2019)*.

Table 1. Statistical test and information criteria for the selection of HS model for the analyzed GNSS CORS stations.

Station	Component	Test Result	AIC	BIC	BIC_{tp}	BIC_c
CLL1	East	H_A (with acceleration)	With acceleration	With acceleration	With acceleration	With acceleration
	North	**H_0 (without acceleration)**	With acceleration	With acceleration	With acceleration	**Without acceleration**
QLAP	East	H_A (with acceleration)	With acceleration	With acceleration	With acceleration	With acceleration
	North	H_A (with acceleration)	With acceleration	With acceleration	With acceleration	With acceleration
MIZU	East	H_A (with acceleration)	With acceleration	With acceleration	With acceleration	With acceleration
	North	H_A (with acceleration)	With acceleration	With acceleration	With acceleration	With acceleration

5. RESULTS AND DISCUSSION

The results obtained by HS and the proposed approach by GLRT are presented and discussed in this section.

5.1. HS information criteria and proposed statistical test

Table 1 presents the results of the proposed GLRT (with a significance value of $\alpha = 5\%$) and the HS model selection by information criteria.

It can be seen that the model with acceleration is selected by AIC, BIC, and BIC_{tp} in all cases, but not for the north component of the CLL1 station in the BIC_c criterion. Our proposed GLRT also reveals an insignificant acceleration for the north component of the CLL1 station. Thus, these experiments showed that different criteria for model selection can lead to different results with the same dataset. In this sense, the proposed GLRT can be another criterion to be added in the HS to help the user in the important task of model selection choice.

However, which criterion is the most suitable deserves future research and thus is outside the scope of this paper. Our goal here is just to propose a new way of model selection using a statistical test rather than AIC or BIC. The proposed GLRT was also conducted at the 1% significance level, with the same test decision in all cases. The role of the significance level on the test result will be further investigated in later experiments.

5.2. MDHA analysis

Since HS does not provide the required matrices in Eq. (10), experiments about the MDHA are conducted in Scilab which is also a free and open-source software. Therefore, data were considered every 6 months for 10 years, since it is not feasible to carry out Hector's entire processing strategy in Scilab due to computational limitations. The BRAZ

and BELE stations located in Brazil and the CLL1 station located in Chile were selected. The choice of these stations is due to their locations at two extremes, with a station located on the edge of a tectonic plate, subject to constant seismic events, and other stations located in the middle of the plate. The data from CLL1, BRAZ, and BELE were obtained from the same database of the NGL.

The displacement of the east and north components of the BRAZ station is shown in Fig. 7. The displacement of the east and north components of the BELE station are presented in Fig. 8.

Table 2 shows the values for acceleration and its standard deviation, the *MDHA* (with $\alpha_0 = 5\%$ and $\beta_0 = 20\%$), and the test result for each GNSS CORS horizontal component, considering a known variance factor.

When analyzing the *MDHA* and the test results, we can observe that in horizontal components where the absolute value of the estimated acceleration is higher than the *MDHA*, the GLRT indicated that the acceleration is statistically significant. This was the case for the east and north components of CLL1 station, as well as the east component of BRAZ station and the north component of BELE station. Therefore, these results indicate that acceleration is present even in the middle of the plate.

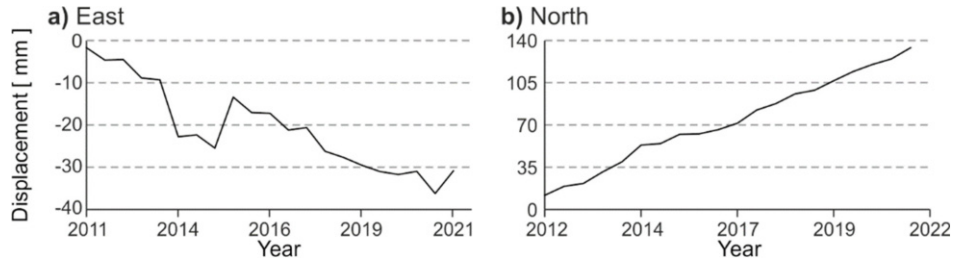

Fig. 7. Original data of **a)** east, and **b)** north component of horizontal displacement at the BRAZ GNSS CORS station.

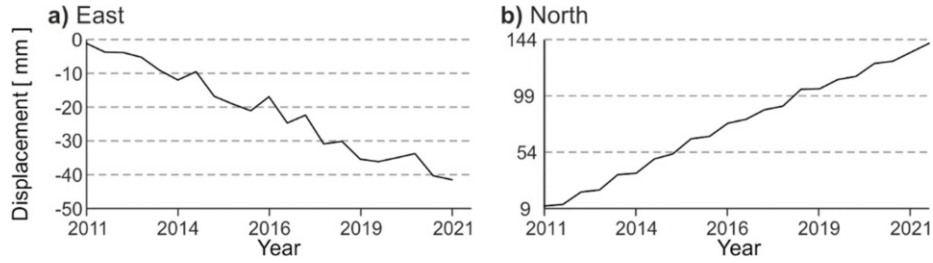

Fig. 8. The same as in Fig. 7, but for the BELE station.

Table 2. Acceleration, Minimal Detectable Horizontal Acceleration (*MDHA*), and acceleration test results for the analyzed GNSS CORS stations.

Station	Component	Acceleration [mm year^{-2}]	*MDHA* [mm year^{-2}]	Acceleration Test Result
CLL1	East	8.46 ± 0.03	0.10	Significant
	North	–1.02 ± 0.04	0.11	Significant
BRAZ	East	0.37 ± 0.03	0.09	Significant
	North	–0.08 ± 0.10	0.40	**Insignificant**
BELE	East	–0.04 ± 0.03	0.10	**Insignificant**
	North	–0.25 ± 0.03	0.09	Significant

On the other hand, when the estimated acceleration was smaller than the *MDHA*, the GLRT indicated insignificance for the acceleration. Importantly, this underscores a key advantage of the proposed statistical test over model selection criteria: the ability to derive reliability measures such as *MDHA*. It is also worth noting that when the uncertainty of acceleration approaches or exceeds its value, the test result appropriately considers it as insignificant, as expected. Additionally, note that the value of *MDHA* is independent of the observations, similar to the *MDB* values in *Baarda's (1968)* original reliability theory. In other words, the *MDHA* can serve as an important tool in the a priori analysis of GNSS CORS time series, just as *Baarda's (1968)* reliability theory is applied in the design of geodetic networks (see, e.g., *Rofatto et al. 2018*).

Although there is the possibility of considering the resultant for horizontal displacement $\sqrt{\Delta E^2 + \Delta N^2}$ (E and N stand for the east and north component, respectively), the test result would be the same using the horizontal components individually as in this research. In the case of utilizing the resultant horizontal displacements, the distribution associated with the horizontal acceleration would follow a Rayleigh distribution.

5.3. Simulation series

To further investigate issues such as the influence of data latency and the time span size on the uncertainty of the estimated acceleration, the validation of the proposed *MDHA* reliability measure, and the role of the significance level on the test results, experiments with simulated data were conducted in the Scilab software. Figure 9 presents an example of a simulated GNSS CORS time series in the east component for a time span of 10 years with a data latency of 0.01 years (≈ 3.65 days). The simulated parameters are: $v = 0.01$ m year^{-1}, $\nabla_i = 0.002$ m year^{-2}, $a_p = 0.003$ m, $b_p = 0.001$ m and $\omega_p = 4\pi$ rad. Thus, we have no jumps and only a semi-annual periodic signal. The stochastic model assumes a white noise with $\sigma_0^2 = 1\,\text{mm}^2$ for all datasets.

First, we analyze the influence of the data latency and the time spansize on the uncertainty (standard deviation) of the estimated acceleration. For this, we considered six scenarios: a 2-year time span with weekly solutions ($n = 105$), 3-day solutions ($n = 244$) and daily solutions ($n = 731$); and a 10-year time span with monthly solutions ($n = 122$), 14-day solutions ($n = 261$) and weekly solutions ($n = 522$). The standard deviations of the estimated acceleration are presented in Table 3.

It is noted that the time span is much more influential than the data latency on the uncertainty of the estimated acceleration: all scenarios with a time span of 10 years provide better results than all scenarios with a time span of 2 years, even when the total number of data (n) is smaller. For a time span of 10-year with 14-day solutions ($n = 261$), the acceleration uncertainty is 0.04 mm year^{-2}, while for a time span of 2-year with 3-days solutions ($n = 244$), the acceleration uncertainty is 0.22 mm year^{-2}. For a time span of

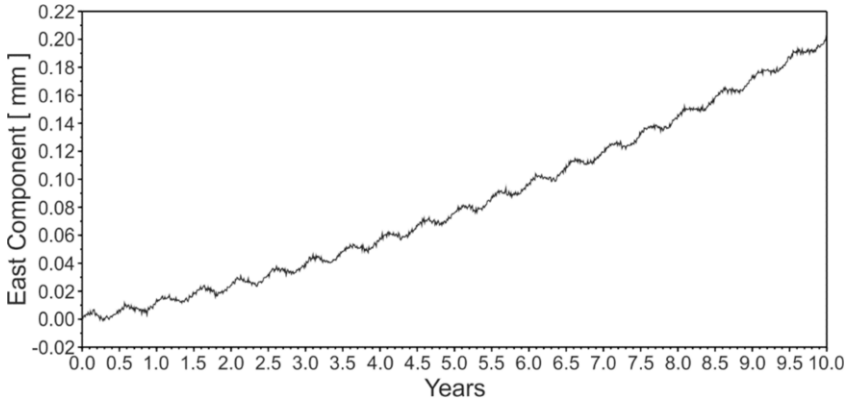

Fig. 9. Simulated time series of the east component of displacement at a GNSS CORS station.

Table 3. Standard deviations of acceleration (in mm year^{-2}) for a different time span and data latency.

Time Span	Data Latency			Time Span	Data Latency		
	Weekly	3 Days	Daily		Weekly	3 Days	Daily
2 years	0.34	0.22	0.13	10 years	0.06	0.04	0.03

10-year with monthly solutions (n = 122), the acceleration uncertainty is 0.06 mm year^{-2}, while for a time span of 2-year with weekly solutions (n = 105), the acceleration uncertainty is 0.34 mm year^{-2}. It is noteworthy that data latency is crucial for modeling cyclical effects and detecting jumps.

In a second experiment, Monte-Carlo simulations were performed to verify the empirical success rate for a given reference value for *MDHA* (see, e.g., *Rofatto et al., 2017, 2020*). Considering a known variance factor, *MDHA* = 0.02 mm year^{-2} was obtained for a time span of 10 years with a data latency of 0.01 years (≈ 3.65 days) by Eq. (10). In this case we have $\lambda_0 = f\left(q=1, \alpha_0 = 0.1\%, \beta_0 = 20\%\right) = 17.075$. Thus, random errors are generated under the alternative hypothesis with ∇_i = 0.02 mm year^{-2} and the GLRT was carried out at the significance level of α_0 = 0.1%. The MC was performed 200 000 times, and in 80.05% of the cases, the test result rejected the null hypothesis. Therefore, the empirical success rate was very close to the theoretical power of the test (γ = 100 – 20 = 80%), as expected.

This experiment confirms the validity of the *MDHA* and ensures that the simulated experiment was conducted correctly. Although the significance level adopted is somewhat arbitrary, the *MDHA* is a valuable metric for analyzing the sensitivity of the model to the test results. To the best of our knowledge, no similar metric exists for model selection by information criteria such as *AIC* or *BIC*, as applied in HS.

Finally, about the significance level adopted, Table 4 presents some critical values for both cases with known ($\chi^2_{q=1,\alpha_0}$) and unknown ($F_{q=1,n-q,\alpha_0}$) variance factor.

The critical value for α_0 = 10% is about four times lower than the critical value for α_0 = 0.1% in all cases presented. The test statistics of Eq. (4) is equivalent to

Table 4. Critical values for different significance levels and model redundancy. See text for the meaning of the variables.

α_0 [%]	$\chi^2_{q=1,\alpha_0}$	$F_{q=1,100,\alpha_0}$	$F_{q=1,500,\alpha_0}$	$F_{q=1,1000,\alpha_0}$	$F_{q=1,5000,\alpha_0}$
0.1	10.83	11.50	10.96	10.89	10.84
1.0	6.63	6.90	6.69	6.66	6.64
5.0	3.84	3.94	3.86	3.85	3.84
10.0	2.71	2.76	2.72	2.71	2.71

$$T_q = \frac{\hat{\nabla}_i^2}{\sigma^2_{\hat{\nabla}_i^2}},$$

where $\hat{\nabla}_i$ is the estimated acceleration and $\sigma_{\hat{\nabla}_i}$ its standard deviation (see, e.g., *Teunissen, 2006*). Given the high redundancy of GNSS CORS time series (in general $n - k$ >100), we can consider $\hat{\nabla}_i$ sufficiently close to ∇_i. Thus, our analysis looks for cases when

$$\frac{\hat{\nabla}_i^2}{\sigma^2_{\hat{\nabla}_i^2}} > \chi^2_{q=1\alpha_0}, \text{ or } \frac{\hat{\nabla}_i^2}{\sigma^2_{\hat{\nabla}_i^2}} > F_{q=1,n-q,\alpha_0}.$$

This yields

$$\hat{\nabla}_i > \sigma_{\hat{\nabla}_i}\sqrt{\chi^2_{q=1.\alpha_0}} \text{ or } \hat{\nabla}_i > \sigma_{\hat{\nabla}_i}\sqrt{F_{q=1,n-k,\alpha_0}}.$$

Considering the values for $\sigma_{\hat{\nabla}_i}$ in Table 3 and the critical values in Table 4, we present the values of $\hat{\nabla}_i \approx \nabla_i > \sigma_{\hat{\nabla}_i}\sqrt{\chi^2_{q=1.\alpha_0}}$ or $\hat{\nabla}_i \approx \nabla_i > \sigma_{\hat{\nabla}_i}\sqrt{F_{q=1,n-k,\alpha_0}}$ that lead to an expected rejection of H_0 in Table 5. In this case, we approximate the critical values $\chi^2_{q=1.\alpha_0}$ and $F_{q=1,n-k,\alpha_0}$ for $\alpha_0 = 0.1\%$ as ≈ 11, for $\alpha_0 = 1\%$ as ≈ 6.7, for $\alpha_0 = 5\%$ as ≈ 3.86 and for $\alpha_0 = 10\%$ as ≈ 2.72 without loss of generality in our conclusions.

It is noted that in general acceleration values lower than 1 mm year^{-2} are sufficient to reject H_0. The only exception was for a time span of 2 years with weekly solutions and a significance level of $\alpha_0 = 0.1\%$, a case that rarely occurs in practice. *Blewitt and Lavallée (2002)* do not recommend time spans of less than 2.5 years in this kind of analysis. For a time span of 10 years, in general, an acceleration of the order of just 0.1 mm year^{-2} would be enough to reject H_0. Considering that usually, the GNSS CORS data latency is daily, we can state that acceleration values lower than 0.1 mm year^{-2} can lead to the rejection of H_0, even for a low significance level such as $\alpha_0 = 0.1\%$.

Finally, Table 6 presents the estimated acceleration values in HS for some GNSS CORS stations with daily data latency and in regions subject to different geodynamic activities.

Table 5. Expected values of acceleration in mm year^{-2} that lead to a rejection of H_0.

Time Span	α_0 [%]	Data Latency			Time Span	α_0 [%]	Data Latency		
		Weekly	3 Days	Daily			Weekly	3 Days	Daily
2 years	0.1	1.12	0.73	0.43	10 years	0.1	0.20	0.13	0.10
	1.0	0.88	0.57	0.34		1.0	0.16	0.10	0.78
	5.0	0.67	0.43	0.25		5.0	0.12	0.08	0.06
	10.0	0.56	0.36	0.21		10.0	0.10	0.07	0.05

Table 6. Time span and estimated acceleration in HS for analyzed GNSS CORS stations.

Station	Time Span [days]	Component	Acceleration [mm year^{-2}]	Incidence of Seismic Events
CLL1	4736	East North	10.34 ± 20644 –0.167 ± 0.053	High
QLAP	4750	East North	20.36 ± 0.046 –0.025 ± 0.024	High
MIZU	4123	East North	–6.770 ± 0.569 2.469 ± 0.235	High
BELE	7017	East North	0.035 ± 0.013 –0.012 ± 0.012	Low
BRAZ	10262	East North	0.042 ± 0.023 0.005 ± 0.009	Low
UFPR	9591	East North	–0.008 ± 0.019 2.096 ± 2.875	Low

It is evident from Tables 5 and 6 that the choice of the significance level should affect only GNSS CORS in regions with low geodynamic activities and/or short time span. However, even in these cases, the proposed MDHA reliability measure remains a very useful tool for analyzing the estimated acceleration and verifying whether the GNSS CORS time series has the desired sensitivity to the phenomenon under observation, considering a pre-stipulated false negative rate. In this case, the MDHA value can be compared with a certain reference value, according to the geophysical phenomenon investigated. This presents an interesting open issue for future work.

6. CONCLUSIONS

This study follows current analyses involving the inclusion of nonlinear trends in the horizontal components of GNSS CORS time series. We presented a statistical test (GLRT) to the model selection problem: with or without acceleration. The proposed GLRT is relevant because the nonlinear long-term trend may be significant even for GNSS CORS stations located in the middle of the tectonic plate, as demonstrated by the results.

Furthermore, we proposed a new application for *Baarda's (1968)* reliability theory to estimate the Minimal Detectable Horizontal Acceleration (*MDHA*), which cannot be derived from model selection by information criteria such as *AIC* or *BIC*. The experiments demonstrated a good agreement between the test results and the *MDHA*, as expected.

Model selection by using information criteria such as *AIC* or *BIC* involves a subjective choice for the penalty function related to the number of parameters (see Eqs (12)–(15)). The proposed statistical test has a subjective choice for the significance level. However, the proposed *MDHA* is a valuable metric to analyze the sensitivity of the model to the test result. To the best of our knowledge, no similar metric exists for model selection by information criterion such as *AIC* or *BIC*, as applied in HS. Besides that, the *MDHA* value can be

compared with a certain reference value corresponding to the geophysical phenomenon under investigation.

Additionally, the *MDHA* value is independent of the vector of observations, such as the *MDB* values of *Baarda (1968)*. Therefore, the *MDHA* value is a very useful tool to analyze whether the GNSS CORS time series has the desired sensitivity to the geophysical phenomenon under investigation. This is an interesting open issue for future works.

We also suggest implementing the proposed statistical test and *MDHA* in software packages such as HS. We recommend an investigation aimed at developing a unified approach that can be tested for all model parameters. Furthermore, this unified approach could address the multivariate distribution while considering the correlation between horizontal components.

References

Akaike H., 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control*, **19**, 716–723

Altamimi Z., Rebischung P., Métivier L. and Collilieux X., 2016. ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions. *J. Geophys. Res.-Solid Earth*, **121**, 6109–6131

Altamimi Z., Rebischung P., Collilieux X., Métivier L. and Chanard K., 2023. ITRF2020: an augmented reference frame refining the modeling of nonlinear station motions. *J. Geodesy*, **97**, Art.No. 47, https://doi.org/10.1007/s00190-023-01738-w

Baarda W., 1968. *A Testing Procedure for Use in Geodetic Networks*. Publications on Geodesy, **2(5)**, Netherlands Geodetic Commission, Delft, The Netherlands

Bertiger W., Bar-Sever Y., Dorsey A., Haines B., Harvey N., Hemberger D., Heflin M., Lu W., Miller M., Moore A. W., Murphy D., Ries P., Romans L., Sibois A., Sibthorpe A., Szilagyi B., Vallisneri M. and Willis P., 2020. GipsyX/RTGx, a new tool set for space geodetic operations and research. *Adv. Space Res.*, **66**, 469–489

Bevis M. and Brown A., 2014. Trajectory models and reference frames for crustal motion geodesy. *J. Geodesy*, **88**, 283–311

Blewitt G., Hammond W.C. and Kreemer C., 2018. Harnessing the GPS data explosion for interdisciplinary science. *Eos*, **99**, https://doi.org/10.1029/2018EO104623

Blewitt G. and Lavallée D., 2002. Effect of annual signals on geodetic velocity. *J. Geophys. Res.-Solid Earth*, **107**, Art.No. 2145, https://doi.org/10.1029/2001JB000570

Bos M.S., 2022. Hector user manual version 2.1. https://teromovigo.com/hector/

Bos M.S., Montillet J.P., Williams S.D.P. and Fernandes R.M.S., 2020. Introduction to geodetic time series analysis. In: Montillet J.P. and Bos M.S. (Eds), *Geodetic Time Series Analysis in Earth Sciences*. Springer International Publishing, Cham, Switzerland, 29–52

Bos M.S., Fernandes R.M.S., Williams S.D.P. and Bastos L., 2013. Fast error analysis of continuous GNSS observations with missing data. *J. Geodesy*, **87**, 351–360

Bogusz J., Klos A., Bos M.S., Hunegnaw A. and Teferle N.F., 2016a. On the impact of a quadratic acceleration term in the analysis of position time series. Abstract. *Geophys. Res. Abs.*, **18**, EGU2016-16816

Bogusz J., Klos A., Gruszczynska M. and Gruszczynski M., 2016b. Towards reliable velocities of permanent GNSS stations. *Rep. Geod. Geoinformatics*, **100**, 17–26

Dong D., Fang P., Bock Y., Cheng M.K. and Miyazaki S., 2002. Anatomy of apparent seasonal variations from GPS-derived site position time series. *J. Geophys. Res.-Solid Earth*, **107**, Art.No. 2075, https://doi.org/10.1029/2001JB000573

Durdag U.M., Hekimoglu S. and Erdogan B., 2018. Reliability of models in kinematic deformation analysis. *J. Surv. Eng.*, **144**, Art.No. 04018004, https://doi.org/10.1061/(ASCE)SU.1943-5428.0000254

He Y., Zhang S., Wang Q., Liu Q., Qu W. and Ho U.X., 2018. HECTOR for analysis of GPS time series. In: Sun J., Yang C. and Guo S. (Eds), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*. CSNC 2018. Lecture Notes in Electrical Engineering, **497**. Springer, Singapore, 10.1007/978-981-13-0005-9_16

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

Klos A., Bos M.S. and Bogusz J., 2018. Detecting time-varying seasonal signal in GPS position time series with different noise levels. *GPS Solut.*, **22**, Art.No. 21, https://doi.org/10.1007/s10291-017-0686-6

Kreemer C., Blewitt G. and Klein C.E., 2014. A geodetic plate motion and Global Strain Rate Model, *Geochem. Geophys. Geosyst.*, **15**, 3849–3889

Kuusniemi H. and Lachapelle G., 2004. GNSS signal reliability testing in urban and indoor environments. In: *Proceedings of the 2004 National Technical Meeting of the Institute of Navigation*, 210–224

Langbein J., Murray R.J. and Snyder A.H., 2006. Coseismic and initial postseismic deformation from the 2004 Parkfield, California, earthquake, observed by Global Positioning System, electronic distance meter, creepmeters, and borehole strainmeters. *Bull. Seismol. Soc. Amer.*, **96**, S304–S320

Lehmann R. and Lösler M., 2016. Multiple outlier detection: hypothesis tests versus model selection by information criteria. *J. Surv. Eng.*, **142**, Art.No. UNSP04016017, https://doi.org/10.1061/(ASCE)SU.1943-5428.0000189

Montillet J.P. and Bos M.S. (Eds), 2020. *Geodetic Time Series Analysis in Earth Sciences*. Springer International Publishing, Cham, Switzerland

Prószyński W. and Lapiński S., 2021. Investigating support by minimal detectable displacement in confidence region determination and significance test of displacements. *J. Geodesy*, **95**, Art.No. 112, https://doi.org/10.1007/s00190-021-01550-4

Ren Y., Lian L. and Wang J., 2021. Analysis of seismic deformation from global three-decade GNSS displacements: implications for a three-dimensional earth GNSS velocity field. *Remote Sens.*, **13**, Art.No. 3369, https://doi.org/10.3390/rs13173369

Rofatto V.F., Matsuoka M.T., Klein I., Veronez M.R., Bonimani M.L. and Lehmann R., 2020. A half-century of Baarda's concept of reliability: a review, new perspectives, and applications. *Surv. Rev.*, **52**, 261–277, https://doi.org/10.1080/00396265.2018.1548118

Rofatto V.F., Matsuoka M.T. and Klein I., 2017. An attempt to analyse Baarda's iterative data snooping procedure based on Monte Carlo simulation. *South Afr. J. Geomat.*, **6**, 416–435, https://doi.org/10.4314/sajg.v6i3.11

Rofatto V.F., Matsuoka M.T. and Klein I., 2018. Design of geodetic networks based on outlier identification criteria: an example applied to the leveling network. *Bol. Cienc. Geod.*, **24**, 152−170, https://doi.org/10.1590/S1982-21702018000200011

Schwarz G., 1978. Estimating the dimension of a model. *Ann. Stat.*, **6**, 461−464

Teunissen P.J.G., 2006. *Testing Theory: an Introduction*. 2nd Edition.VSSD, Delft, The Netherlands, ISBN: 978-90-407-1975-2

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

Ventura M., Saulo H., Leiva V. and Monsueto S., 2019. Log-symmetric regression models: information criteria and application to movie business and industry data with economic implications. *Appl. Stoch. Models. Bus. Ind.*, **35**, 963−977, https://doi.org/10.1002/asmb.2433

Zaminpardaz S., Teunissen P.J.G. and Tiberius C.C.J.M., 2020. A risk evaluation method for deformation monitoring systems. *J. Geodesy*, **94**, Art.No. 28, https://doi.org/10.1007/s00190-020-01356-w