Assessment of Tide Gauge Biases and Precision by the Combination of Multiple Collocated Time Series

Kevin Gobron, Olivier de Viron, Guy Woppelmann, Etienne Poirier, Valérie Ballu, Michel van Camp

Assessment of Tide Gauge Biases and Precision by the Combination of Multiple Collocated Time Series

Assessment of tide gauges biases and precisions by the combination of multiple co-located time series

Kevin Gobron *

LIENSs (University of La Rochelle CNRS UMR7266), La Rochelle, France

Olivier de Viron

LIENSs (University of La Rochelle CNRS UMR7266), La Rochelle, France

Guy Wöppelmann

LIENSs (University of La Rochelle CNRS UMR7266), La Rochelle, France

Étienne Poirier

LIENSs (University of La Rochelle CNRS UMR7266), La Rochelle, France

Valérie Ballu

LIENSs (University of La Rochelle CNRS UMR7266), La Rochelle, France

Michel Van Camp

Royal Observatory of Belgium, Uccle, Belgium

Corresponding author address: Kevin Gobron, Littoral Environnement et Sociétés, 2 rue Olympe de Gouges, La Rochelle, 17000.

17 E-mail: kevin.gobron1@univ-lr.fr

ABSTRACT

This study proposes a method for the cross-calibration of tide gauges. Based on the combination of at least three co-located sea level time series, it takes advantage of the Least-Squares Variance Component Estimation method to assess both sea-level biases and uncertainties in real conditions. The method was applied to a multi-instrument experiment carried out on Aix island, France, in 2016. Six tide gauges were deployed to carry out simultaneous sea level recordings for 11 hours. The best results were obtained with an electrical contact probe, which reaches a 3-millimeter uncertainty. The method allows assessing both the biases and the precision – i.e., the full accuracy – for each instrument. The results obtained with the proposed combination method have been compared to that of a buddy-checking method. It showed that the combination of all time series also provides more precise bias estimates.

1. Introduction

Tide gauges aim at measuring the vertical distance between the sea level and a reference level (or datum). Historically, tide gauges were first used for tide prediction and navigation (Cartwright 2000); today, their applications have been extended (Pugh and Woodworth 2014). Clustered into networks of continuously operating stations, they are the key components of storm surge or tsunami warning systems and climate-related monitoring programs, such as the Global Sea Level Observing System (GLOSS) (IOC et al. 2012).

A wide range of distance meter technologies can serve to implement a tide gauge, as long as it can resolve both sea level and datum along the vertical. The datum of a sea level station is a local and conventional reference level, independent from any instrument. It enables the construction of long time series with successive or overlapping tide gauges. The datum is defined through a network of benchmarks grounded around the sea level station, some of them can be benchmarks from leveling networks (IOC 1985; Pugh and Woodworth 2014). Thus, a preliminary step in field calibrations consists of tying the reference gauge to the station datum or controlling whether it is properly tied.

The simplest and oldest types of tide gauge are graduated poles or tide poles placed against a vertical structure at the coast (Cartwright 2000). Tide poles requiring human-made measurements are still in-use, along with electric tape probes for on-site field calibration of more elaborated self-recording tide gauges. Since 1985, the manuals of the Intergovernmental Oceanographic Commission (IOC) have covered the basic principles of the main types of tide gauges in use across the world, ranging from mechanical float gauges (IOC 1985) to radar technologies (IOC et al. 2016), including pressure and acoustic gauges (IOC et al. 2002, 2006).

Over the past decade, radar-based technologies appeared as the preferred ones (IOC et al. 2016). However, new technologies are emerging, based on Global Navigation Satellite System (GNSS) buoys (André et al. 2013), GNSS reflectometry (GNSS-R) (Larson et al. 2017), or laser distance measurement (MacAulay et al. 2008). A tide gauge complying with GLOSS standards should be capable of measuring instantaneous sea level with an accuracy better than 1 cm, in all conditions of tide, waves, currents and weather (IOC et al. 2016). As laboratory testings do not ensure those performances, the practice has evolved towards field experiments (Míguez Martín et al. 2008b,a; Park et al. 2014; Pérez et al. 2014).

When dealing with accuracy requirements, it is useful to distinguish random and systematic errors. The random error is the error component that, in replicate measurements, varies in an unpredictable manner, whereas the systematic error is the error component that, in replicate measurements, changes in a predictable manner (BIPM et al. 2008).

Given the crucial role of tide gauges in coastal sea level observation, the increasing number of available technologies and the evolution of accuracy requirements, this study aims at providing a cross-calibration method that quantifies both systematic errors – the biases – and random errors – the uncertainties – of sea level time series.

Determining the errors of given time series can be achieved through three approaches : (a) the observed time series can be compared with that from a more precise instrument, (b) it can be compared with theory in cases where the observed phenomena can be very precisely modeled, and (c) observed time series of three or more instruments can be analyzed to obtain enough information to determine the uncertainty of each.

The approach (a), also known as buddy-checking, is routinely used during calibration campaigns where a pair of tide gauges are compared over a tidal cycle, sometimes with the help of the so-called Van de Casteele diagram (Lennon 1968; IOC 1985). During the last decade, several studies

have investigated the performances of radar gauges, pressures gauges, GNSS buoys or GNSS reflectometry based on this approach (Watson et al. 2008; Míguez Martín et al. 2008a, 2012; Pérez et al. 2014; Larson et al. 2017; Pytharouli et al. 2018). Even if this approach can provide bias estimates and general accuracy metrics, such as mean error or root mean square error (RMSE), it cannot rigorously separate the uncertainties of each gauge.

Approach (b) would correspond to removing a tide model from the measured sea level time series. But, because of the complexity of meteorological and ocean dynamics involved in sea level fluctuations, these models cannot help to assess the performance of tide gauges at the targeted centimeter level.

Approach (c) is classically used in metrology (Pálinkáš et al. 2017) and has often been used in geodesy through the three-cornered-hat (TCH) estimation method (Gray and Allan 1974), for example, to determine the stability of reference station positions (Feissel-Vernier et al. 2007; Abbondanza et al. 2015) or the precision of space gravity model (de Viron et al. 2008; Valty et al. 2013). The TCH is not the only possible implementation of the approach (c): the more general framework of Variance Component Estimation (VCE) can similarly address this problem, as shown by the theoretical example 4.10 of Amiri-Simkooei (2007). The TCH and VCE examples can separate the uncertainty of each gauge, but assume the absence of sea level biases.

To take advantage of both approaches (a) and (c), this study proposes a combination model that extends the use of approach (c) to the analysis of potentially biased time series. Obtaining the tide gauge uncertainties in addition to the sea level bias parameters is made possible by the use the Least-Squares Variance Component Estimation (LS-VCE) method (Teunissen and Amiri-Simkooei 2008). As the model can handle an arbitrary number of time series, it is suited for multi-instrument experiments.

The method is applied to an on-site field calibration experiment carried out at Aix Island, mid-Atlantic coast of France, where a permanent radar gauge has operated for several years (Gouriou et al. 2013), and various types of tide gauges (including some emerging technologies) were temporarily deployed during the experiment within meters from each other over a tidal cycle in 2016.

2. The Aix Island experiment

This experiment was carried out on June 7, 2016, by a team of scientists (see Acknowledgment section). They measured one semi-diurnal spring tidal cycle with an amplitude of 5.22 m using 6 different instruments.

Tide gauges recorded the sea level for 11 hours. Each tide gauge record is defined as the average over a 2 minutes acquisition window every 10 minutes.

The 6 tide gauges were a permanent radar gauge (RADAR), a permanent tide pole (POLE), an electrical contact probe (PROBE), two GNSS buoys (BUOY1 and BUOY2) and a laser distance-meter (LASER). RADAR, POLE, PROBE, and LASER are shown in Figure 1 and the two GNSS buoys in Figure 2. All tide gauges and the reference GNSS station were referenced to the station datum by leveling.

The radar gauge (RADAR) is the primary tide gauge of the permanent sea level observatory of the Aix Island. This station contributes to the French sea level observation network (RONIM) operated by the French hydrographic service (SHOM). It is a Krohne Optiwave 7300C gauge that measures the air range between the transmitter fixed above the sea surface and the sea surface with a sampling frequency of 1 Hz using a frequency modulated continuous wave technology (IOC et al. 2016).

The tide pole (POLE) is a permanent instrument made of a plastic staff with graduations every 10 centimeters fixed vertically by a stainless steel structure (Figure 1). The operator estimates the sea level visually over the standard pre-defined 2 minutes acquisition period.

The electrical contact probe (PROBE) is a measuring tape with millimeter graduations ended by an electrical device that emits a short signal when detecting the seawater surface. We used a Schill probe installed within a stilling pipe anchored along the tide pole (Figure 1). A sea level record from PROBE is an average over the 2 minutes of human-made readings every 15 seconds. Electric probes are typically used as the reference gauge in tide gauge calibrations, so was it in our study. The stilling pipe was too short to allow measurements at the lowest sea levels, which resulted in a gap between 10:00 and 12:10 A.M.

The first GNSS buoy (BUOY1), designed at the Institut de Physique du Globe de Paris (IPGP), is a GNSS antenna installed above a lifebuoy and protected from the water by a radome (Figure 2). The second one (BUOY2), designed by the Division Technique de l'Institut National des Sciences de l'Univers (DT INSU), is a GNSS antenna housed in the center of a tripod floating structure (Figure 2). The receivers and batteries of the buoys are located inside a metallic cylinder under each antenna. These two buoys (BUOY1 and BUOY2) were already used in previous campaigns (André et al. 2013). The heights between their phase centers and the water surface are known at the sub-centimeter level thanks to previous testings carried out under calm conditions.

The buoy vertical positions, i.e., ellipsoidal heights, from GPS were assessed by post-processing, using a double-differences strategy with a baseline of about 300 m from the ILDX GNSS reference station. Only satellites with elevation angles above 15 degrees were used, with a combination of both L1 and L2 frequencies. The centimeter level accuracy was achieved, using full ambiguity resolution with the RTKlib software suite with RTKPOST v2.4.2 program (Takasu 2013).

LASER is a reflector-free distance-meter Leica DISTO A6. This type of instrument is built for solid surface ranging but showed fair to good performances during this experiment. This instrument uses an optical laser beam with a wavelength of 635 nm. Each LASER record corresponds to an average of measurements done every 4 seconds.

All instruments time series are presented in Figure 3. Due to data transmission loss and GNSS recording issues during the experiment, some records from the LASER, BUOY1, and BUOY2 instruments are missing.

3. Calibration methods

This study proposes a combination method to go beyond the classical difference methods, allowing a better determination of the biases and their precision. For comparison, we processed the time series using both the combination method and the classical difference method used by the hydrography community, the so-called Van de Casteele (VdC) diagram (Lennon 1968).

a. Sea level error model

Due to the short recording period, this study only considers the influence of the three most common types of range measurement biases on the resulting sea level time-series, namely: the height references, scale, and clock synchronization errors (Watson et al. 2008; Míguez Martín et al. 2008b).

While converting original range measurements into sea level time series, range biases turn into sea level biases that must be quantified and removed. This study proposes a linear sea level bias model, which expresses the sea level bias as a function of the measured sea level itself. More precisely, the model links the i-th sea level time series $y_i(t)$ to the real sea level $h(t)$ through

$$y_i(t) = h(t - \tau_i) + \beta_i \times y_i(t) + \alpha_i + e_i(t), \quad (1)$$

where $\beta_i \times y_i(t) + \alpha_i$ is the linear sea level bias model, and $e_i(t)$ is a random error modeled by a centered normal distribution of unknown variance σ_i^2.

In equation (1), α_i corresponds to the intercept: a constant term representing the sea level bias when $y_i(t) = 0$. It may result from a height reference error, but also from the influence of a scale error, as mentioned by (Pérez et al. 2014). β_i corresponds to the scale error: a multiplying factor that causes a sea level bias proportional to the tidal range. It can result from both instrument or installation defaults. Finally, τ_i is the time delay between different tide gauges: it results from clock synchronization issues.

The measured sea-level $y_i(t)$ depends non-linearly on the time delay τ_i, which makes linear determinations, like the one proposed in this paper, impossible. However, it can be corrected before the other bias estimations, e.g., by computing the delay that maximizes the cross-correlation between a tested signal and a reference signal. Obtaining τ_i by cross-correlation avoids any assumptions on the periodicity of the measured signal. In our case, the time delay estimation showed that the best correlation was achieved with no delays added i.e., $\tau_i = 0, \forall i$.

The sea level bias model directly quantifies the amplitude of the bias associated with a measurement $y_i(t)$. The correction of the sea level time series can be done after the calibration experiment by subtracting the estimated bias model from the measurements. This linear model can be adapted to other types of biases. For example, longer time series analysis (several days, months, or years) may require to consider time-dependent biases such as trends and jumps (Pytharouli et al. 2018).

b. Difference-based calibration methods (DIFF)

Difference-based methods (DIFF) consist in analyzing the differences $\Delta_{y_i}(t) = y_i(t) - y_{ref}(t)$ between the time series of a tested instrument $y_i(t) = h(t) + \beta_i \times y_i(t) + \alpha_i + e_i(t)$ and that of a reference instrument $y_{ref}(t) = h(t) + e_{ref}(t)$.

A commonly used tool for DIFF methods is the Van de Casteele (VdC) diagram, which represents the sea level difference $\Delta_{y_i}(t)$ as a function of $y_i(t)$. Initially developed in 1962, for mechanical tide gauges (IOC 1985), the VdC diagram is nonetheless still applicable for modern sea level measurement technologies (Míguez Martín et al. 2008b). The most attractive feature of this diagram is a fast, visual, detection of possible biases with only one tidal cycle. Figure 4 shows the sea level error patterns resulting from the most common range measurement errors IOC (1985).

In the presence of the linear biases mentioned before, $\Delta_{y_i}(t)$ follows

$$\Delta_{y_i}(t) = \beta_i \times y_i(t) + \alpha_i + e_i(t) - e_{ref}(t). \tag{2}$$

In other words, estimates of the sea level bias parameters α_i and β_i of equation (1) can be obtained by linear regression of $\Delta_{y_i}(t)$ on $y_i(t)$, which corresponds to fitting a line on a VdC diagram.

Assuming that both random errors $e_i(t)$ and $e_{ref}(t)$ are uncorrelated, the term $e_i(t) - e_{ref}(t)$ in equation (2) follows a centered normal distribution with an unknown variance $\sigma_i^2 + \sigma_{ref}^2$. The merge of the random errors $e_i(t)$ and $e_{ref}(t)$ in the differences $\Delta_{y_i}(t)$ imply that, without assumption, DIFF methods can only assess the variance $\sigma_i^2 + \sigma_{ref}^2$, which is just an upper bound to the tested gauge variance σ_i^2 (Lentz 1993; Míguez Martín et al. 2008b; Pytharouli et al. 2018). To separate σ_i^2 and σ_{ref}^2, a piece of additional information is needed: a third time series.

c. The combination-based calibration method (COMB)

When more than 2 time-series are available, it becomes possible to assess the uncertainties and biases from each tide gauge by estimating a weighted combination of all the time series, using a variance component estimation method. In the following, the acronym COMB refers to the combination method.

1) FUNCTIONAL MODEL OF THE OBSERVATIONS

Noting $\mathbf{y}_i$ the i-th gauge $k \times 1$ observation vector (or time series), the full $pk \times 1$ stacked vector $\mathbf{y}$, containing all observations from the p instruments, can be written as

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1^T & \dots & \mathbf{y}_i^T & \dots & \mathbf{y}_p^T \end{bmatrix}^T.$$

The functional model links the expectation $\mathrm{E}(.)$ of the $pk \times 1$ observations vector $\mathbf{y}$ to q unknown parameters by using a model of observation equations. When there is no theoretical model for the observed signal, we can estimate a combined time series $\mathbf{h}$, from the $k \times 1$ vector of the p time series written as

$$\mathbf{h} = \begin{bmatrix} h_1 & \dots & h_j & \dots & h_k \end{bmatrix}^T.$$

In the case of unbiased gauges, the functional model would be $\mathrm{E}(\mathbf{y}_i) = \mathbf{h}$ for every gauge. In the case of the cross-calibration of possibly biased time series, the functional model should also account for the biases. The model of observation equations of the i-th gauge can then be written as

$$\mathrm{E}(\mathbf{y}_i) = \begin{cases} \mathbf{h}, & \text{if the } i\text{-th gauge is unbiased} \\ \mathbf{h} + \beta_i \cdot \mathbf{y}_i + \alpha_i, & \text{otherwise} \end{cases}. \tag{3}$$

Because biases are always defined with respect to a conventional reference, at least one time series must be considered as conventionally unbiased to avoid an ill-posed equation system. Hence, in the following, the first time series $\mathbf{y}_1$ will be considered as conventionally unbiased.

The linear parametric functional model can be expressed using matrix algebra:

$$\mathrm{E}(\mathbf{y}) = \mathbf{A}\mathbf{x} = \begin{bmatrix} \mathbf{A}_{\boldsymbol{h}} & \mathbf{A}_{\boldsymbol{\alpha}} & \mathbf{A}_{\boldsymbol{\beta}} \end{bmatrix} \begin{bmatrix} \mathbf{h} \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{bmatrix}. \tag{4}$$

with $\mathbf{h} = [h_1 \cdots h_k]^T$, the combined solution vector, $\boldsymbol{\alpha} = [\alpha_2 \cdots \alpha_p]^T$, the intercepts parameter vector, and $\boldsymbol{\beta} = [\beta_2 \cdots \beta_p]^T$, the scale error parameter vector.

The combination design $pk \times k$ matrix $\mathbf{A}_{\boldsymbol{h}}$ corresponds to p stacked identity matrices $\mathbf{I}_{k\times k}$ such as

$$\mathbf{A}_{\boldsymbol{h}} = \begin{bmatrix} \mathbf{I}_{k\times k} \\ \vdots \\ \mathbf{I}_{k\times k} \end{bmatrix},$$

and both the intercept design $pk \times (p-1)$ matrix $\mathbf{A}_{\boldsymbol{\alpha}}$ and the scale error design $pk \times (p-1)$ matrix $\mathbf{A}_{\boldsymbol{\beta}}$ are constituted with block non-zeros vectors such as $\mathbf{A}_{\boldsymbol{\alpha}}$ reads

$$\mathbf{A}_{\boldsymbol{\alpha}} = \begin{bmatrix} \mathbf{0}_{k\times 1} & \cdots & \cdots & \mathbf{0}_{k\times 1} \\ \mathbf{1}_{k\times 1} & \ddots & & \vdots \\ \mathbf{0}_{k\times 1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{k\times 1} \\ \mathbf{0}_{k\times 1} & \cdots & \mathbf{0}_{k\times 1} & \mathbf{1}_{k\times 1} \end{bmatrix},$$

and $\mathbf{A}_{\boldsymbol{\beta}}$ follows

$$\mathbf{A}_{\boldsymbol{\beta}} = \begin{bmatrix} \mathbf{0}_{k\times 1} & \cdots & \cdots & \mathbf{0}_{k\times 1} \\ \mathbf{y}_2 & \ddots & & \vdots \\ \mathbf{0}_{k\times 1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{k\times 1} \\ \mathbf{0}_{k\times 1} & \cdots & \mathbf{0}_{k\times 1} & \mathbf{y}_p \end{bmatrix},$$

where $\mathbf{0}_{k\times 1}$, $\mathbf{1}_{k\times 1}$ refer to $k \times 1$ vectors respectively filled with zeros and ones.

2) STOCHASTIC MODEL FOR THE OBSERVATIONS

The stochastic model describes the variance var(.) of the observation vector $\mathbf{y}$. Considering that all measurements are statistically independent and that the uncertainty of the i-th instrument follows a multivariate normal distribution with a precision σ_i^2, the $pk \times pk$ co-variance matrix of the observations $\text{var}(\mathbf{y}) = \mathbf{Q}_{\mathbf{y}}$ reads :

$$\mathbf{Q}_{\mathbf{y}} = \begin{bmatrix} \sigma_1^2\mathbf{I}_{k\times k} & \mathbf{0}_{k\times k} & \cdots & \cdots & \mathbf{0}_{k\times k} \\ \mathbf{0}_{k\times k} & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \sigma_i^2\mathbf{I}_{k\times k} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \mathbf{0}_{k\times k} \\ \mathbf{0}_{k\times k} & \cdots & \cdots & \mathbf{0}_{k\times k} & \sigma_p^2\mathbf{I}_{k\times k} \end{bmatrix}, \tag{5}$$

where $\mathbf{I}_{k\times k}$ and $\mathbf{0}_{k\times k}$ are respectively the $k \times k$ identity and null matrices.

To use the LS-VCE method, $\mathbf{Q}_{\mathbf{y}}$ needs to be expressed as a linear combination of cofactor matrices $\mathbf{Q}_i$ such as

$$\mathbf{Q}_{\mathbf{y}} = \sigma_1^2 \cdot \mathbf{Q}_1 + \ldots + \sigma_p^2 \cdot \mathbf{Q}_p = \sum_{i=1}^{p} \sigma_i^2 \cdot \mathbf{Q}_i, \tag{6}$$

where the σ_i^2 are also referred to as variance components, and correspond to the instrument uncertainties. The $\mathbf{Q}_i$ are known $pk \times pk$ linearly independent cofactor matrices such as

$$\mathbf{Q}_i = \begin{bmatrix} \mathbf{0}_{k\times k} & \cdots & & & & \cdots & \mathbf{0}_{k\times k} \\ \vdots & \ddots & & & & & \vdots \\ & & \mathbf{0}_{k\times k} & & & & \\ & & & \mathbf{I}_{k\times k} & & & \\ & & & & \mathbf{0}_{k\times k} & & \\ \vdots & & & & & \ddots & \vdots \\ \mathbf{0}_{k\times k} & \cdots & & & & \cdots & \mathbf{0}_{k\times k} \end{bmatrix}.$$

3) LEAST-SQUARES ESTIMATION

According to the least-squares estimation theory, for normally distributed observations, a minimum variance estimation of the $q \times 1$ parameter vector $\mathbf{x}$ can be achieved by solving a normal equation system $\mathbf{N}\mathbf{x} = \mathbf{c}$ where $\mathbf{N}$ is the normal $q \times q$ matrix defined by $\mathbf{N} = \mathbf{A}^T\mathbf{Q}_\mathbf{y}^{-1}\mathbf{A}$ and $\mathbf{c}$ is a $q \times 1$ vector defined by $\mathbf{c} = \mathbf{A}^T\mathbf{Q}_\mathbf{y}^{-1}\mathbf{y}$ (Caspary et al. 1987; Teunissen 2000). Hence, the unbiased and minimum variance estimator of the functional parameter $\hat{\mathbf{x}}$ is given by

$$\hat{\mathbf{x}} = \mathbf{N}^{-1}\mathbf{c} = (\mathbf{A}^T\mathbf{Q}_\mathbf{y}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{Q}_\mathbf{y}^{-1}\mathbf{y}, \tag{7}$$

and its co-variance matrix $\mathbf{Q}_{\hat{\mathbf{x}}}$ follows

$$\mathbf{Q}_{\hat{\mathbf{x}}} = \mathbf{N}^{-1} = (\mathbf{A}^T\mathbf{Q}_\mathbf{y}^{-1}\mathbf{A})^{-1}. \tag{8}$$

In the case of a lack of knowledge on the on-site variance of the tide gauges, i.e., on $\mathbf{Q}_\mathbf{y}$, a variance component estimation method can be used to assess the uncertainty of each gauge. As

the minimum variance property of least-squares estimates requires a realistic weighting between sea level time series, the use of a variance component estimation method also allows for more realistic estimates of the parameter vector $\hat{\mathbf{x}}$ and its co-variance matrix $\mathbf{Q}_{\hat{\mathbf{x}}}$.

4) Least-Squares Variance Component Estimation

A review of most variance components estimation methods can be found in Fotopoulos (2003) and Amiri-Simkooei (2007). Here, we consider the application of the Least-Squares-Variance Components Estimation (LS-VCE), which is based on the same least-squares estimation principles used in sub-section 3. LS-VCE was first introduced in 1988 by Teunissen (1988) and further developed by Amiri-Simkooei (2007) and Teunissen and Amiri-Simkooei (2008). Under the hypothesis of the multivariate normal distribution considered in section 2, the method provides an unbiased and minimum variance estimator of the variance components. The method also allows for a direct derivation of the uncertainty of each variance component estimate.

The LS-VCE consists in using the redundancy of information of a system to infer the variance of the observations. In the case of a linear parametric functional model, one can compute a residual $pk \times 1$ vector $\hat{\mathbf{e}}$ such as

$$\hat{\mathbf{e}} = \mathbf{y} - \mathbf{A}\hat{\mathbf{x}} = \mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{y}, \tag{9}$$

where $\mathbf{P}_{\mathbf{A}}^{\perp}$ is a projector matrix defined by

$$\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{Q}_{\mathbf{y}}^{-1}. \tag{10}$$

The residual vector $\hat{\mathbf{e}}$ gives pieces of information about observation quality, potential model miss-specifications, and the presence of outliers. By assuming the absence of outliers and func-

tional model miss-specifications, the LS-VCE provides an estimator of the observation precisions using $\hat{\mathbf{e}}$ and $\mathbf{P}_{\mathbf{A}}^{\perp}$.

As for the standard least-squares estimation, the LS-VCE method estimates the unknown variance components $p \times 1$ vector $\hat{\boldsymbol{\sigma}}^{\mathbf{2}} = \left[\hat{\sigma}_1^2 \cdots \hat{\sigma}_p^2\right]^T$ by solving a normal equations system:

$$\hat{\boldsymbol{\sigma}}^{\mathbf{2}} = \bar{\mathbf{N}}^{-1}\bar{\mathbf{c}} = \begin{bmatrix} \bar{n}_{11} & \cdots & & \cdots & \bar{n}_{1p} \\ \vdots & \ddots & & & \vdots \\ & & \bar{n}_{ij} & & \\ \vdots & & & \ddots & \vdots \\ \bar{n}_{p1} & \cdots & & \cdots & \bar{n}_{pp} \end{bmatrix}^{-1} \begin{bmatrix} \bar{c}_1 \\ \vdots \\ \bar{c}_i \\ \vdots \\ \bar{c}_p \end{bmatrix}, \tag{11}$$

where the normal matrix $\bar{\mathbf{N}}$ and the vector $\bar{\mathbf{c}}$ are specific to the stochastic model, and thus different from the normal matrix $\mathbf{N}$ and vector $\mathbf{c}$ in equation (7).

For the stochastic model defined in sub-section 2, for which all variance components are to be estimated, the elements $\bar{n}_{ij}$ and $\bar{c}_i$ of $\bar{\mathbf{N}}$ and $\bar{\mathbf{c}}$ are defined by (Amiri-Simkooei 2007) :

$$\bar{n}_{ij} = \frac{1}{2}\mathrm{tr}(\mathbf{Q}_i\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{Q}_j\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{P}_{\mathbf{A}}^{\perp}) \tag{12}$$

$$\bar{c}_i = \frac{1}{2}(\hat{\mathbf{e}}^T\mathbf{Q}_{\mathbf{y}}^{-1}\mathbf{Q}_i\mathbf{Q}_{\mathbf{y}}^{-1}\hat{\mathbf{e}}) \tag{13}$$

where $\mathrm{tr}(.)$ stands for the trace operator.

Note that $\hat{\boldsymbol{\sigma}}^{\mathbf{2}}$ is involved in the definition of $\bar{n}_{ij}$ and $\bar{c}_i$ through $\mathbf{Q}_{\mathbf{y}}^{-1}$. Hence, equation (11) expresses $\hat{\boldsymbol{\sigma}}^{\mathbf{2}}$ as a function of $\mathbf{Q}_{\mathbf{y}}$, which is already a function of $\hat{\boldsymbol{\sigma}}^{\mathbf{2}}$ in equation (6). Such system of equations, where the equations for the unknowns include functions of the unknowns, can be numerically solved using an iterative procedure starting with an initial guess on the unknowns: the prior variance component vector $\boldsymbol{\sigma}_{\mathbf{0}}^{\mathbf{2}}$.

The first iteration consists in using the prior vector $\boldsymbol{\sigma}_{\mathbf{0}}^{\mathbf{2}}$ and cofactor matrices $\mathbf{Q}_i$ to compute $\mathbf{Q}_{\mathbf{y}}$ and then $\mathbf{P}_{\mathbf{A}}^{\perp}$, which are necessary to build the normal equations system (11). Solving this normal equations system (11) leads to the estimation of an updated variance component vector $\boldsymbol{\sigma}_{\mathbf{1}}^{\mathbf{2}}$. The next n iterations consist in successively updating the variance component vector $\hat{\boldsymbol{\sigma}}_{\boldsymbol{n}}^{\mathbf{2}}$ by solving the normal equations system (11) built using the previously estimated variance component vector $\hat{\boldsymbol{\sigma}}_{\boldsymbol{n-1}}^{\mathbf{2}}$. The iterations stop when the difference between two estimated variance component vectors becomes negligible. To obtain more details on the implementation of the LS-VCE method, a symbolic algorithm can be found in Figure 4.2 of Amiri-Simkooei (2007).

When encountering a convergence issue with an arbitrary prior variance component vector, using more realistic prior tide gauge uncertainties may be necessary. One could, for example, use the information provided by the tide gauge manufacturers. In the case of convergence, changes in prior variance components should not change the final LS-VCE results.

Once convergence is achieved, an insight into the quality the variance component estimates $\hat{\boldsymbol{\sigma}}^{\mathbf{2}}$ – the co-variance matrix of the variance component estimates – can be obtained by inverting the normal matrix $\bar{\mathbf{N}}$:

$$\mathbf{Q}_{\hat{\boldsymbol{\sigma}}^{\mathbf{2}}} = \bar{\mathbf{N}}^{-\mathbf{1}}, \tag{14}$$

The i-th diagonal element of $\mathbf{Q}_{\hat{\boldsymbol{\sigma}}^{\mathbf{2}}}$ corresponds to the variance of the i-th variance component $\sigma^2_{\hat{\sigma}_i^2}$. As for $\mathbf{Q}_{\hat{\mathbf{x}}}$, the uncertainties of variance component estimates depend on the system redundancy and the precision of the observations.

To get interpretable variance component estimates, one can change variance components $\hat{\sigma}_i^2$ into standard deviation components $\hat{\sigma}_i = \sqrt{\hat{\sigma}_i^2}$. To obtain variance component uncertainties with interpretable units, one can follow Amiri-Simkooei et al. (2009), and approximate the new variance of

the standard deviation component $\sigma^2_{\hat{\sigma}_i}$ by applying variance propagation law through the linearized square root function:

$$\sigma^2_{\hat{\sigma}_i} \approx \sigma^2_{\hat{\sigma}_i^2} \cdot (\frac{1}{2\sqrt{\hat{\sigma}_i^2}})^2, \tag{15}$$

The more interpretable standard deviation of the standard deviation component $\sigma_{\hat{\sigma}_i} = \sqrt{\sigma^2_{\hat{\sigma}_i}}$ can then be derived by taking the square root of both sides of equation (15), which gives:

$$\sigma_{\hat{\sigma}_i} \approx \frac{\sigma_{\hat{\sigma}_i^2}}{2\hat{\sigma}_i}, \tag{16}$$

where $\sigma_{\hat{\sigma}_i^2}$ is the standard deviation of the i-th variance component $\sigma_{\hat{\sigma}_i^2} = \sqrt{\sigma^2_{\hat{\sigma}_i^2}}$.

Hence, one can express the uncertainty estimate of the i-th tide gauge as $\hat{\sigma}_i \pm \sigma_{\hat{\sigma}_i}$ (cm).

4. Results

To compare COMB and DIFF methods on a similar basis, the PROBE time series has been considered conventionally unbiased for both methods.

To remove the influence of potential outliers, residuals time series were computed using equation (9) before the actual processing of both methods. The functional model (4) and the co-variance matrix $\mathbf{Q_y} = \mathbf{I}$ were considered in equation (10). Observations that showed residuals above five times the median absolute deviation of the gauge residual time series were removed from the data-set. In practice, it concerned less than 2 observations by time series.

a. Calibration with the combination (COMB) method

Before the assessment of the unknown bias parameters and the combined solution, a realistic co-variance matrix $\mathbf{Q_y}$ was first computed using the LS-VCE method. An arbitrary standard deviation

of 0.8 cm for all time series was used to build the prior variance component vector. Starting with $\boldsymbol{\sigma}_{\mathbf{0}}^{\mathbf{2}}$, the iterative procedure, summarized in section 4 and fully described in (Amiri-Simkooei 2007), provided the final variance components vector estimate $\hat{\boldsymbol{\sigma}}^{\mathbf{2}}$ and its co-variance matrix $\mathbf{Q}_{\hat{\boldsymbol{\sigma}}^{\mathbf{2}}}$. As the elements of both $\hat{\boldsymbol{\sigma}}^{\mathbf{2}}$ and $\mathbf{Q}_{\hat{\boldsymbol{\sigma}}^{\mathbf{2}}}$ are not directly interpretable, the equation (16) was used to express each tide gauge uncertainty estimate as $\hat{\sigma}_i \pm \sigma_{\hat{\sigma}_i}$ (cm).

Realistic bias parameters and combined solution were estimated by solving the functional model (4) using the variance component estimates: $\hat{\boldsymbol{\sigma}}^{\mathbf{2}}$ was substituted in equations (7) and (8) through equation (6), which led to the estimation of the unknown vector $\hat{\mathbf{x}}$ and its co-variance matrix $\mathbf{Q}_{\hat{\mathbf{x}}}$.

Both estimated sea level bias parameters and uncertainties for 10 min records are given, in centimeter, in Table 1. The electrical PROBE is found to be the most precise gauge in this experiment, with an uncertainty of 0.3 cm. The least precise tide gauges are the tide pole POLE (1.23 cm) and the BUOY1 (1.25 cm). BUOY1 is nearly two times less precise than BUOY2 (0.74 cm).

In Table 1, 4 time series – RADAR, LASER, BUOY1, and BUOY2 – show intercept estimates $\hat{\alpha}_i$ significant at the $3\sigma_{\hat{\alpha}_i}$ – or 99% – confidence level. Their amplitudes range from -1.87 cm (RADAR) to -4.30 cm (BUOY1). For the scale errors $\hat{\beta}_i$, only RADAR and POLE show estimates above $3\sigma_{\hat{\beta}_i}$, with about 0.5 cm m^{-1} and -0.3 cm m^{-1} respectively.

Residual time series of each tide gauge are presented in Figure 5. BUOY1 exhibits a mean shift of about -2 cm between 07:20 and 09:40. This artifact appears in the residual time series because it cannot result from the combination model. It means that the other tide gauges did not observe such a shift, otherwise, it would have been modeled by the combined solution. The presence of this artifact in the BUOY1's residual time series lowers its precision in Table 1. For the other gauges, no clear pattern appears in the residual time series, which suggests that their biases are correctly modeled.

The combined solution $\hat{h}$ and its uncertainty $\sigma_{\hat{h}}$ are presented in Figure 6. Each missing value in one of the time series increases the uncertainty of the combined solution to an extent proportional to its precision. The available measurements are displayed for each tide gauge, in the bottom of Figure 6. When the most precise tide gauge (PROBE) is not recording, between 10:00 and 12:10, the uncertainty $\sigma_{\hat{h}}$ of the combined solution increases by almost a factor of two. Despite the missing values of PROBE, the combined solution is estimated for the entire experiment period because all available observations are taken into account.

To investigate whether PROBE is found to be the most precise gauge because it is the conventionally unbiased gauge, the calibration has been reprocessed by instead considering BUOY1 as conventionally unbiased. The alternative calibration results are presented in Table 2. The choice of another conventionally unbiased gauge does not change uncertainty estimates but changes bias parameter estimates and their uncertainties. Bias parameters are the most affected because because they intrinsically depend on the definition of a convention. As BUOY1 does not exhibit any scale error in Table 1, the changes in scale error estimates in Table 2 are not dramatic. The sea level time series uncertainty estimates are identical in both alternatives because all biases are considered in each case. An alternative functional model ignoring an existing bias would not have provided identical results.

b. Comparison with the difference (DIFF) method

Using PROBE as the reference tide gauge, we plotted the VdC diagram for RADAR, POLE, LASER, BUOY1, and BUOY2. A linear regression on each diagram provided intercept and scale error estimates for each gauge. The DIFF method estimates are presented in Table 3. The differences with the COMB method estimates are summarized in Table 4.

The deviations between the COMB and DIFF methods reach 0.75 cm for the intercepts (BUOY1) and 0.15 cm m^{-1} for the scale errors (BUOY1). In Table 5, the changes in bias uncertainty between methods are expressed in terms of percentage of bias uncertainty reduction. The DIFF method provides slightly different results from the COMB method because it only considers a smaller subset of the data-set for each pair of gauge and because it does not take into account the precision of each time series. In this study, the DIFF method can only take into account the overlapping observation PROBE and the tested gauges. Given that PROBE has no observation between 10:00 and 12:10, the DIFF method ignores several observations, which deteriorates the precision of bias estimates. As a consequence, Table 5 shows that the COMB method provides 30% to 55% smaller uncertainties than the DIFF method for bias parameter estimates.

The presence of the scale error induces a height-dependency of both sea level bias models and their confidence intervals. To illustrate this, Figure 7 displays the estimated sea level bias models and their uncertainties, obtained with both methods, on the VdC diagram for BUOY1, which is the time series with the most substantial differences between the two models. At the lowest tide, sea level bias models obtained with COMB and DIFF method differs of about 3 millimeters. Besides, both sea level bias models are more precise around the mean tide than near the tidal extrema. As a consequence, the combined solution of the COMB method is also less precise near the tidal extrema, which results in the few millimeter changes for $\sigma_{\hat{h}}$ that also appears in Figure 6 at lowest tide: between 10:00 and 12:10.

A representation of all bias estimates obtained with both DIFF and COMB methods is given in Figure 8. Bias estimates are shown as points in the bias parameter space - intercept vs scale error. Their uncertainties appear as 1σ confidence ellipses. The correlations between bias parameters, always around -0.9, induce an inclination of the ellipses. As the cause of the correlation is the same – same signal and same bias model – for every time series, so are the inclinations in Figure 8. The

figure shows that the COMB method globally agrees with the DIFF method for bias detection while providing smaller confidence ellipses and thus, more precise bias parameter estimates.

5. Discussion

a. Performance of the tide gauges

The PROBE time series is twice more precise than that of the next most precise tide gauge. Its good performance results probably from the use of the stilling pipe, which stabilizes the water level and allows accurate readings on the measuring tape. This result comforts the use of electrical probes as references in tide gauge calibration campaigns. The results also show that RADAR, LASER, and BUOY2 uncertainty estimates are below the centimeter level, which confirms that they could provide sea level records with the level of accuracy specified by the IOC with a confidence level of more than 67% if they were not affected by biases.

Among the 6 tested gauges in this work, only two, of which one automatic gauge, present an uncertainty above 1 cm: POLE (1.23 cm) and BUOY1 (1.25 cm). The 1.23 cm uncertainty of POLE might result from the limitation of human eye reading on the 10 cm graduations. The lower performance of BUOY1 compared to BUOY2 is assigned to the presence of the artifact between 07:20 and 09:40. Considering its floating structure is less stable than the more recent model BUOY2, this artifact could be due to the buoy instability in the presence of currents during the ebb tide. BUOY2 did not measure when BUOY1 observed the artifact; one cannot exclude that the artifact is due to a miss-modeling of the GNSS data.

b. Nature of the biases

Separating instrumental and environmental parts of bias estimates is difficult, especially when the gauges are not fully co-located. We can nonetheless draw some hypotheses for bias attribution.

Usually, significant intercept estimates are caused by instrumental height errors. But in this experiment, other explanations are plausible for BUOY1, BUOY2, LASER, and RADAR.

BUOY1 and BUOY2 show similar intercept estimates while being deployed a few tens of meters away from the ground-based instruments. Hence, changes in the dynamic topography due to currents likely impacted their intercept estimates Pérez et al. (2014). In that case, an environmental effects is detected, not instrumental biases.

As LASER is not dedicated to water surface measurements, the intercept estimate is likely caused by a few centimeters penetration of the laser beam into the water. More appropriate laser systems have already been developed, using floating mirrors (MacAulay et al. 2008).

For RADAR, the significant intercept estimate likely results from not an instrumental height error and the influence of the significant scale error.

Theoretically, LASER, RADAR, and POLE could show scale error estimates in the case of range overestimation issues due to vertical alignment defaults. This is a plausible cause for RADAR and LASER. As the vertical alignment of POLE can be considered as reliable, the human-reading is the most likely source of its scale error.

Even though the nature of significant bias parameters α_i and β_i could remain unclear, one can still obtain corrected sea level time series by subtracting the bias model $\beta_i \times y_i(t) + \alpha_i$ to the measured sea level $y_i(t)$.

c. Improvement over difference based methods

The proposed calibration method provides an unbiased and minimum variance estimate of the tide gauge uncertainties, their sea level biases, and the combined solution from all times series. The variance of all estimates, including tide gauge uncertainties, are also determined. Thus, the COMB method leads to a more complete tide gauge calibration than the DIFF method.

The application to the Aix Island experiment revealed that the proposed methodology also leads to more precise bias estimates. This improvement is attributed to the combination of all available observations along with the realistic weighting between each gauge. The drastic precision improvement, from 30% to 55% on the uncertainty of the bias parameters, mostly shows that this method is more robust to the missing values of the most precise time series (PROBE), which is used as a reference to build the VdC diagrams.

For comparison purposes, the study considers only one conventionally unbiased time series. However, the COMB method allows using several unbiased time series and partially unbiased time series at the same time, which is not possible with the DIFF method. Adding unbiased time series should further improve the results of the COMB method.

6. Conclusion

The present contribution proposes a method for the cross-calibration of tide gauges. Based on the combination of multiple co-located time series, it takes advantage of the Least-Squares Variance Component Estimation method to assess both instrumental biases and measurement uncertainties in real conditions. The method was applied to a multi-instrument experiment carried out at Aix Island in 2016. Six instruments were deployed and performed simultaneous sea level recordings for 11 hours, with a 10 minutes sampling.

The electrical probe was found to be two to four times more precise than the other gauges. RADAR, LASER, and BUOY2 uncertainty estimates are below the centimeter level, which confirms that, in those conditions, they could provide sea level records with the level of accuracy specified by the IOC if they were not affected by biases. We showed that, within our time series, significant bias parameters were found for all the tested gauges. Hence, this study shows that it is possible to assess both the biases and the precision – i.e. the full accuracy – for each gauge.

The results obtained with the combination method have been compared to that of a difference based method. It showed that the combination of all time series provides more precise bias estimates.

Because this study is based on an 11 hours experiment, time-dependent biases and random errors have not been considered. Further studies using the COMB methods are thus necessary to investigate the time dependency of sea level bias parameters and tide gauge precisions.

Acknowledgments. This study has been financially supported by the Direction Générale de l'Armement (DGA), the Nouvelle-Aquitaine region, and by the Centre National des Études Spatiales (CNES) as an application of the geodesy missions. The authors are thankful to the other participants of the Aix island experiment : Laurent Testut (CNAP-LEGOS), Médéric Gravelle (CNRS-LIENSs), Álvaro Santamaría-Gómez (CNAP-OMP), Mikaël Guichard (ULR-LIENSs), Elizabeth Prouteau (CNRS-LIENSs), Pierre Chasseloup (SONEL-LIENSs), Cyril Poitevin (LIENSs), Fabien Durand (IRD-LEGOS), Ronan Le Gall (SHOM), and Pascal Le Dû (SHOM). The authors gratefully acknowledge the three anonymous reviewers, whose comments helped improve and clarify this paper.

References

Abbondanza, C., Z. Altamimi, T. Chin, R. Gross, M. Heflin, J. Parker, and X. Wu, 2015: Three-corner hat for the assessment of the uncertainty of non-linear residuals of space-geodetic time series in the context of terrestrial reference frame analysis. *Journal of Geodesy*, **89 (4)**, 313–329.

Amiri-Simkooei, A., 2007: Least-squares variance component estimation: theory and gps applications. Ph.D. thesis, TU Delft, Delft University of Technology.

Amiri-Simkooei, A., P. Teunissen, and C. Tiberius, 2009: Application of least-squares variance component estimation to gps observables. *Journal of Surveying Engineering*, **135 (4)**, 149–160.

André, G., B. M. Míguez, V. Ballu, L. Testut, and G. Wöppelmann, 2013: Measuring sea level with gps-equipped buoys: A multi-instruments experiment at aix island. *The International Hydrographic Review*, **(10)**.

BIPM, I., I. IFcc, I. ISO, and O. IUPAP, 2008: International vocabulary of metrology–basic and general concepts and associated terms, 2008. *JcGM*, **200**, 99–12.

Cartwright, D. E., 2000: *Tides: a scientific history*. Cambridge University Press.

Caspary, W., J. Rüeger, and U. of New South Wales. School of Surveying, 1987: *Concepts of Network and Deformation Analysis*. Monograph (University of New South Wales. School of Surveying), School of Surveying, University of New South Wales.

de Viron, O., I. Panet, V. Mikhailov, M. Van Camp, and M. Diament, 2008: Retrieving earthquake signature in grace gravity solutions. *Geophysical Journal International*, **174 (1)**, 14–20.

Feissel-Vernier, M., O. de Viron, and K. Le Bail, 2007: Stability of vlbi, slr, doris, and gps positioning. *Earth, planets and space*, **59 (6)**, 475–497.

Fotopoulos, G., 2003: An analysis on the optimal combination of geoid, orthometric and ellipsoidal height data. Ph.D. thesis.

Gouriou, T., B. M. Míguez, and G. Wöppelmann, 2013: Reconstruction of a two-century long sea level record for the pertuis dantioche (france). *Continental Shelf Research*, **61**, 31–40.

Gray, J. E., and D. W. Allan, 1974: A method for estimating the frequency stability of an individual oscillator. *28th Annual Symposium on Frequency Control. 1974*, IEEE, 243–246.

IOC, 1985: *Manual on sea level measurement and interpretation. Volume I-Basic procedures.* UNESCO.

IOC, and Coauthors, 2002: *Manual on sea level measurement and interpretation. Volume III-Reappraisals and Recommendations as of the year 2000.* UNESCO.

IOC, and Coauthors, 2006: *Manual on sea level measurement and interpretation. Volume IV-An update to 2006.* UNESCO.

IOC, and Coauthors, 2012: Global sea level observing system (gloss): Implementation plan 2012. *Intergov. Oceanogr. Comm. Tech. Ser*, **100**.

IOC, and Coauthors, 2016: *Manual on sea level measurement and interpretation. Volume V-Radar gauges*. UNESCO.

Larson, K. M., R. D. Ray, and S. D. Williams, 2017: A 10-year comparison of water levels measured with a geodetic gps receiver versus a conventional tide gauge. *Journal of Atmospheric and Oceanic Technology*, **34 (2)**, 295–307.

Lennon, G., 1968: The evaluation of tide-gauge performance through the van de casteele test. *Cah. Oceanogr*, **20**, 867–877.

Lentz, S. J., 1993: The accuracy of tide-gauge measurements at subtidal frequencies. *Journal of atmospheric and oceanic technology*, **10 (2)**, 238–245.

MacAulay, P., C. OReilly, and K. Thompson, 2008: Atlantic canadas real-time water level system observations, predictions, forecasts and datums on the web.

Míguez Martín, B., R. Le Roy, and G. Wöppelmann, 2008a: The use of radar tide gauges to measure variations in sea level along the french coast. *Journal of coastal research*, **24 (sp3)**, 61–68.

Míguez Martín, B., L. Testut, and G. Wöppelmann, 2008b: The van de casteele test revisited: an efficient approach to tide gauge error characterization. *Journal of Atmospheric and Oceanic Technology*, **25 (7)**, 1238–1244.

Míguez Martín, B., L. Testut, and G. Wöppelmann, 2012: Performance of modern tide gauges: towards mm-level accuracy. *Scientia Marina*, **76 (S1)**, 221–228.

Pálinkáš, V., and Coauthors, 2017: Regional comparison of absolute gravimeters, euramet. mg-k2 key comparison. *Metrologia*, **54 (1A)**, 07 012.

Park, J., R. Heitsenrether, and W. Sweet, 2014: Water level and wave height estimates at noaa tide stations from acoustic and microwave sensors. *Journal of Atmospheric and Oceanic Technology*, **31 (10)**, 2294–2308.

Pérez, B., A. Payo, D. López, P. Woodworth, and E. A. Fanjul, 2014: Overlapping sea level time series measured using different technologies: an example from the redmar spanish network. *Natural Hazards and Earth System Sciences*, **14 (3)**, 589.

Pugh, D., and P. Woodworth, 2014: *Sea-level measuring systems*, 1735. Cambridge University Press, doi:10.1017/CBO9781139235778.005.

Pytharouli, S., S. Chaikalis, and S. C. Stiros, 2018: Uncertainty and bias in electronic tide-gauge records: Evidence from collocated sensors. *Measurement*, **125**, 496–508.

Takasu, T., 2013: Rtklib ver. 2.4. 2 manual. *RTKLIB: An Open Source Program Package for GNSS Positioning*, 29–49.

Teunissen, P., 2000: *Adjustment theory, Series on Mathematical geodesy and positioning*. Delft University Press.

Teunissen, P. J. G., 1988: Towards a least-squares framework for adjusting and testing of both functional and stochastic models.

Teunissen, P. J. G., and A. R. Amiri-Simkooei, 2008: Least-squares variance component estimation. *Journal of Geodesy*, **82 (2)**, 65–82, doi:10.1007/s00190-007-0157-x, URL https://doi.org/10.1007/s00190-007-0157-x.

Valty, P., O. de Viron, I. Panet, M. Van Camp, and J. Legrand, 2013: Assessing the precision in loading estimates by geodetic techniques in southern europe. *Geophysical Journal International*, **194 (3)**, 1441–1454.

Watson, C., R. Coleman, and R. Handsworth, 2008: Coastal tide gauge calibration: a case study at macquarie island using gps buoy techniques. *Journal of Coastal Research*, 1071–1079.

LIST OF TABLES

563 TABLE 1. Tide gauge cross-calibration results obtained using the COMB method. PROBE scale error and
564 intercept are conventionally set to zero.

Gauges	$\hat{\alpha}_i \pm \sigma_{\hat{\alpha}_i}$ (cm)	$\hat{\beta}_i \pm \sigma_{\hat{\beta}_i}$ (cm m^{-1})	$\hat{\sigma}_i \pm \sigma_{\hat{\sigma}_i}$ (cm)
RADAR	-1.87 $\pm$ 0.30	0.52 $\pm$ 0.07	0.81 $\pm$ 0.08
PROBE	.	.	0.31 $\pm$ 0.10
POLE	-0.13 $\pm$ 0.39	-0.32 $\pm$ 0.09	1.23 $\pm$ 0.12
BUOY1	-4.30 $\pm$ 0.41	0.00 $\pm$ 0.11	1.25 $\pm$ 0.14
LASER	-3.42 $\pm$ 0.35	0.13 $\pm$ 0.08	0.90 $\pm$ 0.10
BUOY2	-3.53 $\pm$ 0.30	0.17 $\pm$ 0.07	0.74 $\pm$ 0.09

565 TABLE 2. Alternative tide gauge cross-calibration results obtained using the COMB method and by defining
566 BUOY1 as the conventionally unbiased gauge. BUOY1 scale error and intercept are conventionally set to zero.

Gauges	$\hat{\alpha}_i \pm \sigma_{\hat{\alpha}_i}$ (cm)	$\hat{\beta}_i \pm \sigma_{\hat{\beta}_i}$ (cm m^{-1})	$\hat{\sigma}_i \pm \sigma_{\hat{\sigma}_i}$ (cm)
RADAR	2.34 ± 0.42	0.55 ± 0.11	0.81 ± 0.08
PROBE	4.18 ± 0.42	0.03 ± 0.11	0.31 ± 0.10
POLE	4.07 ± 0.49	-0.29 ± 0.13	1.22 ± 0.12
BUOY1	.	.	1.25 ± 0.14
LASER	0.72 ± 0.45	0.15 ± 0.12	0.90 ± 0.10
BUOY2	0.68 ± 0.42	0.19 ± 0.11	0.74 ± 0.09

TABLE 3. Tide gauge calibration results obtained using the DIFF method. PROBE is the reference gauge.

Gauges	$\hat{\alpha}_i \pm \sigma_{\hat{\alpha}_i}$ (cm)	$\hat{\beta}_i \pm \sigma_{\hat{\beta}_i}$ (cm m^{-1})
RADAR	-1.54 $\pm$ 0.47	0.42 $\pm$ 0.10
PROBE	.	.
POLE	0.09 $\pm$ 0.66	-0.36 $\pm$ 0.14
BUOY1	-5.05 $\pm$ 0.72	0.15 $\pm$ 0.18
LASER	-3.07 $\pm$ 0.77	0.12 $\pm$ 0.17
BUOY2	-3.42 $\pm$ 0.47	0.18 $\pm$ 0.10

TABLE 4. Difference between DIFF and COMB calibration results.

Gauges	$\Delta\hat{\alpha}_i$ (cm)	$\Delta\sigma_{\hat{\alpha}_i}$ (cm)	$\Delta\hat{\beta}_i$ (cm m^{-1})	$\Delta\sigma_{\hat{\beta}_i}$ (cm m^{-1})
RADAR	-0.33	-0.17	0.10	-0.03
PROBE	.	.	.	.
POLE	-0.22	-0.27	0.04	-0.05
BUOY1	0.75	-0.31	-0.15	-0.07
LASER	-0.35	-0.42	0.01	-0.09
BUOY2	-0.11	-0.17	-0.01	-0.03

TABLE 5. Reduction of the standard deviations of the bias parameters obtained using the COMB method with respect to the DIFF method.

Gauge	$\Delta\sigma_{\hat{\alpha}_i}$ (%)	$\Delta\sigma_{\hat{\beta}_i}$ (%)
RADAR	-36	-30
PROBE	.	.
POLE	-41	-36
BUOY1	-43	-39
LASER	-55	-53
BUOY2	-36	-30

LIST OF FIGURES

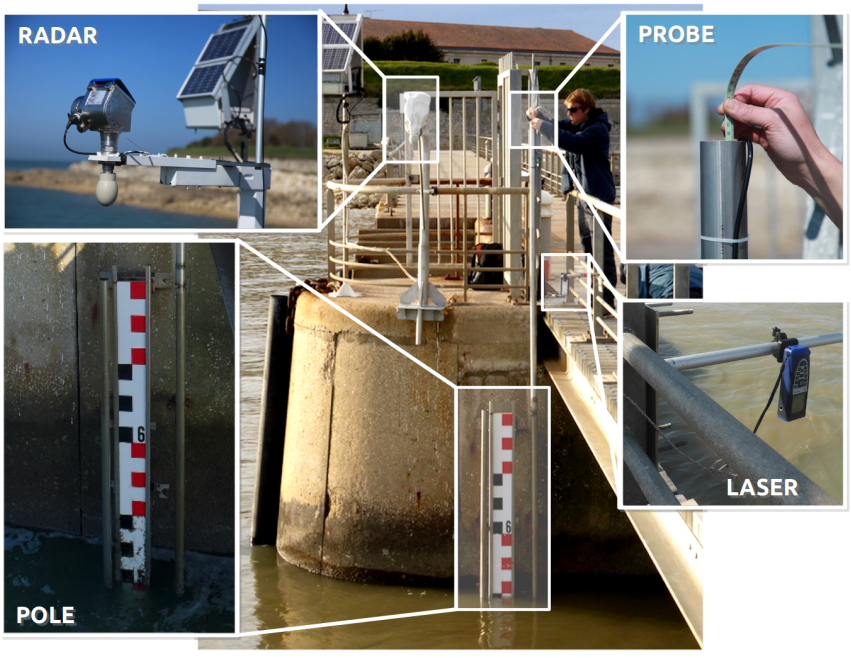


FIG. 1. The four ground-based tide gauges: RADAR, POLE, PROBE and LASER.

FIG. 2. The two GNSS buoys: BUOY1 and BUOY2.

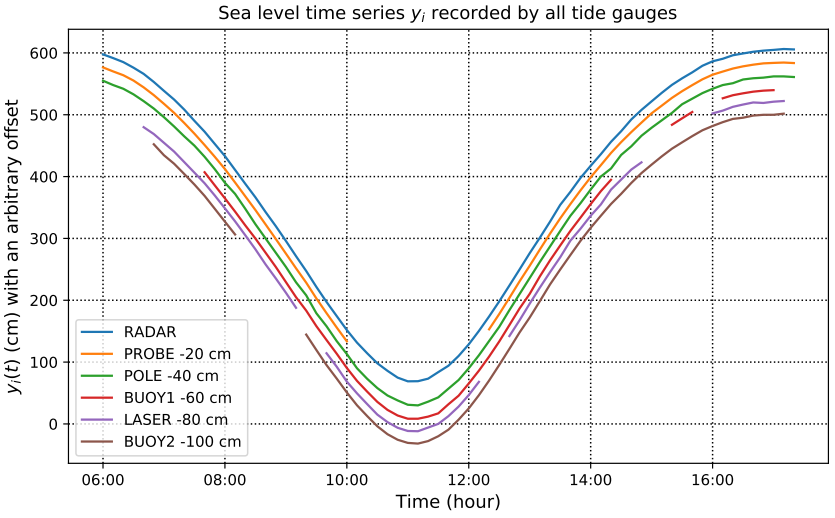


FIG. 3. Sea level time series y_i recorded by all tide gauges.

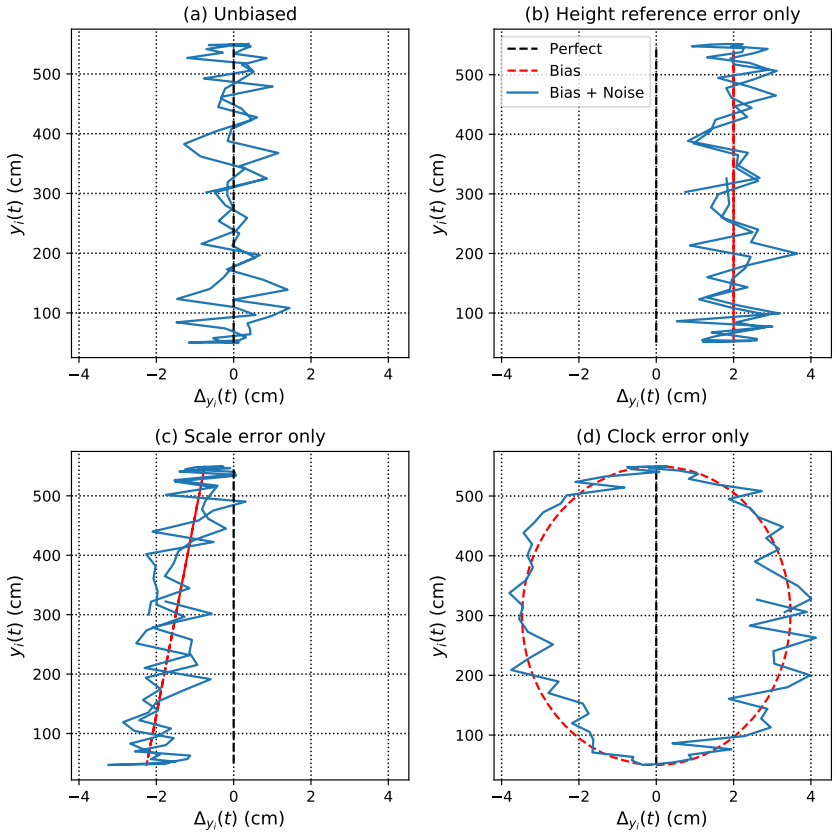


FIG. 4. Synthetic examples of Van de Casteele diagrams for the most common types of range measurement errors: (a) random measurement errors only; (b) random measurement errors and a height reference error; (c) random measurement errors and a scale error; (d) random measurement errors and a clock error.

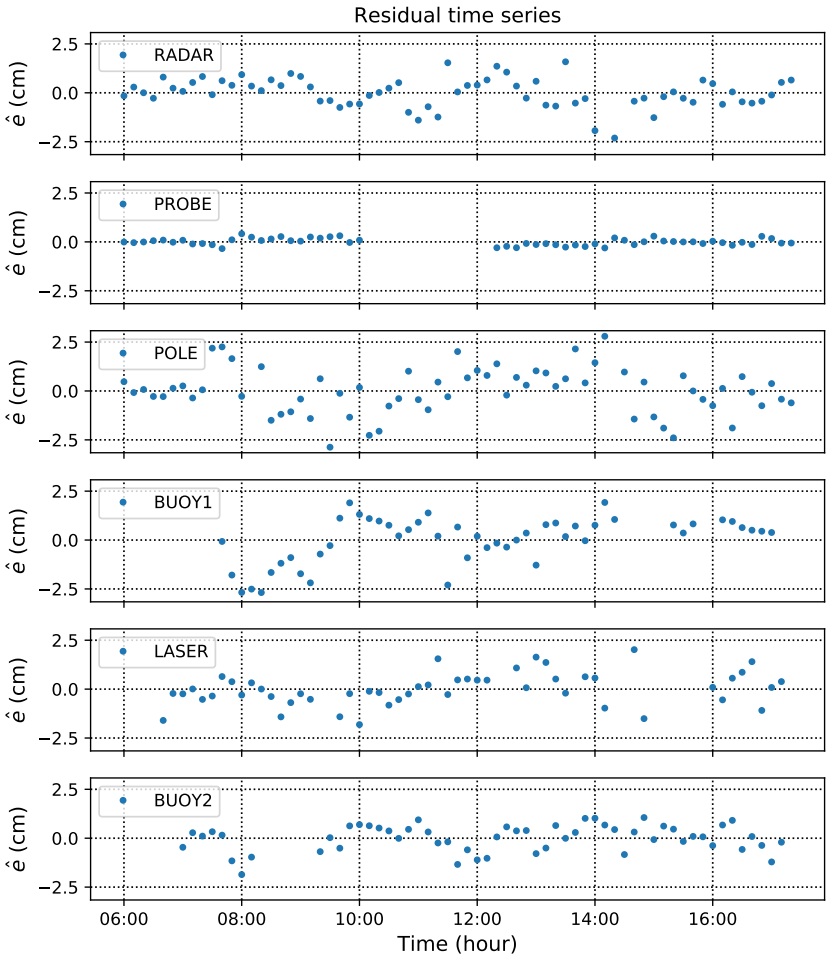


FIG. 5. Residual time series of each tide gauge for the estimated linear combination model.

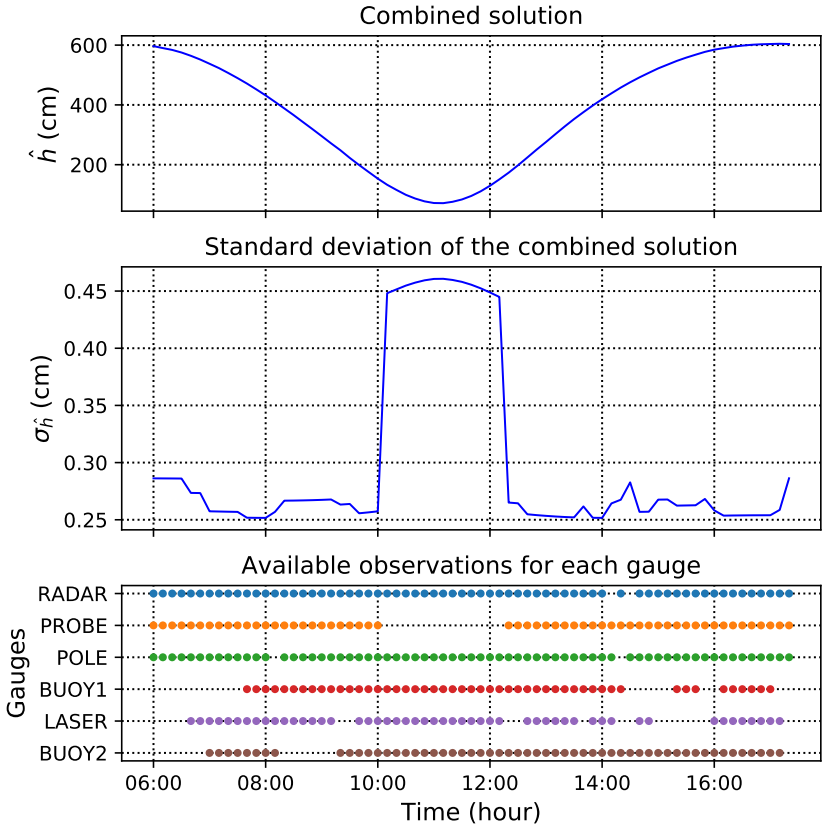
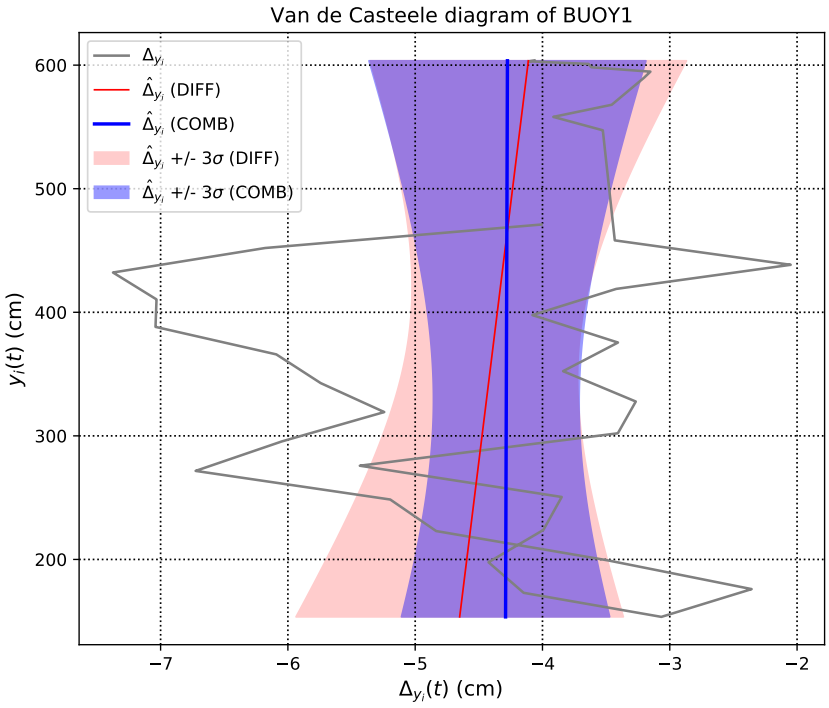
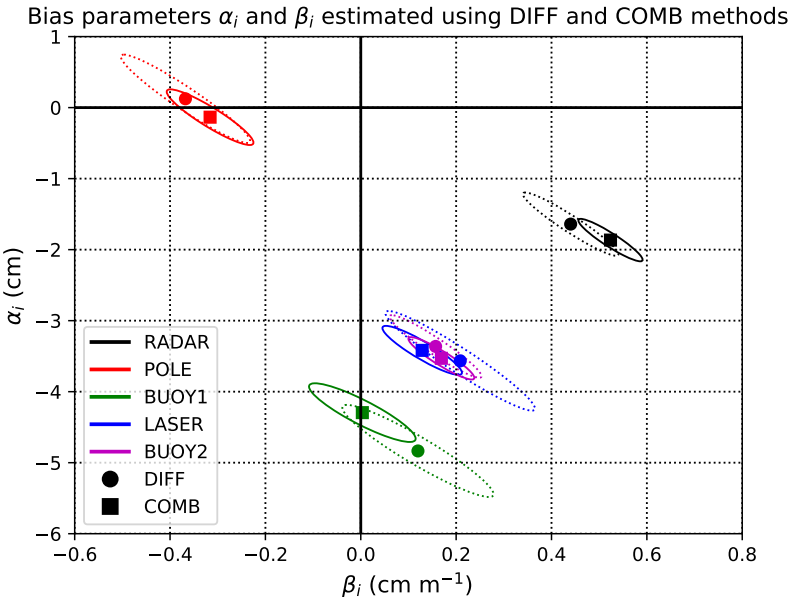


FIG. 6. Combined solution (top), the standard deviation of the combined solution (middle) and available observations for each gauge (bottom).



589 FIG. 7. Van de Casteele diagram of BUOY1. The sea level bias model estimated with the COMB method is
590 displayed in blue, and the one estimated with the DIFF method is displayed in red.



591 FIG. 8. Representation of the bias parameter estimates in the parameter space for both COMB and DIFF
592 methods.