UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA CARTOGRÁFICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS GEODÉSICAS E TECNOLOGIAS DA GEOINFORMAÇÃO

MARCELLA FERNANDES DE OLIVEIRA MELO

ASSESSING TERRESTRIAL WATER STORAGE ANOMALIES UNCERTAINTIES IN SOUTH AMERICA USING THE THREE-CORNERED HAT METHOD

Recife
2023

MARCELLA FERNANDES DE OLIVEIRA MELO

ASSESSING TERRESTRIAL WATER STORAGE ANOMALIES UNCERTAINTIES IN SOUTH AMERICA USING THE THREE-CORNERED HAT METHOD

Dissertation presented to the Graduate Program in Ciências Geodésicas e Tecnologias da Geoinformação at the Universidade Federal de Pernambuco, as a partial requirement for obtaining the Mestre em Ciências Geodésicas e Tecnologia da Geoinformação.

Concentration area: Geodetic Sciences and Geoinformation Technologies.

Advisor: Prof. Dr. Vagner Gonçalves Ferreira.

Recife

2023

Catalogação na fonte:
Bibliotecária Sandra Maria Neri Santiago CRB-4 / 1267

M528a Melo, Marcella Fernandes de Oliveira.
Assessing terrestrial water storage anomalies uncertainties in south america using the three-cornered hat method / Marcella Fernandes de Oliveira Melo. – 2023.
39 f. : il., fig., tab., abrev. e siglas.

Orientador: Prof. Dr. Vagner Gonçalves Ferreira.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-graduação em Ciências Geodésicas e Tecnologias da Geoinformação. Recife, 2023.
Inclui referências.

1. Ciências geodésicas. 2. GRACE. 3. Mascons. 4. Método-do-chapéu-de-três-pontas. 5. TWSA. 6. Incertezas. I. Ferreira, Vagner Gonçalves (Orientador). II. Título.

UFPE

526.1 CDD (22. ed.) BCTG/2023-81

MARCELLA FERNANDES DE OLIVEIRA MELO

ASSESSING TERRESTRIAL WATER STORAGE ANOMALIES UNCERTAINTIES IN SOUTH AMERICA USING THE THREE-CORNERED HAT METHOD

Dissertation presented to the Graduate Program in Ciências Geodésicas e Tecnologias da Geoinformação at the Universidade Federal de Pernambuco, as a partial requirement for obtaining the Mestre em Ciências Geodésicas e Tecnologia da Geoinformação.

Concentration area: Geodetic Sciences and Geoinformation Technologies.

Approved on: 10/04/2023.

EXAMINATION BOARD

_____

Prof. Dr. Vagner Gonçalves Ferreira (Advisor)
Universidade Federal de Pernambuco e Universidade de Hohai

_____

Prof. Dr. Henry D. Montecino C. (Internal Examiner)
Universidade Federal de Pernambuco e Universidad de Concepción

_____

Prof. Dr. Regiane Dalazoana (External Examiner)
Universidade Federal do Paraná

ABSTRACT

The GRACE mission (Gravity Recovery and Climate Experiment) and its successor, the GRACE Follow-On (GRACE-FO) have been monitoring the terrestrial water storage anomalies (TWSA) since April 2002. The Space Research Center (CSR), Jet Propulsion Laboratory (JPL), and Goddard Space Flights (GSFC) are operationally providing individual monthly solutions based on the mass concentration (mascon) solution as the radial basis function. The inverted TWSA maps (or regional averaged values) from these centers are being used in many applications. However, as terrestrial data are unavailable, the uncertainties and their confidence intervals are unknown. Consequently, this work aims to evaluate the quality of each TWSA solution from the three processing centers using a generalized formulation of the three-cornered hat method (TCH) and estimate the respective confidence intervals for the uncertainties. Overall, the TCH results for the study period from April 2002 to June 2017 over South America indicate that the uncertainties of TWSA from CSR, JPL, and GSFC are 34.4, 47.4, and 38.7 mm, respectively. At basin scale, the good performance of the TWSA based on CSR mascon solution is observed. The THC-based results agree with an ensemble mean from the three solutions. Regarding the confidence intervals, in general, the reliability of the data decreases as the area decreases. In this sense, this study supports hydrologists in selecting the best TWSA solution for hydrological studies.

Keywords: GRACE; mascons; three-cornered hat method; TWSA; uncertainties.

RESUMO

A missão GRACE (Gravity Recovery and Climate Experiment) e sua sucessora, a GRACE Follow-On (GRACE-FO) têm monitorado as anomalias de armazenamento d'água terrestre (do inglês, terrestrial water storage anomaly - TWSA) desde abril de 2002. O Centro de Pesquisa Espacial (CSR) da Universidade do Texas, o Laboratório de Propulsão a Jato (JPL), e o Centro de Voos Espaciais Goddard (GSFC) estão operacionalmente fornecendo grades mensais contendo TWSA baseadas nas funções de base radial empregando-se a solução concentração de massa (do inglês, mass concentration - mascon). As grades contendo TWSA (ou valores médios regionais) produzidas por estes centros estão sendo usadas em muitas aplicações. No entanto, como observações terrestres não estão disponíveis, as incertezas e seus intervalos de confiança são desconhecidos. Consequentemente, este trabalho visa avaliar a precisão de cada solução TWSA dos três centros de processamento usando uma formulação generalizada do método do chapéu-de-três-pontas (do inglês, three-cornered hat - TCH) e estimar os respectivos intervalos de confiança para as incertezas. No geral, os resultados do método TCH para o período de estudo de abril de 2002 a junho de 2017 indicam que as incertezas das grades TWSA para a América do Sul provenientes do CSR, JPL e GSFC são 34,4, 47,4 e 38,7 mm, respectivamente. Para a média regional considerando as bacias hidrográficas, observou-se o bom desempenho da solução TWSA produzida pelo CSR. As incertezas baseadas no método THC concordam com as incertezas considerando média de conjunto para as três soluções como referência. Com relação aos intervalos de confiança, em geral, a confiabilidade dos dados diminui à medida que a área diminui. Neste sentido, esta dissertação auxilia hidrologistas a selecionarem a melhor solução TWSA para estudo hidrológicos.

Palavras-chave: GRACE; mascons; método-do-chapéu-de-três-pontas; TWSA; incertezas.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

CSR	Center for Space Research
GRACE	Gravity Recovery and Climate Experiment
GRACE-FO	Gravity Recovery and Climate Experiment Follow-On
GSFC	Goddard Space Flight Center
JPL	Jet Propulsion Laboratory
mascon	mass concentration
SNR	Signal-to-Noise Ratio
TCH	Three-Cornered Hat
TWSA	Terrestrial Water Storage Anomalies
ubRMSE	unbiased Root Mean Squared Error

SUMMARY

1 INTRODUCTION

Since March 17, 2002, the Gravity Recovery and Climate Experiment (GRACE) mission and its successor, the GRACE Follow-On (GRACE-FO), have been used to monitor Earth's gravitational field in the space and time domains. The variability in Earth's gravity field as observed by GRACE, represents geophysical responses associated with mass distributions and redistribution at or near the Earth's surface. The mass variations are likely to occur at the Earth's surface on the time scales examined by GRACE measurements. Generally, the largest time-varying gravity signals observable in GRACE data are expected to come from changes in the distribution of water and stored snow on continents (WAHR et al., 1998, WAHR et al., 2009). Consequently, temporal variations in the Earth's gravitational field can be used to estimate global and regional terrestrial water storage anomalies (TWSA) (cf., WAHR et al., 2004).

GRACE and GRACE-FO missions significantly expanded the frontiers of satellite geodesy applications in hydrology. For two decades, the GRACE mission has been enabling the hydrological community to explore how TWS evolves in space and time (TAPLEY et al., 2019), analyze the water balance (HASSAN and JIN, 2016), investigate groundwater exploitation, flooding, and drought severity (RODELL et al., 2007; LONG et al., 2014; REAGER et al., 2014; FRAPPART and RAMILLIEN, 2018; HASAN et al., 2021), estimate regional evapotranspiration (RODELL et al., 2004), and precipitation minus evapotranspiration (SWENSON and WAHR, 2006), and study lake/reservoir storage dynamics (MOORE and WILLIAMS, 2014), among other applications. In the context of geodesy, GRACE missions allow geodesists to model geoid height, glacial isostatic adjustment, hydrological loading (VAN DAM et al., 2007), and glacial mass variations (PAULSON et al., 2007; GERUO et al., 2013; SUTTERLEY et al., 2014). These applications are possible due to an ever-increasing range of new products and data, and continuous improvements in existing processing and analysis methodologies produced by the GRACE processing centers.

The TWSA fields, expressed as equivalent water height, can be estimated from the spherical harmonics coefficients or based on the mass concentration (mascon) solution. The spherical harmonic coefficients are available through different parameterizations, and their inversion to TWSA requires different filtering strategies, smoothing, and scaling factors (cf. LANDERER and SWENSON, 2012; LONG et al., 2015). Alternatively, the Center for Space Research (CSR) at the University of Texas, the Jet Propulsion Laboratory (JPL), and the

Goddard Space Flight Center (GSFC) are producing direct TWSA solutions by inverting the GRACE raw data based on the mascon solutions. TWSA solutions based on mascons have the advantages of retaining enhanced gravity signals using regional constraints, reducing residual noise, and minimizing spatial leakage error. In most cases, mascons solutions can be used directly without applying scaling factors or post-processing schemes (cf. WATKINS et al., 2015). Consequently, mascon solutions present advantages to the hydrological community. However, evaluation of available mascons solutions is still necessary to assist hydrologists in selecting the best product and their limitations in terms of the noise level.

Error estimates for satellite-based products require validation using ground datasets. Nevertheless, assessing the uncertainties is challenging without ground data, especially for regions such as South America, where data is limited or nonexistent. This situation is even worse for the TWSA, where no direct measurements exist, since it is virtually the sum of soil moisture, inland waters, groundwater, and ice and snow storage. Noise decoupling problems, such as the three-cornered hat (TCH) method, offer an approach to assess the uncertainty of each GRACE processing center without the need for ground measurements (FERREIRA et al., 2016). Efforts were made to evaluate the GRACE uncertainties using the TCH method. For example, Ferreira et al. (2016) evaluated the uncertainties of four GRACE processing centers on a global scale using spherical harmonic solutions and found results between 9.4 and 14.8 mm. Furthermore, Yao et al. (2019) analyzed the uncertainties related to TWSA changes derived from GRACE of five spherical harmonic solutions over mainland China and the results showed that the average uncertainties of TWSA changes over mainland China are between 15 to 56 mm. However, the confidence interval for the errors estimated by the TCH method was not presented in previous works, as well as the use of mascons solutions. In this study, the generalized formulation of the TCH method is applied to evaluate the different TWSA solutions produced by three mascon solutions. Specifically, the main objectives of this study are:

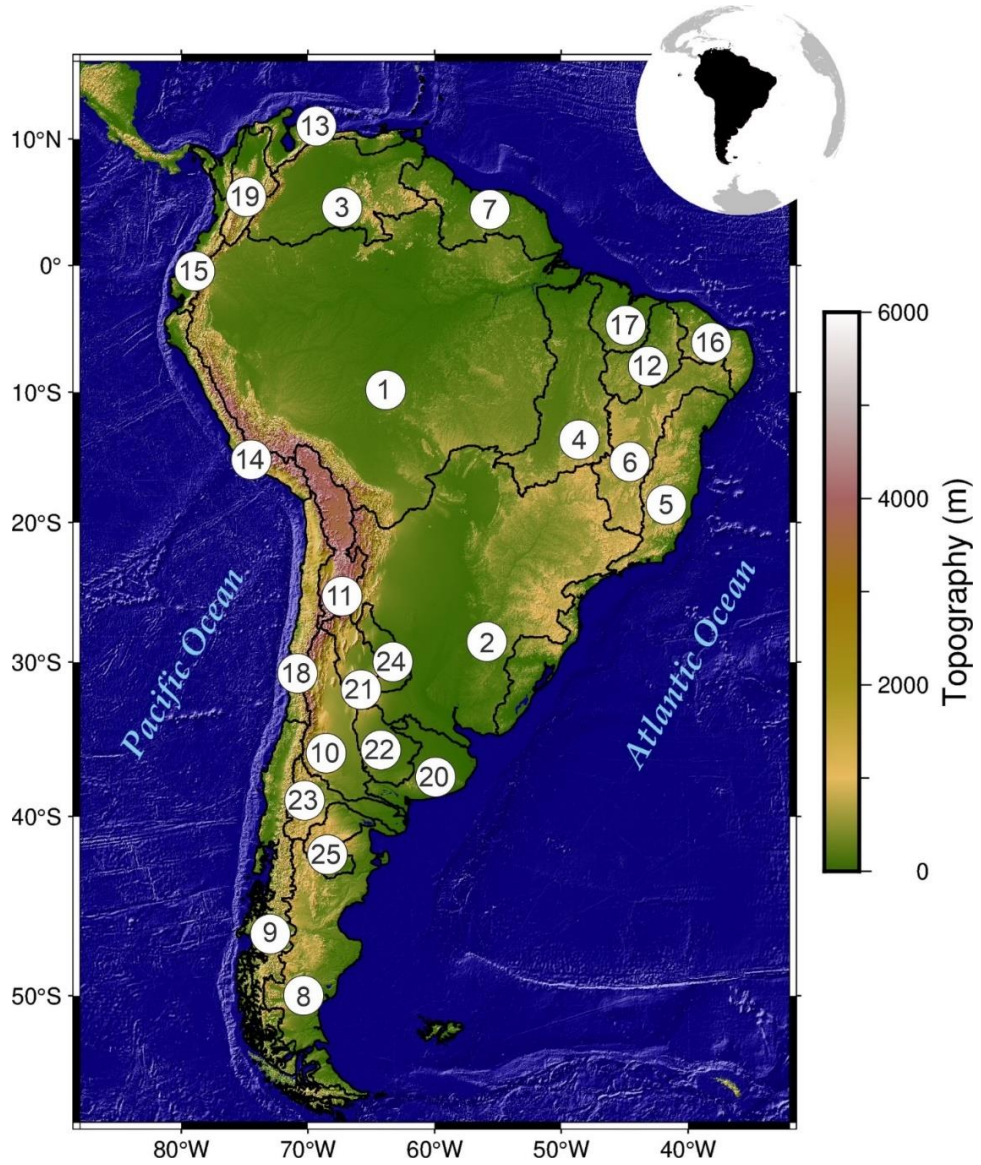
(i) to calculate the validation metrics for each of the three mascon solutions (CSR, JPL, and GSFC), and

(ii) to estimate their respective confidence intervals.

2 STUDY AREA

The study area is composed of South American watersheds, which comprise a total of 25 basins with areas equal to or larger than 100,000 km^2 (Figure 1). This region contains the world's largest watershed (the Amazon Basin) and drier regions, such as eastern Patagonia and the extremely arid Atacama Desert. In addition, it has two other important hydrographic basins, La Plata and Orinoco, the two largest and most productive hydrographic basins on the continent.

Figure 1 - Watersheds of South America extracted from the South American Hydrological Basins, available in FAO (2006). The numbers assigned to each catchment here can be used to identify the catchment names (Table 2).



Source: Author (2023).

The South American continent, with unique ecosystems and high biodiversity, presents extreme geographic variations and diverse weather and climate patterns, including tropical, subtropical, and extratropical characteristics (GARREAUD et al., 2009). The continent is responsible for approximately 20% of global freshwater discharge (DE LINAGE et al., 2013), thus crucial in the global biogeochemical cycle. The region is largely under the influence of large-scale atmospheric-oceanic phenomena, including mainly the El Niño Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO), which affect the climate and its phases associated with droughts, floods, and extreme weather events in different parts of the continent (MAGRIN et al., 2007, TEDESCHI and COLLINS, 2016).

Climate variability across South America can be categorized based on distance from the equator and the altitude of the area. These climate variabilities have significant impacts on the continent's water storage. Other important essential factors such as excessive water use, especially for agricultural purposes, widely threaten water resources (GRAU and AIDE, 2008, MAGRIN et al., 2014). Therefore, considering the diversity of South America presented and the limitations of terrestrial observations, it becomes interesting to use it for this study.

3 MATERIAL AND METHODS

This section describes the dataset as well as the methods used in this study.

3.1 Datasets

This topic aims to present the datasets, divided into (i) GRACE mascon solutions and (ii) pre-processing of GRACE mascon solution.

3.1.1 GRACE mascon solutions

The TWSA can be described as a functional of the Earth's time-varying gravitational field, related to hydrological applications, and represents the sum of water stored in the soil, water contained in glaciers, surface and underground runoff and precipitation for a studied area or basin (e.g., AHI and CEKIM, 2021). TWSA products can be estimated based on mascons solutions. The mascons solutions represent changes in the Earth's gravitational field, which can be defined as blocks of mass concentration in a sphere (MULLER and SJÖGREN, 1968). Mascons have some advantages over spherical harmonic functions. First, the geophysical constraints applied support noise filtering. This is more rigorous than the empirical methods used to remove the north-south stripes present in spherical harmonic solutions. In addition to better noise filtering, mascons also prevent information leakage between land and ocean, leading to a better representation of coastal zones (SAVE et al., 2016). The mascon blocks are provided by the GRACE mission's CSR, JPL and GSFC processing centers. Mascon solutions provided by these individual processing centers differ by processing variants and sizes, shapes, and spatial resolution of the blocks on the Earth’s surface (generally taken as a spherical approximation).

In this study, mascon GRACE solutions were used in their most recent version labeled as RL06 (Release-06) provided in the form of monthly grided TWSA values by CSR (SAVE et al., 2016; SAVE, 2020), JPL (WATKINS et al., 2015; WIESE et al., 2016; WIESE et al., 2018; LANDERER et al., 2020), and GSFC (LOOMIS et al., 2019). Table 1 presents the main processing techniques and spatial resolution of blocks in a sphere used by each GRACE center.

Table 1 - Main processing techniques and spatial resolution of each GRACE center.

Centers	**Replacement of C20 Coefficients**	**Reduced effect of Glacial Isostatic Adjustment (GIA)**	**Noise reduction**	**spatial resolution**
CSR	Satellite Laser Range (SLR) Estimates	Using the ICE6G_D model (PELTIER et al., 2018)	Application of restrictions based only on GRACE observations (SAVE et al., 2016)	0.25°
JPL	Satellite Laser Range (SLR) Estimates	Using the ICE6G_D model (PELTIER et al., 2018)	Coastline Resolution Improvement (CRI) method (WATKINS et al., 2015)	0.5°
GSFC	Satellite Laser Range (SLR) Estimates	Using the ICE6G_D model (PELTIER et al., 2018)	Adaptive decomposition filter using anisotropic constraints (LUTHCKE et al., 2013)	0.5°

Source: Author (2023).

3.1.2 Pre-processing of GRACE mascon solutions

For the evaluation of uncertainties, it is initially necessary to perform a pre-processing step in order to find correspondences in space and time between estimates that have different spatial resolutions since the datasets are sampled in different grids and/or acquired at different times (GRUBER et al., 2020). The process of finding matches between these data is commonly referred to as collocation, being essentially a resampling task (LOEW et al., 2017). This step is used to minimize the impact of representativeness errors (primarily spatial) on validation metrics (GRUBER et al., 2020).

As the spatial resolutions of the gridded TWSA products differ from each other, a reference grid must be selected on which the other products are resampled for collocation purposes. For this investigation, a spatial resolution of 0.25° was chosen as a reference, this being the resolution of the CSR solution and the resolution of the JPL and GSFC centers were resampled from 0.5° to 0.25° using the conservation of mass (see, WIESE et al., 2016) in order to promote the reduction of leakage errors along coastlines. It is important to note that

the resampling consisted only of dividing the cell into other smaller cells under the condition that the volume must be conserved. So, no mathematical manipulation was used.

Subsequently, the GRACE-derived monthly TWS fields were re-sampled precisely in the middle of each month, and the missing 20-month data were interpolated using data from neighboring months as suggested by Mo et al. (2022). This was necessary as the number of days differs by a few days between GRACE processing centers, and all datasets must be time-aligned to use the TCH method (FERREIRA et al., 2016). Monthly TWSA observations during the interval between GRACE and GRACE-FO are absent, leading to discontinuity in the time series and therefore preventing full utilization and analysis of the data (SUN et al., 2021, YI and SNEEUW, 2021). Hence, due to this 11-month GAP between the GRACE mission and GRACE-FO, the data used are only for the first mission, ranging from April 2002 to June 2017.

3.2 Methods

This topic aims to present the methods used in this study, divided into (i) TWSA time series processing, (ii) Gaussian filtering, (iii) The three-cornered hat method, (iv) Validation metrics and (v) Confidence Intervals from block-bootstrapping.

3.2.1 TWSA time series processing

To evaluate the TWSA time series derived from GRACE on a grid scale, the annual amplitude and linear trend were calculated using the following functional model:

$$y(t) = a_0 + a_1(t - t_0) + \sum_{k=1}^{n_f} (S_k \sin \omega_k t + C_k \cos \omega_k t) \quad (1)$$

$y(t)$ is the original input series as a function of time, a_0 is the value of y at the reference epoch t_0, a_1 is the linear trend, ω_k is the angular frequency ($\omega_k = 2\pi/T_k$, where T_k is the period , one year in our analyses), and the coefficients S_k and C_k contain information on the annual amplitude A_k, where it is calculated by:

$$A_k = \sqrt{S_k^2 + C_k^2} \tag{2}$$

3.2.2 Gaussian filtering

In order to make comparisons between the GRACE processing centers, a spatial smoothing of the observed amplitudes is performed. Here, it is proposed to average the annual amplitudes and linear trends using the Gaussian smoothing function, a low-pass filter, to reduce the signal strength. The weighted Gaussian mean is calculated by the expression (BOMFIM et al. 2013):

$$\overline{\Delta r} = \frac{\sum_i \Delta r_i W(\psi_i)}{\sum_i W(\psi_i)} \tag{3}$$

where the Gaussian operator $W(\psi_i)$ is a function of the spherical distance, ψ, between the computation point and the execution point, calculated by Wahr et al. (1998) as:

$$W(\psi) = \frac{2be^{-b(1-\cos\psi)}}{1 - e^{-2b}} \tag{4}$$

where b is dimensionless and defines the smoothing process of the Gaussian operator, which is calculated as (WAHR et al. 1998):

$$b = \frac{\ln(2)}{1 - \cos(\frac{r}{R})} \tag{5}$$

where, r is the average radius (or smoothing radius), which is equivalent to the distance on Earth at which W drops to half its value at the origin (ψ = 0°), in this case, the smoothing radius used was 150 km, and R is the radius of a spherical Earth (6378 km).

3.2.3 The three-cornered hat method

The TCH method was initially proposed to estimate the relative accuracy of oscillators and atomic clocks (GRAY and ALLAN, 1974) but has been widely used in geodetic studies (CHIN et al., 2005; KOOT et al., 2006; VALTY et al., 2013). The TCH is applicable when at least three observations of the same quantity obtained by different sensors are available. The presupposed hypothesis is that the observations contain the same signal in common, that is,

the true signal, plus independent noise responsible for the observed differences. The method consists of introducing covariances of the measured data, with the proposal of an adequate optimization criterion, allowing estimation of the complete covariance matrix of the three measures, provided that its positive definition is guaranteed (PREMOLI and TAVELLA, 1993).

Regarding the GRACE TWS variable, in the absence of a reference dataset, the TCH method can be used to estimate the relative uncertainties of the TWS derived from different sources if at least three products are available. To estimate the uncertainty in the TWS datasets, consider the time series of stored available products as $\{X_i\}_{i=1,2,\dots,N}$ where i corresponds to each solution center (that is, $N = 3$, corresponding to the GRACE processing centers that produce mascons solutions). Let each time series be expressed as:

$$X_i = S + \varepsilon_i, \quad \forall\, i = 1, \dots, N, \tag{6}$$

where S is the true sign and ε_i is a zero mean white noise process, representing the noise deviation of the GRACE processing center i.

Since no true estimate of S is available, the differences between $N - 1$ processing centers and an arbitrarily chosen reference center are calculated as (KOOT et al., 2006):

$$Y_{iN} \equiv X_i - X_N = \varepsilon_i - \varepsilon_N, \quad \forall\, i = 1, \dots, N - 1, \tag{7}$$

with X_N being the time series designated as reference. For this study, the TWSA time series derived from the CSR was selected as the reference series. However, the results of the uncertainty estimates are independent of the choice of the time series of a given GRACE processing center as a reference since the calculations are based on the covariance of differences (Equation 7) (see, for example, TAVELLA and PREMOLI, 1994; KOOT et al., 2006).

The samples of the $N - 1$ differences of solution centers (Equation 7) are stored in the columns of an $M\ x\ (N - 1)$ matrix as:

$$Y = \begin{bmatrix} Y_{1N} & Y_{1N} & \dots & Y_{(N-1)N} \end{bmatrix}, \tag{8}$$

where each row contains a monthly observation (here, M = 183; that is, 183 months from April 2002 to June 2017, after allowing for the missing months as shown in Section 3.2), and

each column represents the difference between each time series and the reference. The covariance matrix of the difference series is calculated as:

$$S = cov(Y), \tag{9}$$

where cov (∘) is the covariance operator, and elements of $S(s_{i,j})$ being variance estimates (for $i = j$) or covariance estimates (for $i \neq j$) otherwise. Presenting the unknown N x N covariance matrix of the individual noises R, it is related to S by (PREMOLI and TAVELLA, 1993):

$$r_{ij} = s_{ij} - r_{NN} + r_{iN} + r_{jN}, \quad i, j = 1, \dots, N-1. \tag{10}$$

Equation 9 is indeterminate because there are $N \times (N+1)/2$ unknowns, but only $N \times (N-1)/2$ equations. Thus, there remain N 'free' parameters that must be determined to obtain a unique solution (GALINDO and PALACIO, 2003).

An important restriction on the solution domain for free parameters, however, is that the estimated covariance matrix should be positive defined (KOOT et al., 2006), i.e., $|R| > 0$ (GALINDO and PALACIO, 2003). This condition restricts the solution domain to the free parameters ($r_{iN}, \dots, r_{NN}, i = 1, \dots, N-1$), but it is still not sufficient to determine them (Koot et al., 2006). The free parameters are thus chosen so that the sum of the estimated correlations between all-time series is minimal, considering the restriction $|R| > 0$. To determine the free N parameters, an appropriate objective function must be defined. The objective function used is given by (PREMOLI and TAVELLA, 1993) as:

$$F(r_{1N}, \dots, r_{NN}) = \frac{K[G(r_{1N}, \dots, r_{NN})]^2}{H(r_{1N}, \dots, r_{NN})} \tag{11}$$

where $K = 3\sqrt{|S|}$, $G(r_{1N}, \dots, r_{NN})$ is a quadratic average of covariance and $H(r_{iN}, \dots, r_{NN}) = |R|$ both depending on the free parameters. Solution of the objective function minimization problem can be done analytically as Premoli and Tavella (1993). Therefore, when the free parameters were estimated, the solution to the other unknown elements of R is given by Equation 9.

3.2.4 Validation metrics

In this study, the metrics used to validate the TWS-GRACE uncertainties based on the TCH method were the unbiased Root Mean Squared Error (ubRMSE) and the Signal-to-Noise Ratio (SNR). According to GRUBER et al. (2020), the ubRMSE is the root mean squared error without the bias component, which can be interpreted as the standard deviation of the random error, calculated as:

$$ubRMSE_1 = \sqrt{\left| r_{11} - \frac{r_{12} r_{13}}{r_{23}} \right|} \tag{12}$$

Where r_{11}, r_{12}, r_{13} and r_{23} the covariances of the matrix.

The signal-to-noise ratio is the dimensionless ratio between the true signal power and the noise power. The higher the signal-to-noise ratio, the less the effect of background noise on signal detection or measurement. In terms of random variables, the signal-to-noise ratio can be defined as:

$$SNR = \frac{std(TWS)}{\sqrt{r_{TWS}}} \tag{13}$$

where std refers to the standard deviation of the TWSA series and r_{TWS} the value of the variance of the error for the respective TWS time series.

3.2.5 Confidence Intervals from block-bootstrapping

Estimating confidence intervals for validation metrics is not always straightforward because the sampling error probability density function (pdf) of the various estimators is often poorly understood or contains parameters that are typically unknown (ZWIEBACK et al., 2012). For TCH-based metrics, it is necessary to use bootstrapping to approximate the sampling error pdf. Bootstrapping is a non-parametric method for obtaining confidence intervals of estimators with unknown sampling distribution (EFRON and TIBSHIRANI, 1986).

The calculation of the above metrics is based on finite-sized samples, with the sample size directly affecting the statistical performance. To quantify the statistical uncertainties

caused by sampling errors, this study relied on the block-bootstrapping resampling technique (Ólafsdóttir and Mudelsee, 2014) to construct confidence intervals (CI) for the ubRMSE validation metric. Relative differences were considered greater than the sampling uncertainty (i.e., they are statistically significant) only when the CI of various cases does not overlap.

Compared with resampling single data points, block-bootstrapping can prevent the autocorrelation in the TWSA time-series from generating an enlarged CI (GRUBER et al., 2020). The specific steps to apply block-bootstrapping to estimate CI of validation metrics are summarized as follows: 1) Resample blocks of collocation input datasets 1000 times, with replacement and preserve the original sample size. For detailed information about the calculation of the optimal block length see GRUBER et al. (2020). 2) Repeatedly calculate the validation metrics in each resampling procedure. 3) Construct the empirical probability distribution of these iteration metrics and obtain the corresponding 95% CI to visually express the sampling errors.

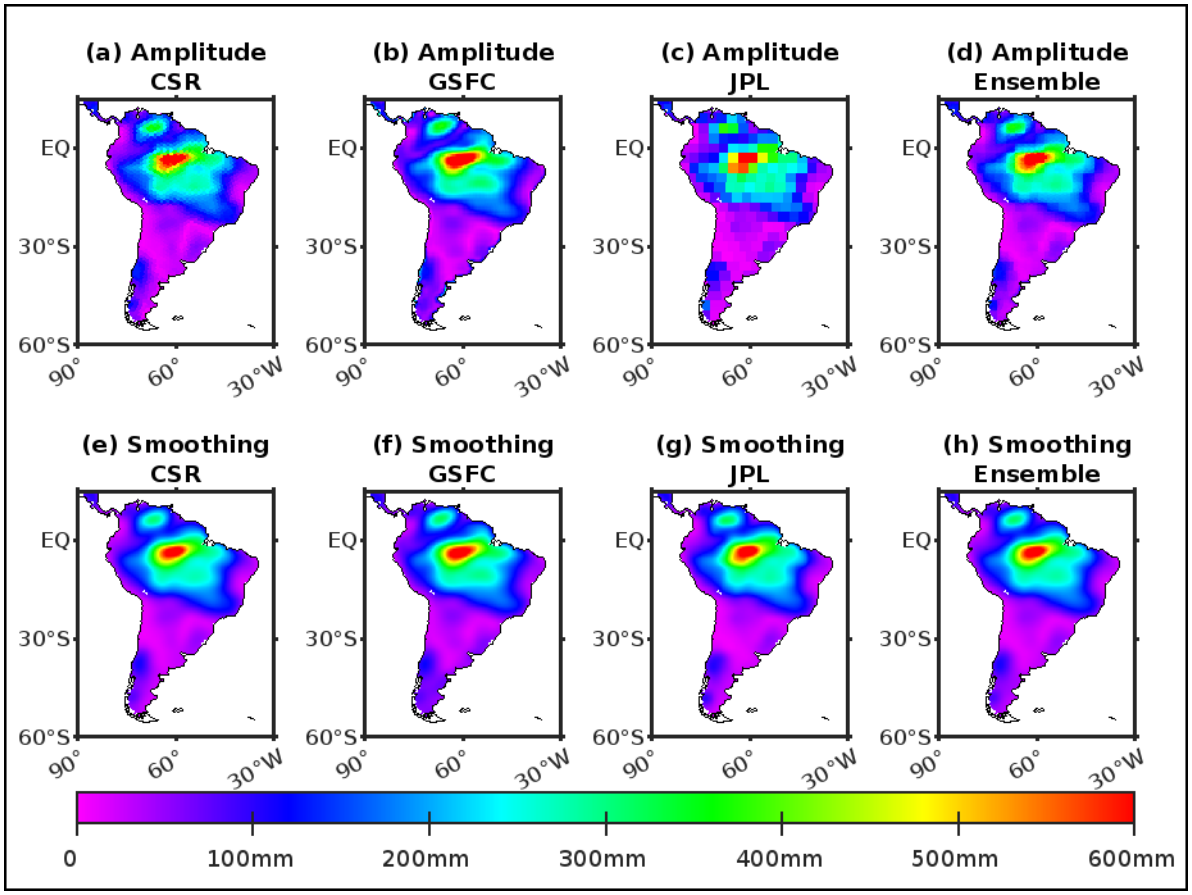
4 RESULTS AND DISCUSSIONS

In this section, the main results of this study are presented. Section 4.1 shows the results related to the comparisons of the solutions of the three GRACE processing centers on a regional scale and Section 4.2 shows the results of the solutions on a basin scale.

4.1 Grid-Scale Comparison

First, the pre-processing stage described in section 2.3 was performed by obtaining the monthly TWSA fields. Then, annual amplitudes were estimated using least-squares at each grid point for the respective series. In addition, the average between the three processing centers was calculated (Figure 2(d)). This was necessary to verify the differences/similarities of the intercomparison in the spatial domain. The resulting annual amplitudes are presented in Figure 2 (a)-(d). It is possible to observe that the signals from the CSR and GSFC solutions showed greater agreement and had a gradual change along adjacent 1° grid cells compared to the marked change observed between adjacent 3° cells seen in the JPL solution.

Figure 2 - Annual amplitudes of processing centers (a) CSR, (b) GSFC, (c) JPL and (d) Ensemble and smoothed annual amplitudes of processing centers (e) CSR, (f) GSFC, (g) JPL and (h) Ensemble of regional TWSA fields for the period April 2002 to June 2017.

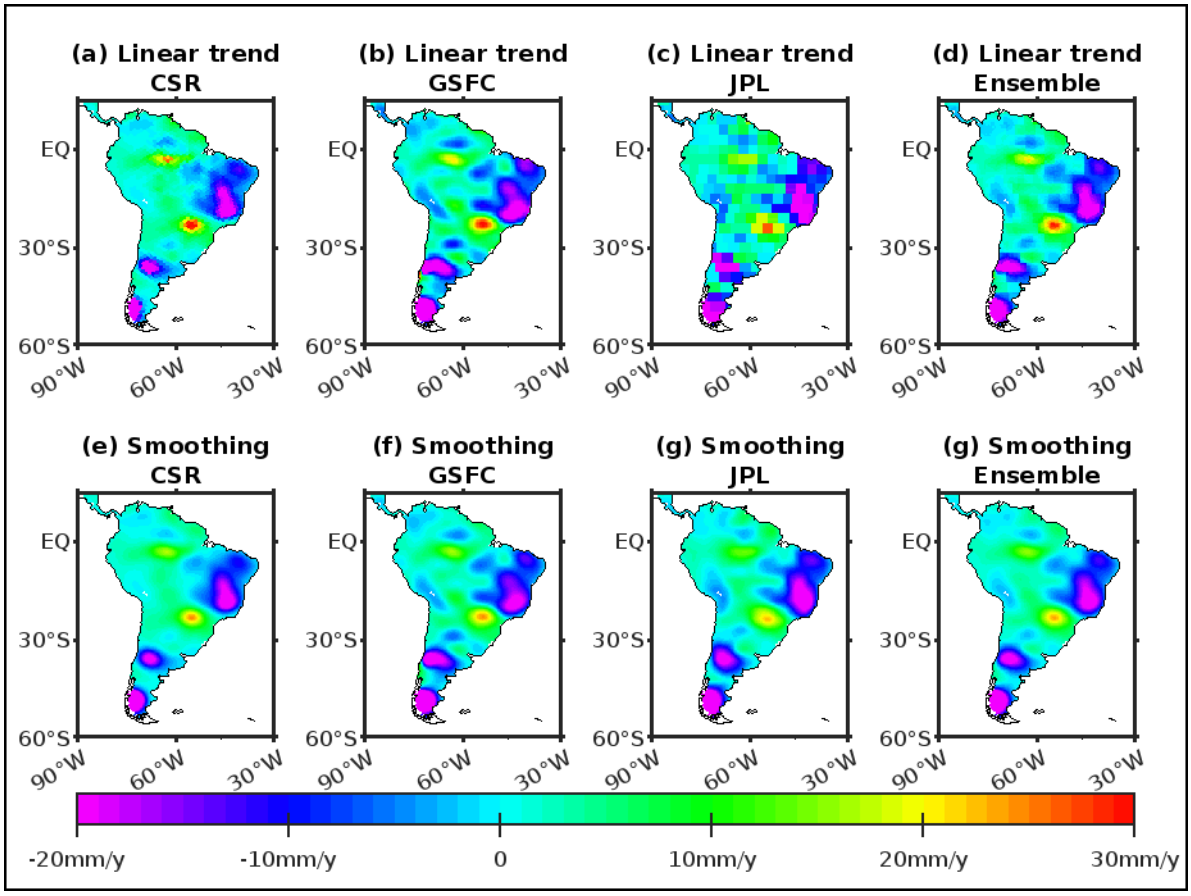


Source: Author (2023).

Furthermore, spatial smoothing of observed amplitudes was performed to reduce signal intensity using the Gaussian filter per section 3.2.2. After applying the Gaussian filter, the results show the same spatial patterns with insignificant differences and good agreement between the three solutions and the ensemble (Figure 2 (d)-(g)). Relative comparisons of annual amplitudes between processing center show correlations of 0.97 between CSR and GSFC, 0.97 between CSR and JPL, and 0.96 between GSFC and JPL. The relative comparisons of the smoothed annual amplitudes between each set of two products showed correlations greater than 0.99.

Figure 3 shows the linear trends of TWSA series based on a least square fitting procedure at each grid point.

Figure 3 - Linear trends of processing centers (a) CSR, (b) JPL, (c) GSFC and (d) ensemble and smoothed linear trends of processing centers (e) CSR, (f) JPL, (g) GSFC and (h) Ensemble of regional TWS fields for the period April 2002 to June 2017.



Source: Author (2023).

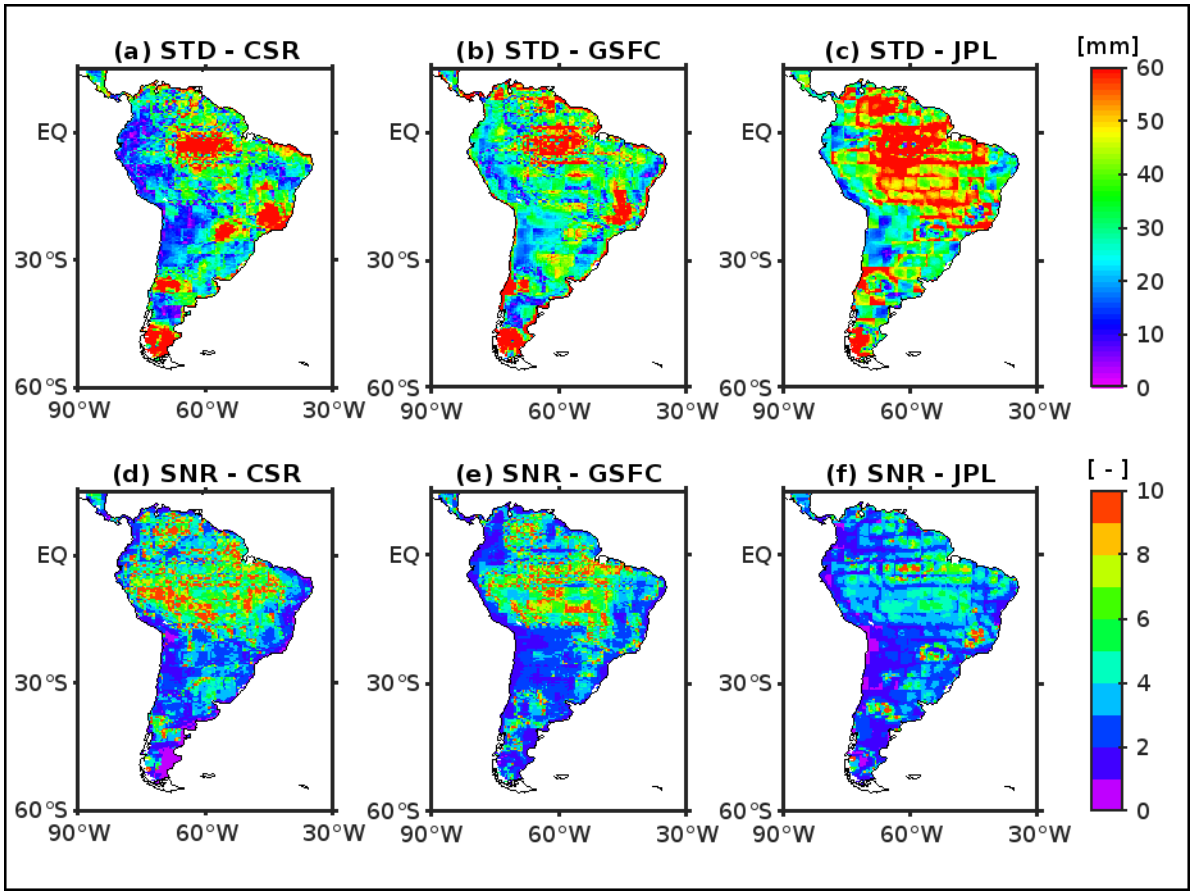
Like the annual amplitudes, it is possible to observe in the linear trends that the signals from the CSR and GSFC solutions show greater agreement, with a gradual change along the grid cells, whereas in the JPL solution, it is possible to observe a more accentuated change (Figure 3c), which is reduced after applying Gaussian smoothing (Figure 3f), where it shows the same patterns with insignificant differences. In addition, the CSR shows lower linear trend values compared with the other solutions, followed by the GSFC and JPL, respectively. It is also possible to notice that the linear trend of the ensemble resembles the linear trend of the GSFC.

The quality of the results published by the three processing centers was evaluated through the uncertainties calculated from the TCH method. First, Equation (10) was minimized at each grid point to estimate the three (N = 3) free parameters, which were used to calculate the other elements of matrix using Equation (9). The diagonal elements of matrix R,

which contain noise variations (r_{11},r_{22},r_{33}), were then used to calculate the noise deviations, which express the quality of each mascon solution (i.e., CSR, JPL, and GSFC). Despite the differences between the strategies adopted for each mascon solution to estimate the parameters, the GRACE measurements are the same among the centers. Thus, it is possible to suppose that GRACE detects the same geophysical phenomena and the common signals are canceled using Equation (7), making the TCH method capable of providing dependent errors in the methodology of each solution, which agrees with the finds reported by FERREIRA et al. (2016). The results of TCH-derived uncertainties are presented in Figure 4.

The CSR (Figure 4a) seems to present the slightest uncertainties compared to the other two solutions (Figures 4b and 4c), especially in the regions of Colombia, Ecuador, Peru, Bolivia, and Chile. The JPL uncertainty map (Figure 4c) shows larger general error patterns than GSFC, which shows higher uncertainties in Brazil. The area-weighted uncertainties for each processing center were calculated considering the uncertainties distribution presented in Figure 4. The results showed that the CSR has a regional (weighted) average uncertainty equals to 34.4 mm, JPL 47.4 mm, and GSFC 38.7 mm. It is known that the average of the datasets is effective in reducing noise compared to the individual members, as shown by Sakumura et al. (2014). Here, the average of the CSR, JPL, and GSFC time series was calculated and compared to each individual solution. The standard deviation of differences was 28.2 mm for CSR, 32.5 mm for JPL, and 32.3 mm for GSFC, which agrees with the results derived by the TCH method. The results of the uncertainties derived from the mean of the ensemble of solutions are presented in Figure 5, where it is possible to observe the similarity with Figure 4 of the results of the uncertainties derived from the TCH.

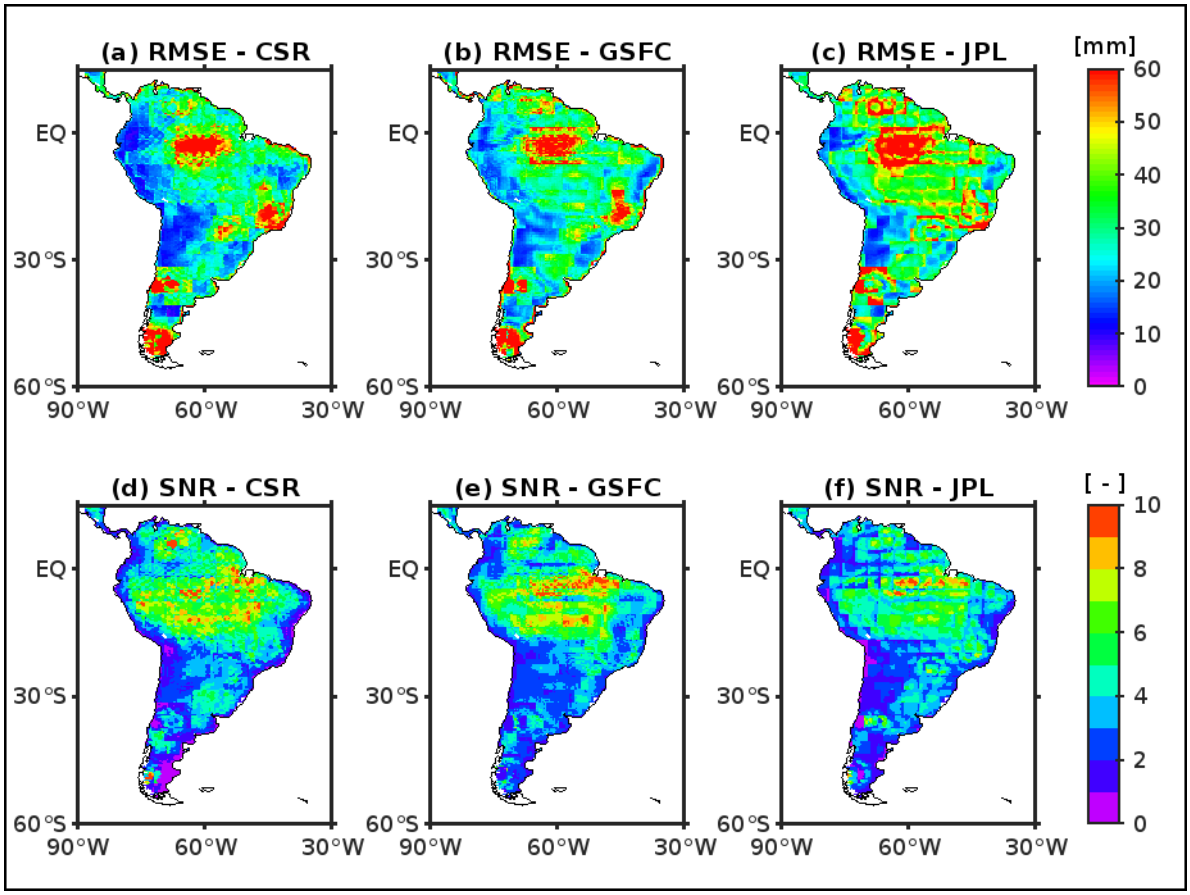
Figure 4 - TCH-derived TWSA uncertainties of processing centers (a) CSR, (b) GSFC and (c) JPL and SNR based of processing centers (d) CSR , (e) GSFC and (f) JPL.



Source: Author (2023).

However, the noise itself is insufficient to describe the relative quality of each processing center. Thus, for each mascon product (i.e., CSR, JPL, and GSFC), the Signal-Noise-Ratio (SNR) was estimated using Equation (13). The results are presented in panels (d), (e) and (f) of Figure 4. The spatial patterns of SNRs are different from those of noise distribution (panels (a), (b), and (c) Figure 4), indicating higher values in regions with strong hydrological signals (e.g., Amazon, La Plata, Orinoco, and Tocantins Basins). The areal weighted SNRS values were calculated for each mascon solution. The results show that CSR has a SNR value equal to 5.3, JPL 3.5, and GSFC 4.7, indicating the good overall performance of the CSR product compared to the other two centers. Again, compared to the ensemble mean, the average SNRs show 7.1 for CSR, 5.0 for JPL, and 6.1 for GSFC. As for the uncertainties, the results of the SNRs derived from the mean of the ensemble of solutions are presented in Figure 5, where it is possible to observe the similarity with Figure 4 of the results of the uncertainties derived from the TCH.

Figure 5 - TWSA uncertainties derived from the ensemble of processing centers (a) CSR, (b) GSFC and (c) JPL and SNR of processing centers (d) CSR, (e) GSFC and (f) JPL.



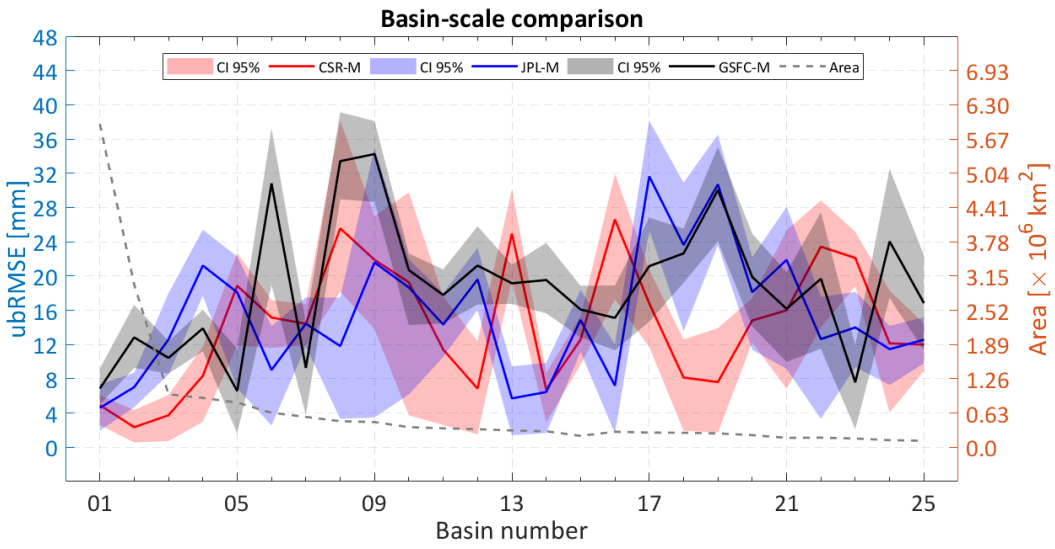
Source: Author (2023).

4.2 Watershed-scale Comparison

South America's watersheds (Figure 1 and Table 2) were chosen to quantify the validation metrics of TWs' relative differences, as well as the confidence intervals relating to the ubRMSE of the three mascon solutions on basin-scale to analyze whether variations strongly impact the content of information within the data. The limits of the 25 basins selected with areas equal to larger than 90 × 10^3 km² were obtained from the South America hydrological basins, available in FAO (2006). This choice was based on the fact that smaller basins than the spatial resolution of the GRACE-derived TWSA (approx. 300 km) cannot be properly captured by GRACE. The watersheds were characterized according to Long et al. (2017) as large basins (basin area > 1,000,000 km²), medium-sized basins (200,000 km² < basin area ≤ 1,000,000 km²), and small basins (basin area ≤ 200,000 km²). Based on basin size, 8% (2) of the basins are grouped into large basins, 72% (18) in medium basins, and 20%

(5) in small basins. In addition, the basin average TWSA for all 25 basins was calculated by applying a regional average by defining a mask with the perimeter of the basins. TCH-based ubRMSE values are shown in Figure 6.

Figure 6 - TCH -based UbRMSE for the 25 watersheds.



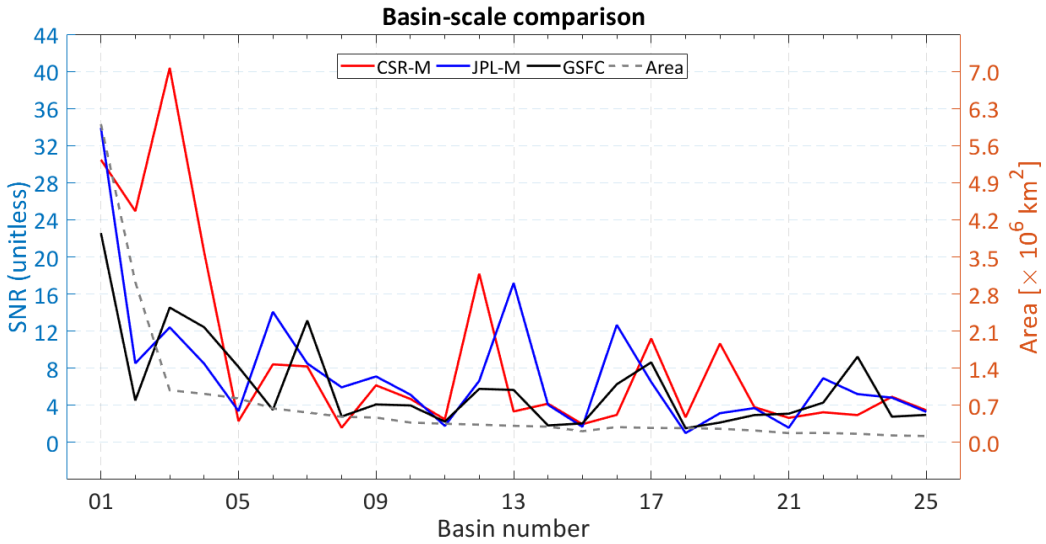
Source: Author (2023).

It is evident that in terms of large basins, the CSR center provides the best performance (Figure 6), followed by the JPL and the GSFC. It is possible to note that ubRMSES vary concerning the size of the basin, which is following the fact that the GRACE is more sensitive and accurate for large watersheds (see, for example, WAHR et. al, 2006; SUN, 2013; REAGER and FAMIGLIETTI, 2013). Regarding medium-size basins, JPL performs better compared to other centers, except among the Atlantic Western Northeast basins (Basin N° 17), Pacific Coast, North Chile (Basin N° 18), Magdalena Basin (Basin N° 19), Atlantic Southeast Coast (Basin N° 20) and Salinas Grande Basin (Basin N° 21). The CSR has greater variations in the results, presenting greater ubRMESs for the Caribbean Coast (N° 13) and Atlantic Eastern Northeast Coast (N° 16) basins. The GSFC overall had larger ubRMSES except for the Atlantic North Coast (N° 07), Caribbean Coast (N° 13), Atlantic Eastern Northeast Coast (N° 16), Atlantic Western Northeast (N° 17), Pacific Coast, North Chile (N° 18) and Magdalena Basin (N° 19). In terms of small basins, each mascon solution's performance differs from basin to basin.

To determine whether variations in the hydrological signal of the relative basins influence the magnitude of the ubRMSES, SNR values were calculated for each watershed (Figure 7). For the large basins it is evident that the dispersion of the SNRs is relatively

decreasing. That is, it usually depends on the size of the basin. This finding agrees with the results by Sakumura et al. (2014) and Ferreira et al. (2016), where it was shown that solutions are due to random errors and not signal differences. Regardless of the basin's size, the hydrological signal's amplitude also plays an essential role in the performance of the GRACE -derived TWS fields. For example, the SNRs of the Atlantic East Coast (N° 05), Atlantic South Coast (N° 08), and La Puna Region (N° 11) are relatively lower than those of the basins with smaller areas. Overall, all three solutions presented high performance in recovering hydrological variations on the basins considered here.

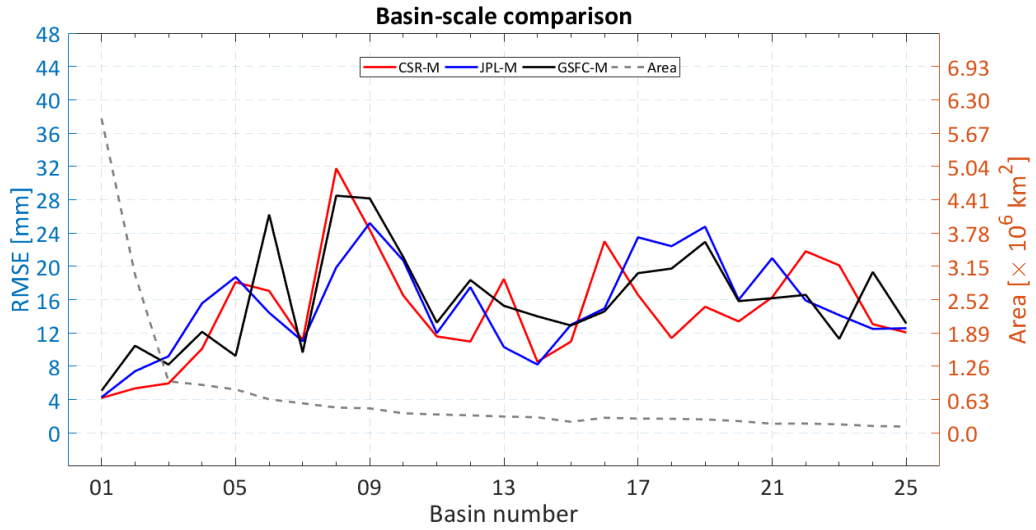
Figure 7 - TCH-based SNR for the 25 watersheds.



Source: Author (2023).

The ubRMSE (Figure 6) and the SNR (Figure 7) of the CSR, for the Amazon (N° 01), La Plata (N° 02), Orinoco (N° 03), and Parnaiba (N° 12) basins, present the noise deviations relatively lowest and highest SNRs. The JPL and GSFC processing centers have the highest SNR values for the Amazon Basin (N° 01), due to its size and the strength of its hydrological signal (Table 2), while the CSR has the highest SNR for the Orinoco Basin (N° 03). The ubRMSE for the Amazon basin is 5.0 mm based on CSR, 4.6 mm for JPL solutions, and 6.8 mm for GSFC (Table 2). The relative comparison between the CSR and the GSFC, in terms of SNR, shows that the first presents values approximately four times greater than the second for the Orinoco (N° 03) and Parnaiba (N° 12) watersheds. Furthermore, for the La Plata basin (N° 02), the SNR of the CSR is five times greater than that of the GSFC.

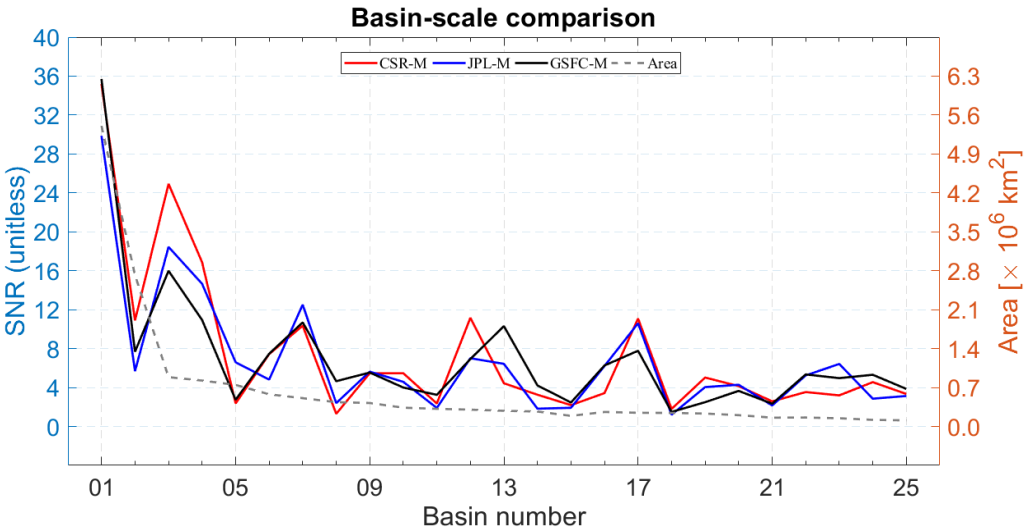
Figure 8 - The ensemble mean based standard deviations for the 25 watersheds.



Source: Author (2023).

Similar to section 4.1, the ensemble solution calculated as the arithmetic mean of the TWSA time series derived from the three processing centers was used to evaluate each ensemble member. The results are shown in Figure 8 and can be compared to those in Figure 6 where they are very similar. The results of the SNRs shown in Figure 9 are also comparable to those in Figure 7.

Figure 9 - SNRs for the 25 watersheds.



Source: Author (2023).

Table 2 - ubRMSE and its respective confidence intervals for the 25 largest basins investigated. The lines [l. u.] are the 95% confidence intervals of the estimates, where l. means low, av. means average and u. means upper.

Basin		**Area (km²)**	**Confidence Interval – ubRMSE (mm)**								
			CSR			**JPL**			**GSFC**		
Nº	**Name**		**l.**	**av.**	**u.**	**l.**	**av.**	**u.**	**l.**	**av.**	**u.**
1	Amazon Basin	5,970,775	2.6	5.0	6.1	1.8	4.6	7.3	4.9	6.8	9.4
2	La Plata Basin	3,016,800	0.5	2.3	4.3	4.6	7.1	8.8	9.3	12.9	16.4
3	Orinoco Basin	974,772	0.8	3.7	6.2	8.4	12.7	17.9	8.6	10.4	12.6
4	Tocantins Basin	915,661	3.0	8.4	11.8	17.8	21.2	25.0	11.5	14.0	16.1
5	Atlantic East Coast	830,359	12.0	18.9	23.0	6.7	18.3	22.4	1.5	6.3	11.1
6	Sao Francisco Basin	635,159	12.1	15.2	17.2	1.9	9.0	13.7	19.0	30.4	37.1
7	Atlantic North Coast	561,413	11.9	14.5	17.0	11.8	14.5	17.3	3.5	9.1	13.9
8	Atlantic South Coast	484,180	18.4	25.8	37.9	2.7	11.8	17.5	28.9	33.5	39.1
9	Pacific Coast, South Chile	469,783	15.3	21.9	27.1	3.3	21.3	34.6	28.6	34.4	38.1
10	Colorado Basin	373,863	3.9	19.2	29.7	5.8	18.8	22.0	14.3	20.7	22.6
11	La Puna Region	348,890	3.1	11.7	18.0	10.8	14.1	17.6	14.5	17.8	20.6
12	Parnaiba Basin	331,643	1.3	6.7	12.4	15.8	19.6	23.1	16.6	21.2	25.7
13	Caribbean Coast	317,043	18.9	24.8	30.2	1.2	5.7	9.3	16.8	19.1	21.2
14	Pacific Coast, Colombia / Ecuador	290,939	3.0	6.7	8.7	1.6	6.5	9.8	15.7	19.6	23.7

Source: Authors (2023).

Table 2 - ubRMSE and its respective confidence intervals for the 25 largest basins investigated. The lines [l. u.] are the 95% confidence intervals of the estimates, where l. means low, av. means average and u. means upper.

Basin		**Area (km²)**	**Confidence Interval – ubRMSE (mm)**								
			CSR			**JPL**			**GSFC**		
Nº	**Name**		**l.**	**av.**	**u.**	**l.**	**av.**	**u.**	**l.**	**av.**	**u.**
15	Pacific Coast, Peru	290,939	9.7	12.7	15.2	11.0	15.0	18.4	12.8	16.1	19.0
16	Atlantic Eastern Northeast Coast	285,877	17.1	26.5	32.0	1.5	7.5	12.0	11.1	15.0	18.6
17	Atlantic Western Northeast	271,751	11.5	16.6	20.6	25.4	31.6	38.4	14.9	21.5	27.0
18	Pacific Coast, North Chile	268,036	2.0	8.0	12.6	14.7	23.0	30.9	19.2	22.6	25.4
19	Magdalena Basin	259,632	1.7	7.4	14.0	23.4	30.5	36.8	24.5	30.2	35.4
20	Atlantic Southeast Coast	224,076	12.1	14.8	17.3	11.8	18.3	22.4	14.4	19.8	25.4
21	Salinas Grandes Basin	177,187	7.1	16.4	25.7	9.0	21.8	28.4	10.8	15.9	20.5
22	Pampas Region	175,610	13.9	23.5	29.2	3.1	12.6	17.6	11.5	18.4	28.0
23	Negro Basin	162,658	18.3	22.0	25.3	8.8	14.1	17.9	1.6	7.7	12.2
24	Mar Chiquita Basin	129,715	3.8	12.3	18.8	6.4	11.3	14.1	17.7	24.2	33.3
25	Central Patagonia Highlands	121,293	9.0	11.9	14.5	9.7	12.7	15.1	11.6	16.9	22.1

Source: Author (2023).

Finally, a block bootstrapping resampling method described in Section 3.5 was used to estimate the 95% confidence interval of the TCH-based ubRMSE metric of the CSR, JPL, and GSFC solutions in the 25 watersheds. The results are shown in Table 2 and Figure 6. For the large basins, as well as the ubRMSE, the reliability of the data decreases as the area decreases. That is, it has larger confidence intervals for smaller areas. In terms of medium and small

basins, as well as the ubRMSE, each processing center's performance differs from basin to basin. The calculation of the weighted mean of the confidence intervals showed that the CSR_l and CSR_u present values of 5.2 and 12.2, respectively. The JPL_l and JPL_u have values of 6.0 and 14.3, respectively. The GSFC_l and GSFC_u have values of 10.0 and 17.0, respectively. Thus, in terms of confidence intervals for medium and small basins, there is no significant difference between the CSR and the GSFC. However, there is a significant difference between JPL and the other two.

5 CONCLUSIONS

The main goal of this work was to estimate the uncertainties of GRACE TWSA mascon solutions provided by CSR, JPL, and GSFC based on the generalized TCH method. Furthermore, the bootstrapping method was investigated to estimate the confidence intervals for the TCH-derived uncertainties, which was not addressed in the previous works. Experiments were carried out at South America's grid-scale and over its largest river basins using GRACE data from April 2002 to June 2017. GRACE-FO TWSA datasets were not used due to the 11 months gap between the two missions (July 2017 to June 2018).

The results are summarized as follows:

1) On a regional scalever South America's continental scale, the CSR processing center has a weighted average uncertainty of 34.4 mm and an SNR of 5.3, while JPL and GSFC have uncertainties of 47.4 and 38.7 mm, with SNRs of 3.5 and 4.7, respectively.
2) Comparisons of the averaged TWSA time series for the 25 largest river basins in South America showed that CSR, JPL, and GSFC have mean uncertainties of 14.3, 15.3 and 18.6 mm.
3) The uncertainties at grid- and basin-scale using TCH were compared with those using the arithmetic mean of the three centers. It was found that the uncertainty values from both approaches agree.
4) In addition, the 95% confidence intervals of the TCH-based ubRMSE metric of the CSR, JPL, and GSFC solutions in the 25 watersheds were estimated. It was possible to identify that for the large basins, the reliability of the data decreases as the area decreases. In terms of medium and small basins, each processing center's performance differs from basin to basin. Thus, in terms of confidence intervals for medium and small basins, there is no significant difference between the CSR and the GSFC.

Overall, the CSR center provides the most accurate monthly solution in terms of TWSA during the period of the comparisons (2002-2017) at the regional and basin scale, although the performance of each processing center differs from basin to basin. It can be concluded that the applicability of GRACE for a specific watershed or region would depend on the signal strength and not on the size of the watershed. Although using the average of the set of three centers is recommended from a practical point of view, the present study contributed to the possibility of choosing appropriate weights so that the average series of the TWSA set will have a low noise variation.

REFERENCES

AHI, G. O., and CEKIM, H. O. Long-term temporal prediction of terrestrial water storage changes over global basins using GRACE and limited GRACE-FO data. **Acta Geodaetica et Geophysica**, v. 56, n. 2, pp. 321-344, 2021.

BOMFIM, E. P., BRAITENBERG, C. and MOLINA, E. C. Mutual evaluation of global gravity models (EGM2008 and GOCE) and terrestrial data in Amazon Basin, Brazil. **Geophysical Journal International**, v. 195, n. 2, p. 870-882, 2013.

CHIN, T. M., GROSS, R. S., and DICKEY, J. O. Multi-reference evaluation of uncertainty in Earth orientation parameter measurements. **Journal of Geodesy**, v. 79, n. 1, pp. 24-32, 2005.

DE LINAGE, C, KIM, H., FAMIGLIETTI, J. S., and YU, J. Y. Impact of Pacific and Atlantic Sea surface temperatures on interannual and decadal variations of GRACE land water storage in tropical South America. **Journal of Geophysical Research: Atmospheres,** v. 118, n. 19, pp. 10,811-10,829, 2013.

EFRON, B., and TIBSHIRANI, R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. **Statistical science**, pp. 54-75, 1986.

FAO. Geonetwork: find and analyse geo-spatial data. 2006. Available: <https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/home> [Accessed: 10 December 2021].

FERREIRA, V. G., MONTECINO, H. D., YAKUBU, C. I., and HECK, B. Uncertainties of the Gravity Recovery and Climate Experiment time-variable gravity-field solutions based on three-cornered hat method. **Journal of applied remote sensing,** v. 10, n. 1, pp. 015015, 2016.

FRAPPART, F., and RAMILLIEN, G. Monitoring groundwater storage changes using the Gravity Recovery and Climate Experiment (GRACE) satellite mission: A review. **Remote Sensing,** v. 10, n. 6, pp. 829, 2018.

GALINDO, F. J., and PALACIO, J. (2003). Post-processing ROA data clocks for optimal stability in the ensemble timescale. **Metrologia**, v. 40, n. 3, p. S237, 2003.

GARREAUD, R. D., VUILLE, M., COMPAGNUCCI, R, & MARENGO, J. Present-day south american climate. **Palaeogeography, Palaeoclimatology, Palaeoecology**, v. 281, n. 3-4, pp. 180-195, 2009.

GERUO, A., WAHR, J., and ZHONG, S. Computations of the viscoelastic response of a 3-d compressible earth to surface loading: an application to glacial isostatic adjustment in antarctica and canada. **Geophysical Journal International**, v. 192, n. 2, pp. 557-572, 2013.

GRAU, H. R., and AIDE, M. Globalization and land-use transitions in Latin America. **Ecology and society**, v. 13, n. 2, 2008.

GRAY, J. E., and ALLAN, D. W. A method for estimating the frequency stability of an individual oscillator. In: 28th annual symposium on frequency control. IEEE, 1974. p. 243-246.

GRUBER, A., DE LANNOY, G., ALBERGEL, C., AL-YAARI, A., BROCCA, L., CALVET, J. C., COLLIANDER, A., COSH, M., CROW, W., DORIGO, C., DRAPER, C., HIRSCHI, M., KERR, Y., KONINGS, A. LAHOZ, W. et al. Validation practices for satellite soil moisture retrievals: What are (the) errors?. **Remote sensing of environment**, v. 244, p. 111806, 2020.

HASAN, E., TARHULE, A., and KIRSTETTER, P. E. Twentieth and twenty-first century water storage changes in the Nile river basin from grace/grace-fo and modeling. **Remote Sensing**, v. 13, n. 5, p. 953, 2021.

HASSAN, A. and JIN, S. Water storage changes and balances in Africa observed by GRACE and hydrologic models. **Geodesy and Geodynamics**, v. 7, n. 1, p. 39-49, 2016.

KOOT, L., VIRON, O. D., and DEHANT, V. Atmospheric angular momentum time-series: characterization of their internal noise and creation of a combined series. Journal of Geodesy, v. 79, p. 663-674, 2006.

LANDERER, F. W., and SWENSON, S. C. Accuracy of scaled GRACE terrestrial water storage estimates. **Water resources research**, v. 48, n. 4, 2012.

LANDERER, F. W., FLECHTNER, F. M., SAVE, H., WEBB, F. H., BANDIKOVA, T., BERTIGER, W. I., BETTADPUR, S. V., BYUN, S. H., DAHLE, C. et al. Extending the global mass change data record: GRACE Follow-On instrument and science data performance. **Geophysical Research Letters**, v. 47, n. 12, p. e2020GL088306, 2020.

LOEW, A., BELL, W., BROCCA, L., BULGIN, C. E., BURDANOWITZ, J., CALBET, X., DONNER, R. V., GHENT, D., GRUBER, A., KAMINSKI, T., KINZEL, J. KLEPP, C. LAMBERT, J. C. SCHAEPMAN-STRUB, G., SCHRÖDEr, M., and VERHOELST, T. Validation practices for satellite-based Earth observation data across communities. Reviews of Geophysics, v. 55, n. 3, p. 779-817, 2017.

LONG, D., SHEN, Y., SUN, A., HONG, Y., LONGUEVERGNE, L., YANG, Y., LI, B. and CHEN, L. Drought and flood monitoring for a large karst plateau in Southwest China using extended GRACE data. **Remote Sensing of Environment**, v. 155, p. 145-160, 2014.

LONG, D., LONGUEVERGNE, L., and SCANLON, B. R. Global analysis of approaches for deriving total water storage changes from GRACE satellites. **Water Resources Research**, v. 51, n. 4, p. 2574-2594, 2015.

LONG, D., PAN, Y., ZHOU, J., CHEN, Y., HOU, X., HONG, Y., LONGUEVERGNE, L. Global analysis of spatiotemporal variability in merged total water storage changes using multiple GRACE products and global hydrological models. **Remote sensing of environment**, v. 192, p. 198-216, 2017.

LOOMIS, B. D., LUTHCKE, S. B., and SABAKA, T. J. Regularization and error characterization of GRACE mascons. **Journal of geodesy**, v. 93, p. 1381-1398, 2019.

LUTHCKE, S. B., SABAKA, T. J., LOOMIS, B. D., ARENDT, A. A., MCCARTHY, J. J., and CAMP, J. Antarctica, Greenland and Gulf of Alaska land-ice evolution from an iterated GRACE global mascon solution. **Journal of Glaciology**, v. 59, n. 216, p. 613-631, 2013.

MO, S., ZHONG, Y., FOROOTAN, E., MEHRNEGAR, N., YIN, X., WU, J., FENG, W., and SHI, X. Bayesian convolutional neural networks for predicting the terrestrial water storage anomalies during GRACE and GRACE-FO gap. **Journal of Hydrology**, v. 604, p. 127244, 2022.

MOORE, P., and WILLIAMS, S. D. P. Integration of altimetric lake levels and GRACE gravimetry over A frica: Inferences for terrestrial water storage change 2003–2011. **Water Resources Research**, v. 50, n. 12, p. 9696-9720, 2014.

MULLER, P. M., and SJOGREN, W. L. Mascons: Lunar mass concentrations. Science, v. 161, n. 3842, p. 680-684, 1968.

ÓLAFSDÓTTIR, K. B., and MUDELSEE, M. More accurate, calibrated bootstrap confidence intervals for estimating the correlation between two time series. **Mathematical Geosciences**, v. 46, n. 4, p. 411-427, 2014.

PAULSON, A., ZHONG, S., and WAHR, J. Inference of mantle viscosity from GRACE and relative sea level data. **Geophysical Journal International**, v. 171, n. 2, p. 497-508, 2007.

PELTIER, W. R., ARGUS, D. F., and DRUMMOND, R. Comment on "An assessment of the ICE-6G_C (VM5a) glacial isostatic adjustment model" by Purcell et al. **Journal of Geophysical Research: Solid Earth**, v. 123, n. 2, p. 2019-2028, 2018.

PREMOLI, A., and TAVELLA, P. A revisited three-cornered hat method for estimating frequency standard instability. IEEE Transactions on instrumentation and measurement, v. 42, n. 1, p. 7-13, 1993.

REAGER, J. T. and FAMIGLIETTI, J. S. Characteristic mega-basin water storage behavior using GRACE. **Water resources research**, v. 49, n. 6, p. 3314-3329, 2013.

REAGER, J. T., THOMAS, B. F., and FAMIGLIETTI, J. S. River basin flood potential inferred using GRACE gravity observations at several months lead time. **Nature Geoscience**, v. 7, n. 8, p. 588-592, 2014.

RODELL, M., FAMIGLIETTI, J. S., CHEN, J., SENEVIRATNE, S. I., VITERBO, P., HOLL, S., and WILSON, C. R. Basin scale estimates of evapotranspiration using GRACE and other observations. **Geophysical Research Letters**, v. 31, n. 20, 2004.

RODELL, M., CHEN, J., KATO, H., FAMIGLIETTI, J. S., NIGRO, J., and WILSON, C. R. Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. **Hydrogeology Journal**, v. 15, n. 1, pp. 159-166, 2007.

SAKUMURA, C., BETTADPUR, S. and BRUINSMA, S. Ensemble prediction and intercomparison analysis of GRACE time-variable gravity field models. **Geophysical Research Letters**, v. 41, n. 5, p. 1389-1397, 2014.

SAVE, H., BETTADPUR, S., and TAPLEY, B. D. High-resolution CSR GRACE RL05 mascons. **Journal of Geophysical Research: Solid Earth**, v. 121, n. 10, pp. 7547-7569, 2016.

SAVE, H., 2020. CSR GRACE and GRACE-FO RL06 Mascon Solutions v02.

SUN, A. Y. Predicting groundwater level changes using GRACE data. **Water resources research**, v. 49, n. 9, pp. 5900-5912, 2013.

SUN, A. Y., SCANLON, B. R., SAVE, H., and RATEB, A. Reconstruction of GRACE total water storage through automated machine learning. **Water Resources Research**, v. 57, n. 2, pp. e2020WR028666, 2021.

SUTTERLEY, T. C., VELICOGNA, I., CSATHO, B., VAN DEN BROEKE, M., REZVAN-BEHBAHANI, S., and BABONIS, G. Evaluating Greenland glacial isostatic adjustment corrections using GRACE, altimetry and surface mass balance data. **Environmental Research Letters,** v. 9, n. 1, pp. 014004, 2014.

SWENSON, S., and WAHR, J. Estimating large-scale precipitation minus evapotranspiration from GRACE satellite gravity measurements. **Journal of Hydrometeorology**, v. 7, n. 2, pp. 252-270, 2006.

TAPLEY, B. D., WATKINS, M. M., FLECHTNER, F., REIGBER, C., BETTADPUR, S., RODELL, M., SASGEN, I., FAMIGLIETTI, J. S., LANDERER, F. W., CHAMBERS, D. P., REAGER, J. T., GARDNER, A. S., SAVE, H., IVINS, E. R., SWENSON, S. C., BOENING, C., DAHLE, C., WIESE, D. N., DOBSLAW, H., TAMISIEA, M. E. and VELICOGNA, I. Contributions of GRACE to understanding climate change. **Nature climate change**, v. 9, n. 5, pp. 358-369, 2019.

TAVELLA, P., and PREMOLI, A. Estimating the instabilities of N clocks by measuring differences of their readings. **Metrologia**, v. 30, n. 5, pp. 479, 1974.

TEDESCHI, R. G., and COLLINS, M. The influence of ENSO on South American precipitation during austral summer and autumn in observations and models. **International Journal of Climatology**, v. 36, n. 2, pp. 618-635, 2016.

VALTY, P., DE VIRON, O., PANET, I., VAN CAMP, M., and LEGRAND, J. Assessing the precision in loading estimates by geodetic techniques in Southern Europe. **Geophysical Journal International**, v. 194, n. 3, pp. 1441-1454, 2013.

VAN DAM, T., WAHR, J., and LAVALLÉE, D. A comparison of annual vertical crustal displacements from GPS and Gravity Recovery and Climate Experiment (GRACE) over Europe. **Journal of Geophysical Research: Solid Earth**, v. 112, n. B3, 2007.

WAHR, J., MOLENAAR, M., BRYAN, F. Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. **Journal of Geophysical Research: Solid Earth**, v. 103, n. B12, p. 30205-30229, 1998.

WAHR, J., SWENSON, S., ZLOTNICKI, V., and VELICOGNA, I. Time-variable gravity from GRACE: First results. **Geophysical Research Letters,** v. 31, n. 11, 2004.

WAHR, J., SWENSON, S., and VELICOGNA, I. Accuracy of GRACE mass estimates. **Geophysical Research Letters**, v. 33, n. 6, 2006.
WAHR, J., SELVANS, Z. A., MULLEN, M. E., BARR, A. C., COLLINS, G. C., SELVANS, M. M., and PAPPALARDO, R. T. Modeling stresses on satellites due to nonsynchronous rotation and orbital eccentricity using gravitational potential theory. Icarus, v. 200, n. 1, p. 188-206, 2009.

WATKINS, M. M., WIESE, D. N., YUAN, D. N., BOENING, C., and LANDERER, F. W. Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. **Journal of Geophysical Research: Solid Earth**, v. 120, n. 4, pp. 2648-2671, 2015.

WIESE, D. N., LANDERER, F. W., and WATKINS, M. M. Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. **Water Resources Research,** v. 52, n. 9, pp. 7490-7502, 2016.

WIESE, D. N., YUAN, D. N., BOENING, C., LANDERER, F. W., and WATKINS, M. M. JPL GRACE mascon ocean, ice, and hydrology equivalent water height release 06 coastal resolution improvement (CRI) filtered version 1.0. **DAAC: Pasadena**, CA, USA, 2018.

YAO, C., LI, Q., LUO, Z., WANG, C., ZHANG, R., & ZHOU, B. Uncertainties in GRACE-derived terrestrial water storage changes over mainland China based on a generalized three-cornered hat method. **Chinese Journal of Geophysics**, v. 62, n. 3, p. 883-897, 2019.

YI, S., and SNEEUW, N. Filling the data gaps within GRACE missions using singular spectrum analysis. **Journal of Geophysical Research: Solid Earth**, v. 126, n. 5, pp. e2020JB021227, 2021.

ZWIEBACK, S., SCIPAL, K., DORIGO, W., and WAGNER, W. Structural and statistical properties of the collocation technique for error characterization. **Nonlinear Processes in Geophysics**, v. 19, n. 1, p. 69-80, 2012.