



**HAL**  
open science

## 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

► **To cite this version:**

Kevin Gobron, Olivier de Viron, Guy Woppelmann, Etienne Poirier, Valérie Ballu, et al.. Assessment of Tide Gauge Biases and Precision by the Combination of Multiple Collocated Time Series. *Journal of Atmospheric and Oceanic Technology*, 2019, 36 (10), pp.1983-1996. 10.1175/JTECH-D-18-0235.1 . hal-02384242

**HAL Id: hal-02384242**

**<https://hal.science/hal-02384242v1>**

Submitted on 28 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

The following manuscript is a pre-print. The final version, published by the American Meteorological Society (AMS), is available at the following link:

<https://doi.org/10.1175/JTECH-D-18-0235.1>

Please use the following citation when referring to this work:

Gobron, K., de Viron, O., Wöppelmann, G., Poirier, É., Ballu, V., & Van Camp, M. (2019). Assessment of Tide Gauge Biases and Precision by the Combination of Multiple Collocated Time Series. *Journal of Atmospheric and Oceanic Technology*, 36(10), 1983-1996, <https://doi.org/10.1175/JTECH-D-18-0235.1>.

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

3                                     Kevin Gobron \*

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

5                                     Olivier de Viron

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

7                                     Guy Wöppelmann

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

9                                     Étienne Poirier

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

11                                    Valérie Ballu

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

13                                    Michel Van Camp

14                     *Royal Observatory of Belgium, Uccle, Belgium*

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

<sup>17</sup> E-mail: [kevin.gobron1@univ-lr.fr](mailto:kevin.gobron1@univ-lr.fr)

## ABSTRACT

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

## 30 **1. Introduction**

31 Tide gauges aim at measuring the vertical distance between the sea level and a reference level  
32 (or datum). Historically, tide gauges were first used for tide prediction and navigation (Cartwright  
33 2000); today, their applications have been extended (Pugh and Woodworth 2014). Clustered  
34 into networks of continuously operating stations, they are the key components of storm surge or  
35 tsunami warning systems and climate-related monitoring programs, such as the Global Sea Level  
36 Observing System (GLOSS) (IOC et al. 2012).

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

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

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

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

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

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

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

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

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

85 Approach (c) is classically used in metrology (Pálinkáš et al. 2017) and has often been used  
86 in geodesy through the three-cornered-hat (TCH) estimation method (Gray and Allan 1974), for  
87 example, to determine the stability of reference station positions (Feissel-Vernier et al. 2007; Ab-  
88 bondanza et al. 2015) or the precision of space gravity model (de Viron et al. 2008; Valtý et al.  
89 2013). The TCH is not the only possible implementation of the approach (c): the more gen-  
90 eral framework of Variance Component Estimation (VCE) can similarly address this problem, as  
91 shown by the theoretical example 4.10 of Amiri-Simkooei (2007). The TCH and VCE examples  
92 can separate the uncertainty of each gauge, but assume the absence of sea level biases.

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



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

## 103 **2. The Aix Island experiment**

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

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

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

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

120 The tide pole (POLE) is a permanent instrument made of a plastic staff with graduations every  
121 10 centimeters fixed vertically by a stainless steel structure (Figure 1). The operator estimates the  
122 sea level visually over the standard pre-defined 2 minutes acquisition period.

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

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

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

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

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

### 150 **3. Calibration methods**

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

#### 155 *a. Sea level error model*

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

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

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

164 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  
 165 centered normal distribution of unknown variance  $\sigma_i^2$ .

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

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

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

183 *b. Difference-based calibration methods (DIFF)*

184 Difference-based methods (DIFF) consist in analyzing the differences  $\Delta_{y_i}(t) = y_i(t) - y_{ref}(t)$   
185 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  
186 reference instrument  $y_{ref}(t) = h(t) + e_{ref}(t)$ .

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

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

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

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

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

203 *c. The combination-based calibration method (COMB)*

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

## 208 1) FUNCTIONAL MODEL OF THE OBSERVATIONS

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

## 231 2) STOCHASTIC MODEL FOR THE OBSERVATIONS

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

$$\mathbf{Q}_y = \begin{bmatrix} \sigma_1^2 \mathbf{1}_{k \times k} & \mathbf{0}_{k \times k} & \cdots & \cdots & \mathbf{0}_{k \times k} \\ \mathbf{0}_{k \times k} & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \sigma_i^2 \mathbf{1}_{k \times k} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \mathbf{0}_{k \times k} \\ \mathbf{0}_{k \times k} & \cdots & \cdots & \mathbf{0}_{k \times k} & \sigma_p^2 \mathbf{1}_{k \times k} \end{bmatrix}, \quad (5)$$

236 where  $\mathbf{1}_{k \times k}$  and  $\mathbf{0}_{k \times k}$  are respectively the  $k \times k$  identity and null matrices.

237 To use the LS-VCE method,  $\mathbf{Q}_y$  needs to be expressed as a linear combination of cofactor ma-  
 238 trices  $\mathbf{Q}_i$  such as

$$\mathbf{Q}_y = \sigma_1^2 \cdot \mathbf{Q}_1 + \dots + \sigma_p^2 \cdot \mathbf{Q}_p = \sum_{i=1}^p \sigma_i^2 \cdot \mathbf{Q}_i, \quad (6)$$



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

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

### 241 3) LEAST-SQUARES ESTIMATION

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

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

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

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

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

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

#### 253 4) LEAST-SQUARES VARIANCE COMPONENT ESTIMATION

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

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

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

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

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

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

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

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

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

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

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

$$\bar{n}_{ij} = \frac{1}{2} \text{tr}(\mathbf{Q}_i \mathbf{Q}_y^{-1} \mathbf{P}_A^\perp \mathbf{Q}_j \mathbf{Q}_y^{-1} \mathbf{P}_A^\perp) \quad (12)$$

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

276 where  $\text{tr}(\cdot)$  stands for the trace operator.

277 Note that  $\hat{\boldsymbol{\sigma}}^2$  is involved in the definition of  $\bar{n}_{ij}$  and  $\bar{c}_i$  through  $\mathbf{Q}_y^{-1}$ . Hence, equation (11)  
 278 expresses  $\hat{\boldsymbol{\sigma}}^2$  as a function of  $\mathbf{Q}_y$ , which is already a function of  $\hat{\boldsymbol{\sigma}}^2$  in equation (6). Such system  
 279 of equations, where the equations for the unknowns include functions of the unknowns, can be  
 280 numerically solved using an iterative procedure starting with an initial guess on the unknowns: the  
 281 prior variance component vector  $\boldsymbol{\sigma}_0^2$ .

282 The first iteration consists in using the prior vector  $\sigma_0^2$  and cofactor matrices  $Q_i$  to compute  $Q_y$   
 283 and then  $P_A^\perp$ , which are necessary to build the normal equations system (11). Solving this normal  
 284 equations system (11) leads to the estimation of an updated variance component vector  $\sigma_1^2$ . The  
 285 next  $n$  iterations consist in successively updating the variance component vector  $\hat{\sigma}_n^2$  by solving  
 286 the normal equations system (11) built using the previously estimated variance component vector  
 287  $\hat{\sigma}_{n-1}^2$ . The iterations stop when the difference between two estimated variance component vectors  
 288 becomes negligible. To obtain more details on the implementation of the LS-VCE method, a  
 289 symbolic algorithm can be found in Figure 4.2 of Amiri-Simkooei (2007).

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

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

$$Q_{\hat{\sigma}^2} = \bar{N}^{-1}, \quad (14)$$

297 The  $i$ -th diagonal element of  $Q_{\hat{\sigma}^2}$  corresponds to the variance of the  $i$ -th variance component  $\sigma_{\hat{\sigma}_i^2}^2$ .  
 298 As for  $Q_{\hat{x}}$ , the uncertainties of variance component estimates depend on the system redundancy  
 299 and the precision of the observations.

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

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

$$\sigma_{\hat{\sigma}_i}^2 \approx \sigma_{\hat{\sigma}_i^2}^2 \cdot \left( \frac{1}{2\sqrt{\hat{\sigma}_i^2}} \right)^2, \quad (15)$$

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

$$\sigma_{\hat{\sigma}_i} \approx \frac{\sigma_{\hat{\sigma}_i^2}}{2\hat{\sigma}_i}, \quad (16)$$

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

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

#### 309 4. Results

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

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

##### 317 *a. Calibration with the combination (COMB) method*

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

320 of 0.8 cm for all time series was used to build the prior variance component vector. Starting  
321 with  $\sigma_0^2$ , the iterative procedure, summarized in section 4 and fully described in (Amiri-Simkooei  
322 2007), provided the final variance components vector estimate  $\hat{\sigma}^2$  and its co-variance matrix  $\mathbf{Q}_{\hat{\sigma}^2}$ .  
323 As the elements of both  $\hat{\sigma}^2$  and  $\mathbf{Q}_{\hat{\sigma}^2}$  are not directly interpretable, the equation (16) was used to  
324 express each tide gauge uncertainty estimate as  $\hat{\sigma}_i \pm \sigma_{\hat{\sigma}_i}$  (cm).

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

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

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

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

343 The combined solution  $\hat{h}$  and its uncertainty  $\sigma_{\hat{h}}$  are presented in Figure 6. Each missing value in  
344 one of the time series increases the uncertainty of the combined solution to an extent proportional  
345 to its precision. The available measurements are displayed for each tide gauge, in the bottom  
346 of Figure 6. When the most precise tide gauge (PROBE) is not recording, between 10:00 and  
347 12:10, the uncertainty  $\sigma_{\hat{h}}$  of the combined solution increases by almost a factor of two. Despite  
348 the missing values of PROBE, the combined solution is estimated for the entire experiment period  
349 because all available observations are taken into account.

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

#### 360 *b. Comparison with the difference (DIFF) method*

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

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

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

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



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

## 391 **5. Discussion**

### 392 *a. Performance of the tide gauges*

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

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

### 408 *b. Nature of the biases*

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

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

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

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

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

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

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

### 429 *c. Improvement over difference based methods*

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

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

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

## 444 **6. Conclusion**

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

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

457 The results obtained with the combination method have been compared to that of a difference  
458 based method. It showed that the combination of all time series provides more precise bias esti-  
459 mates.

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

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

## 472 **References**

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

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

- 478 Amiri-Simkooei, A., P. Teunissen, and C. Tiberius, 2009: Application of least-squares variance  
479 component estimation to gps observables. *Journal of Surveying Engineering*, **135** (4), 149–160.
- 480 André, G., B. M. Míguez, V. Ballu, L. Testut, and G. Wöppelmann, 2013: Measuring sea level  
481 with gps-equipped buoys: A multi-instruments experiment at aix island. *The International Hy-*  
482 *drographic Review*, (10).
- 483 BIPM, I., I. IFcc, I. ISO, and O. IUPAP, 2008: International vocabulary of metrology–basic and  
484 general concepts and associated terms, 2008. *JcGM*, **200**, 99–12.
- 485 Cartwright, D. E., 2000: *Tides: a scientific history*. Cambridge University Press.
- 486 Caspary, W., J. Rüeger, and U. of New South Wales. School of Surveying, 1987: *Concepts of*  
487 *Network and Deformation Analysis*. Monograph (University of New South Wales. School of  
488 Surveying), School of Surveying, University of New South Wales.
- 489 de Viron, O., I. Panet, V. Mikhailov, M. Van Camp, and M. Diament, 2008: Retrieving earthquake  
490 signature in grace gravity solutions. *Geophysical Journal International*, **174** (1), 14–20.
- 491 Feissel-Vernier, M., O. de Viron, and K. Le Bail, 2007: Stability of vlbi, slr, doris, and gps posi-  
492 tioning. *Earth, planets and space*, **59** (6), 475–497.
- 493 Fotopoulos, G., 2003: An analysis on the optimal combination of geoid, orthometric and ellip-  
494 soidal height data. Ph.D. thesis.
- 495 Gouriou, T., B. M. Míguez, and G. Wöppelmann, 2013: Reconstruction of a two-century long sea  
496 level record for the pertuis dantioche (france). *Continental Shelf Research*, **61**, 31–40.
- 497 Gray, J. E., and D. W. Allan, 1974: A method for estimating the frequency stability of an individual  
498 oscillator. *28th Annual Symposium on Frequency Control*. 1974, IEEE, 243–246.

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

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

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

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

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

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

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

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

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

518 Míguez Martín, B., R. Le Roy, and G. Wöppelmann, 2008a: The use of radar tide gauges to  
519 measure variations in sea level along the french coast. *Journal of coastal research, 24 (sp3),*  
520 *61–68.*

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

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

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

528 Park, J., R. Heitsenrether, and W. Sweet, 2014: Water level and wave height estimates at noaa tide  
529 stations from acoustic and microwave sensors. *Journal of Atmospheric and Oceanic Technology*,  
530 **31 (10)**, 2294–2308.

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

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

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

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

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

- 542 Teunissen, P. J. G., 1988: Towards a least-squares framework for adjusting and testing of both  
543 functional and stochastic models.
- 544 Teunissen, P. J. G., and A. R. Amiri-Simkooei, 2008: Least-squares variance component es-  
545 timation. *Journal of Geodesy*, **82** (2), 65–82, doi:10.1007/s00190-007-0157-x, URL https:  
546 //doi.org/10.1007/s00190-007-0157-x.
- 547 Valty, P., O. de Viron, I. Panet, M. Van Camp, and J. Legrand, 2013: Assessing the precision  
548 in loading estimates by geodetic techniques in southern europe. *Geophysical Journal Interna-*  
549 *tional*, **194** (3), 1441–1454.
- 550 Watson, C., R. Coleman, and R. Handsworth, 2008: Coastal tide gauge calibration: a case study  
551 at macquarie island using gps buoy techniques. *Journal of Coastal Research*, 1071–1079.



552 **LIST OF TABLES**

553 **Table 1.** Tide gauge cross-calibration results obtained using the COMB method.  
554 PROBE scale error and intercept are conventionally set to zero. . . . . 32

555 **Table 2.** Alternative tide gauge cross-calibration results obtained using the COMB  
556 method and by defining BUOY1 as the conventionally unbiased gauge. BUOY1  
557 scale error and intercept are conventionally set to zero. . . . . 33

558 **Table 3.** Tide gauge calibration results obtained using the DIFF method. PROBE is the  
559 reference gauge. . . . . 34

560 **Table 4.** Difference between DIFF and COMB calibration results. . . . . 35

561 **Table 5.** Reduction of the standard deviations of the bias parameters obtained using the  
562 COMB method with respect to the DIFF method. . . . . 36

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

567 TABLE 5. Reduction of the standard deviations of the bias parameters obtained using the COMB method with  
 568 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

569 **LIST OF FIGURES**

570 **Fig. 1.** The four ground-based tide gauges: RADAR, POLE, PROBE and LASER. . . . . 38

571 **Fig. 2.** The two GNSS buoys: BUOY1 and BUOY2. . . . . 39

572 **Fig. 3.** Sea level time series  $y_i$  recorded by all tide gauges. . . . . 40

573 **Fig. 4.** Synthetic examples of Van de Castelee diagrams for the most common types of range mea-  
574 surement errors: (a) random measurement errors only; (b) random measurement errors and  
575 a height reference error; (c) random measurement errors and a scale error; (d) random mea-  
576 surement errors and a clock error. . . . . 41

577 **Fig. 5.** Residual time series of each tide gauge for the estimated linear combination model. . . . . 42

578 **Fig. 6.** Combined solution (top), the standard deviation of the combined solution (middle) and avail-  
579 able observations for each gauge (bottom). . . . . 43

580 **Fig. 7.** Van de Castelee diagram of BUOY1. The sea level bias model estimated with the COMB  
581 method is displayed in blue, and the one estimated with the DIFF method is displayed in red. . . 44

582 **Fig. 8.** Representation of the bias parameter estimates in the parameter space for both COMB and  
583 DIFF methods. . . . . 45

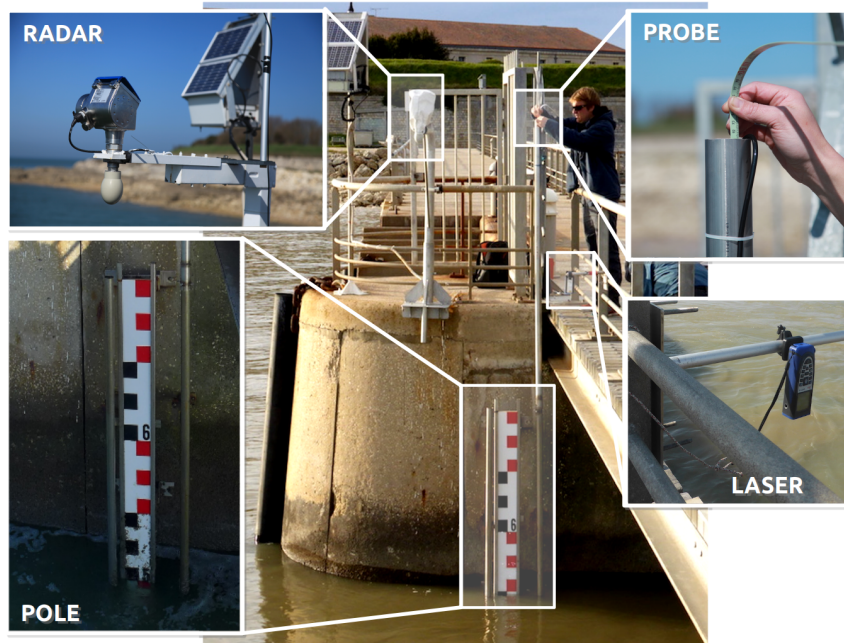


FIG. 1. The four ground-based tide gauges: RADAR, POLE, PROBE and LASER.





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

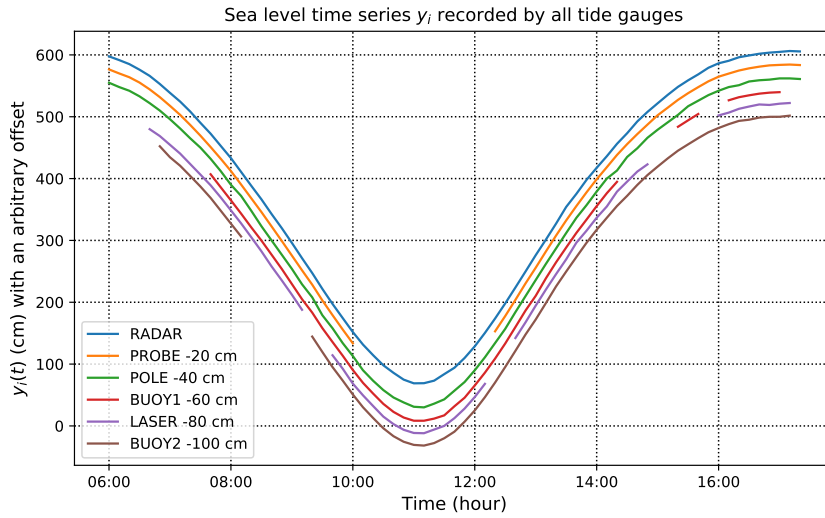
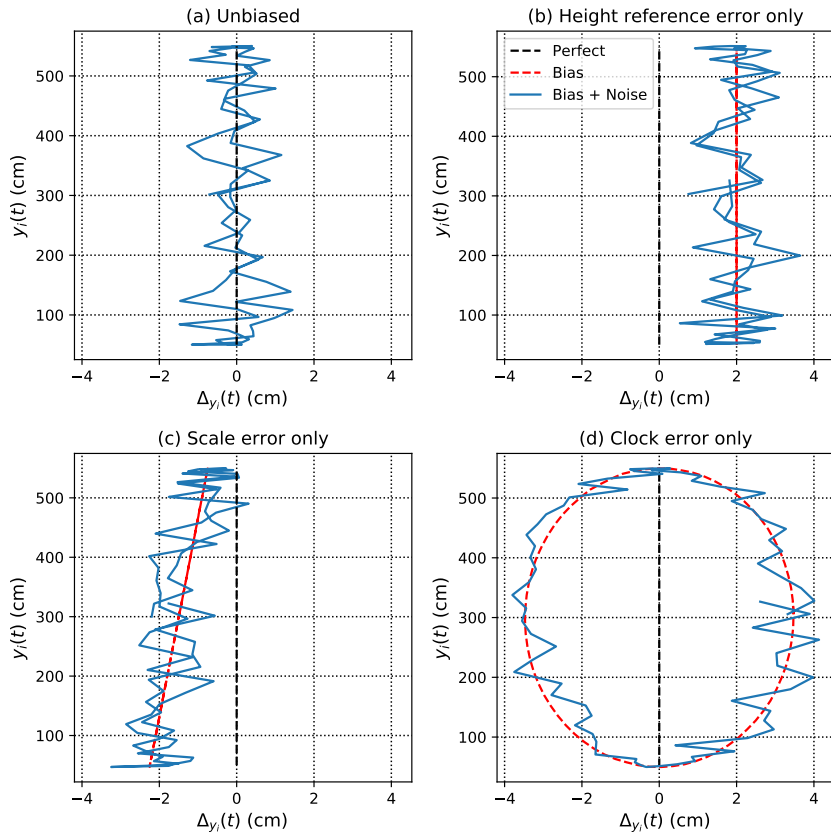


FIG. 3. Sea level time series  $y_i$  recorded by all tide gauges.



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

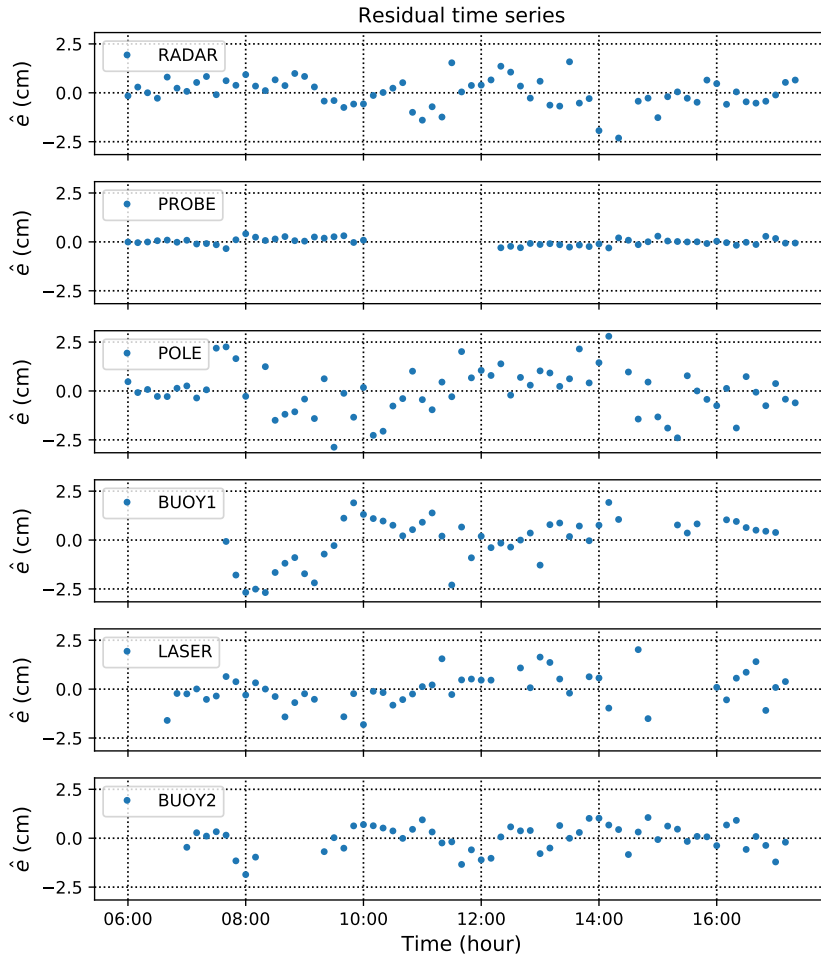
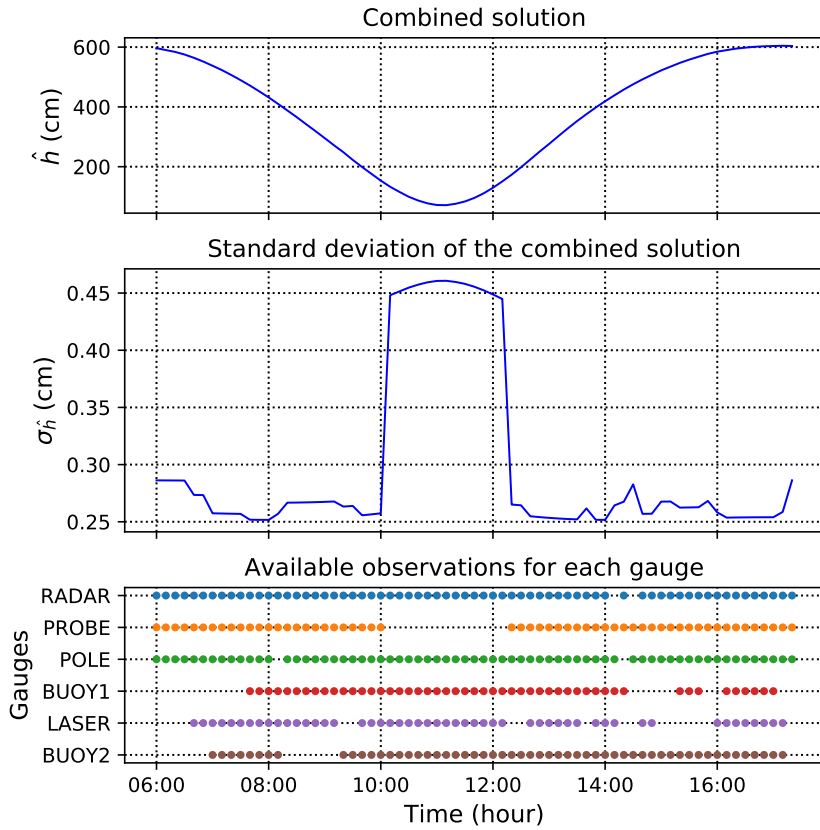
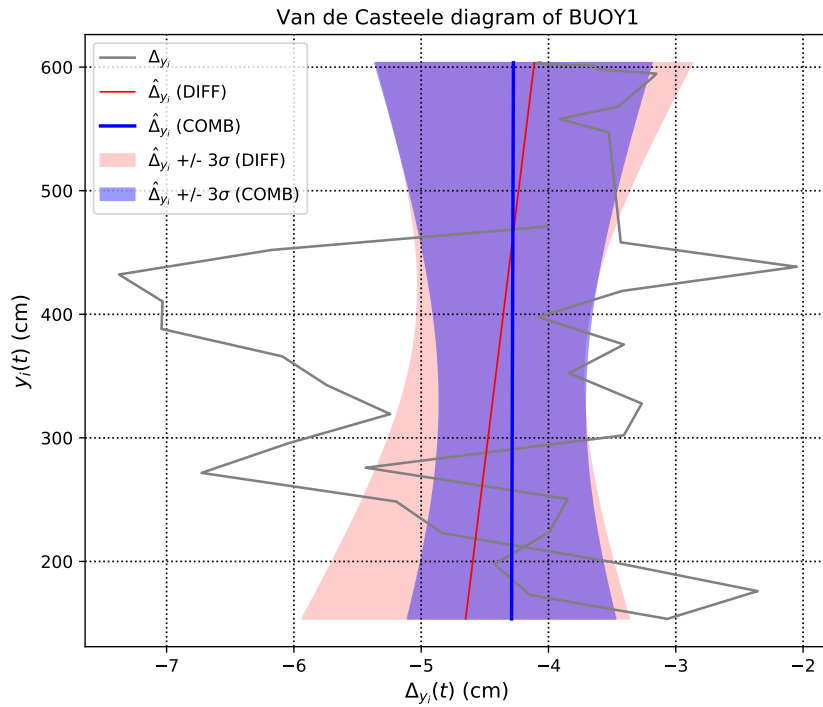


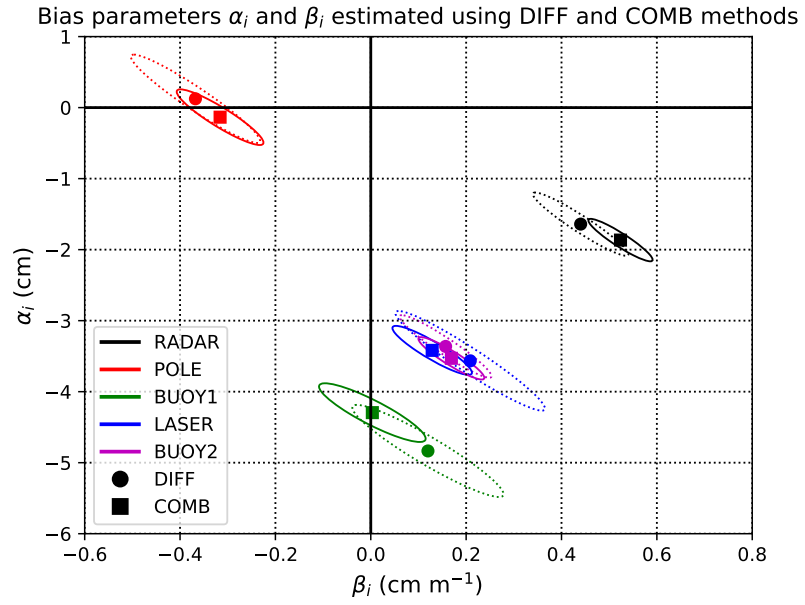
FIG. 5. Residual time series of each tide gauge for the estimated linear combination model.



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



589 FIG. 7. Van de Castele 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.