The Van de Casteele Test Revisited: An Efficient Approach to Tide Gauge Error Characterization

BELEN MARTIN MIGUEZ*
University of la Rochelle–C.L.D.G., La Rochelle, France

LAURENT TESTUT
LEGOS, Toulouse, France

GUY WÖPPELMANN
University of la Rochelle—C.L.D.G., La Rochelle, France

(Manuscript received 6 February 2007, in final form 29 October 2007)

ABSTRACT

The classical question of metrology related to the quality of the tide gauge measurements has become more important this last decade or so as new technologies have emerged and tide gauge networks are modernized. The Global Sea Level Observing System (GLOSS) target of 1-cm accuracy in the individual sea level measurement is motivated by more demanding applications than the traditional hydrographic works and tide predictions, for instance, the monitoring of the long-term trends in sea level or the calibration of satellite radar altimeters. To examine and further assess the performance of modern tide gauge measurements, the Van de Casteele test is revisited. This test is based on a diagram plotting readings taken with a reference probe against the tide gauge readings over at least one tidal cycle. The application of the test to different sets of data at different locations in the world under different environmental conditions shows the test as a simple procedure that immediately gives a qualitative and quantitative illustration of the errors involved in the sea level measurement, capable of sensing the presence of a fault with whatever tide gauge technology is involved. It is recommended that such quality control tests are brought back into fashion and are conducted on a regular basis, in particular following the upgrading of the tide gauge stations.

1. Introduction

In the last two decades considerable progress has been made in the modernization of tide gauge networks. This progress originally arose out of research purposes, in particular storm surges and climatic-related sea level changes. Furthermore, more modern observation technologies have become available: traditional mechanical float devices have progressively been replaced by electronic and digital ones, which are mainly based either on the measurement of the subsurface pressure or on the measurement of the time of flight of a pulse, either acoustic or radar (e.g., see IOC 2002; Wöppelmann and Pirazzoli 2005). In parallel, concern over the performance of the new tide gauge installations has been mounting over recent years. The need to assess the performances of the new technologies is brought about by more demanding applications, like the monitoring of the long-term trends in sea level or the calibration of satellite radar altimeters (Nerem and Mitchum 2001). The Intergovernmental Oceanographic Commission (IOC) manuals (IOC 1985, 1994, 2002) on sea level observation and interpretation provide valuable detailed information on each type of tide gauge, their respective advantages, drawbacks, perfor-

---

* Current affiliation: IOC UNESCO, Paris, France.

Corresponding author address: Guy Wöppelmann, University of la Rochelle–C.L.D.G., Avenue Michel Crépeau, 17042 La Rochelle, France.
E-mail: gwoppelm@univ-lr.fr

DOI: 10.1175/2007JTECHO554.1

mances, and limitations, as well as advice on operational methods and environmental conditions of use. Acoustic and radar tide gauge technologies are, however, relatively recent and more work needs to be done to assess their quality and ensure they meet the Global Sea Level Observing System (GLOSS) target of 1-cm accuracy in the individual sea level height measurement over long periods (IOC 1997, p. 26).

To estimate the accuracy of sea level measurements, laboratory or field experiments can be undertaken in which the tide gauge is compared with an independent higher quality standard or reference gauge. The laboratory tests require, however, that the instrument can easily be removed temporarily from its installation. A rigorous procedure to assure the stability of the datum before and after reinstallation of the gauge is then mandatory. In addition, the on-site operational conditions may affect the electronics of the instrument differently from the more benign laboratory conditions. Therefore, on-site checks are important. They, however, require a standard or reference gauge operating simultaneously with an accuracy of one order of magnitude higher than the required performance of the checked tide gauge itself. Despite the latter being a difficult requirement to meet, the observed differences represent an upper bound on errors in the sea level data, provided the errors at each gauge are uncorrelated with each other and with the true sea level signal. Classical methods applied to analyze the data of such on-site experiments include (e.g., Lentz 1993; Woodworth and Smith 2003; Martin Miguez et al. 2005; Testut et al. 2006)

(i) examination of the time series of the computed differences between the tide gauge measurement and the standard or reference gauge measurement;
(ii) computation of the root-mean-square (rms) of the time series of the differences;
(iii) visualization of one tide gauge's data against the other (scatterplot) and computation of the slope of the linear regression trend between both sea level series. This slope expresses the distinct sensitivities of the gauges to the tidal range;
(iv) inspection of the spectral power of nontidal residuals after tidal variations have been removed by means of the harmonic analysis; and
(v) comparison of the tidal constituents obtained from the harmonic analysis.

The computation of the rms of the differences between the tide gauge and the reference gauge can be used to estimate the upper bound of the noise in the sea level data and their precision (Woodworth and Smith 2003). Evaluating the accuracy, on the contrary, remains a more arduous issue. One should keep in mind here the distinction between precision and accuracy. The precision refers to how closely individual measurements agree with each other, whereas the accuracy refers to how closely a measured value agrees with the correct value. The rms is a synthetic figure insufficient to assess the level of accuracy of the data because it can mask important systematic errors underlying the signal.

To further examine the errors in the tide gauge measurements, in particular the systematic ones, IOC (1985) recommends applying a procedure that was devised by Charles Van de Casteele (1903–77) in the 1960s. Though this procedure is applicable to all sea level measuring techniques, surprisingly its use has been generally restricted to mechanical tide gauges. Our study investigates whether the Van de Casteele test could be usefully applied to modern tide gauge technology. What can be learned from the application of the Van de Casteele test to radar, pressure, or acoustic tide gauges?

To answer these questions several datasets have been considered from different site tests in Spain, the United Kingdom, France, and the far sub-Antarctic island of Kerguelen in the Indian Ocean. Our point when selecting those datasets is double. On the one hand, we have chosen the examples that better exemplify some of the most common faults of the equipment. On the other hand, it is important to highlight that those data have been acquired with very different types of tide gauge, different types of installations, and different environments, thus showing the potential of the Van de Casteele test regardless of the conditions.

2. Data and methods

a. The Van de Casteele test

A detailed description of the Van de Casteele test can be found in Lennon (1968) and in IOC (1985). In short the test involves taking simultaneous sea level heights with both a tide gauge (the gauge to be checked) and a standard (the reference gauge) over a full tidal cycle. The data obtained are then used to construct a simple diagram in which the sea level elevation (y axis) is plotted against the gauge error (x axis). The gauge error (ΔH) is determined as the difference in sea level height measured by the reference tide gauge (H) and the sea level height measured by the tide gauge we are checking (H'). In the case of a perfect gauge the gauge error $\Delta H = H - H' = 0$, and the diagram results in a vertical line centered at zero. In practice the diagram indicates the magnitude of the expected error in the recorded elevations. More importantly, the shape of the diagram is most instructive as it immediately provides a qualitative illustration of the type of error in-

TABLE 1. Main features of the datasets considered in the article.

Example	Site	Type of installation	Technologies compared	Main type of error	Slope	rms (cm)
a	Vilagarcia	open air	Pulse radar/Pulse radar	Instrumental noise	1.0012	0.48
b	Liverpool	open air/subsurface	FMCW radar/Bubbler	Scale	0.9937	1.59
c	Vilagarcia	open air	Pulse radar/Pulse radar	Time shift	0.9999	2.84
d	Vilagarcia	open air	FMCW radar/Pulse radar	Instrumental	1.0018	0.97
e	Brest	stilling well	FMCW radar/Acoustic	Nonlinear	1.0025	0.77
F	Kerguelen	Stainless steel tube	Pulse radar/Pression	Installation	1.0082	0.98

volved. A number of typical Van de Casteele diagrams can be found in IOC (1985, p. 28) with some comments on the probable cause of error in a mechanical tide gauge.

b. The datasets

The Van de Casteele test will be applied to six sets of sea level data obtained in four different test sites. The first experiment was carried out in Liverpool (United Kingdom), and compared the performance of a new Frequency Modulated Continuous Wave (FMCW) radar tide gauge with the performance of the traditional bubbler gauge. The second test station was located in Vilagarcia de Arousa (Spain), where the two pulse radar and one FMCW radar examined in this paper were among up to seven tide gauges measuring together. These two experiments have already been presented in previous works (Woodworth and Smith 2003; Martin Miguez et al. 2005; respectively) and the reader is referred therein for further details. The third experiment was performed in Brest (France), where an acoustic gauge and an FMCW radar tide gauge were installed in the same stilling well. Finally we will present the results obtained in Kerguelen (France, in the austral Indian Ocean), comparing a pressure tide gauge and a FMCW radar gauge, both installed in a stainless tube. Table 1 summarizes the most important features of each of the test sites and the datasets obtained in them.

3. Results: Application to modern gauges

Evidently, the Van de Casteele diagram can reflect the combination of several of the errors listed in Table 1. Nonetheless, for the sake of clarity we have selected the cases where there is one type of predominant error. The diagram will be constructed with time series comprising a few tidal cycles and the mean of the errors will be subtracted so that the eventual offset is eliminated. Finally, the rms and the slope of the linear regression trend between each pair of sea level time series will also be computed.

Figure 1a illustrates a case where no systematic errors can be detected. Both tide gauges used the same type of transducer (a pulse radar of the same manufacturer) and they were placed in the same quay at a distance of approximately 2 m apart inside a sheltered harbor, thus being submitted to the same environmental conditions. Besides, the time series were processed similarly to reduce the differences due to the sampling strategy (see Martin Miguez et al. 2005 for further details). This implies that most sources of discrepancies such as the type of sensor, the location, or the sampling strategy were minimized. Consequently, the dispersion of data in the diagram (the differences between the signals detected by the two tide gauges) is expected to reflect mainly the instrumental noise. Indeed, the plot is near a straight vertical line centered at zero. We can assume that both gauges have a similar contribution to the final random error estimated by the rms, which would yield an individual precision of $0.48/\sqrt{2} = 0.34$ cm.

Figure 1b is a very good example of one of the most common types of systematic errors, namely the scale error. This problem appears when the two instruments are measuring different tidal ranges and this translates into an inclination of the Van de Casteele diagram, whose slope is proportional to the scale error. Suppose that H is the sea level measured by the reference tide gauge and H' is the sea level registered by the tide gauge to be checked. In that case, the Van de Casteele diagram is a plot of H as a function of ΔH; that is to say,

$$H = f(\Delta H) = f(H - H'). \tag{1}$$

In Fig. 1b, the differences between the sea level time series recorded by a bubbler pressure gauge and the sea level recorded by a pulse radar are presented. There is a clear linear slope in the Van de Casteele diagram, so we can express the sea level height H as

$$H = b\Delta H = b(H - H'). \tag{2}$$

The sea level recorded by the bubbler pressure gauge depends on the proper estimation of seawater density. We can express the sea level H and H' in terms of the seawater density and reordering terms we obtain:

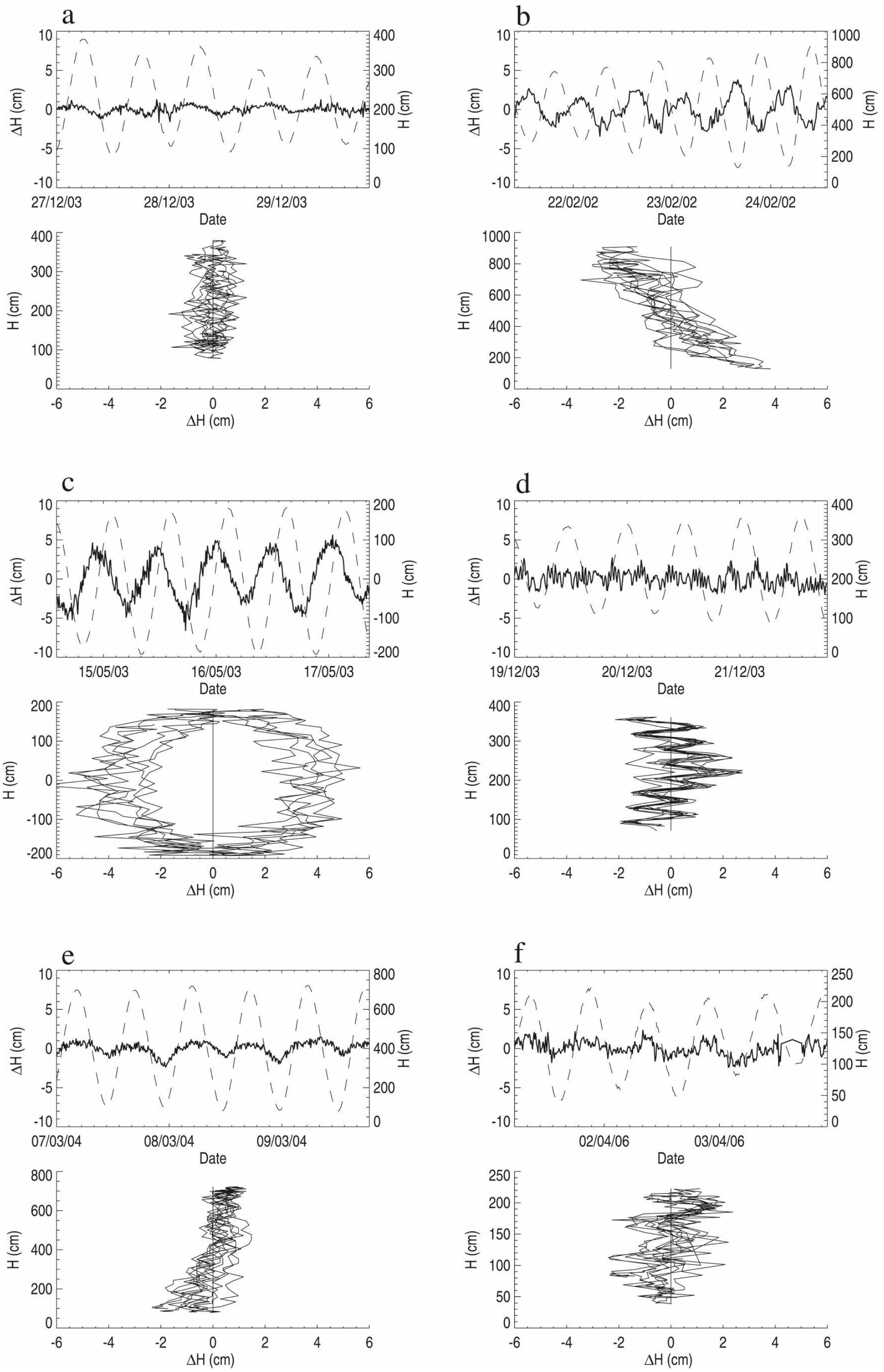

FIG. 1. (a)–(f) Results obtained in each of the examples referred to in Table 1. Upper subplot is time series of the difference between the two instruments under comparison (Δ*H*; left *y* axis) and time series of the sea level height (*H*; right *y* axis). Lower subplot is Van de Casteele diagram representing the sea level (*H*, *y* axis) vs the difference between the records of the two instruments (Δ*H*; *x* axis). The sampling rate of data is 10 min except for example b (15 min).

$$\frac{1}{b} = 1 - \frac{H'}{H} = 1 - \frac{\rho}{\rho'}. \qquad (3)$$

We thus obtain the relationship between the slope b of the linear regression and the scale error. In the Liverpool experiment this scale error was indeed due to an overestimation of the seawater density used for the computation of the sea level (Woodworth and Smith 2003). Our results are consistent with those obtained by those authors.

In Fig. 1c the comparison is made between the same two pulse radars of Fig. 1a. In this case the errors are due to a time shift between the clocks of both instruments. The reference gauge had a GPS-controlled time system; hence, the time shift was due to a lag in the clock of the second one. This lag is reflected in the ellipsoidal shape of the Van de Casteele diagram, and as we shall see, its axis ratio can allow us to approach the estimation of the time shift.

For the sake of simplicity we will assume that the evolution of the sea level H and H' recorded by both tide gauges can be described by means of a single M_2 semidiurnal harmonic constituent characterized by a common amplitude A and a phase that differs between the two instruments by an offset $\Delta\gamma$; that is to say,

$$H = A\sin(\omega t), \qquad (4)$$

$$H' = A\sin(\Delta\gamma + \omega t). \qquad (5)$$

Assume now that the sea level height corresponds to the mean value

$$H = \overline{H} = 0. \qquad (6)$$

In that case,

$$\sin(\omega t) = 0, \qquad (7)$$

$$\Delta H = H - H' = A\sin(\Delta\gamma). \qquad (8)$$

The phase lag (thus, the time shift) between the two instruments is given by

$$\sin(\Delta\gamma) = \frac{\Delta H}{A}. \qquad (9)$$

In the case of Fig. 1c for $H = \overline{H}$, $\Delta H \gg 0.035$ m and $A \gg 2$ m, so $\Delta\gamma = 0.018$ rad. Consequently, for an M_2 wave characterized by a period of 12 h and 25 min, this phase offset corresponds to a time shift of 2 min between the two instruments.

It is interesting to notice that the rms would yield a relatively bad result (2.84 cm) while actually both sensors are providing basically the same information (yet lagged).

The fourth case (Fig. 1d) is an eloquent example of a type of error that is not easily detected in other comparison analyses. This time we are comparing the sea level recorded by two types of radar gauges, an FMCW radar and a pulse radar. As we see in Fig. 1d, the plot shows systematic periodic oscillations that, despite the good rms (<1 cm), clearly suggest a malfunctioning of one of the sensors. The error was found to be due to a problem in the interpolation algorithm used internally in the FMCW sensor. The visualization of the diagram turns out to be most effective to detect a fault in the instrument that would be otherwise ignored if we only took into account the rms. This problem could also be detected after comparing the spectral content of both signals (Martin Miguez et al. 2005). However, this type of spectral analysis cannot be performed on short time series or time series with gaps whereas the Van de Casteele diagram gives an immediate warning even with only one or two tidal cycles of measurements.

In the example of Fig. 2e the dubious performance of one of the gauges is more evident during low tides. The two instruments compared were an FMCW radar tide gauge and an acoustic tide gauge. Former works show the good performance of the radar gauge (Le Roy 2006; Martin Miguez et al. 2007); hence, errors are expected to have their origin in the acoustic gauge. The acoustic sensors need to estimate the speed of sound as a previous step to calculate the distance between the sensor and the water surface. This estimation takes into account the air density, which in turn depends on the air temperature. The existence of temperature gradients between the sensor and the water surface can bring about errors in the estimation of the water level, and these errors increase as the distance between the sea level and the sensors increases (see IOC 1993 for a review). Hence the greatest differences are likely to show up in the diagram at low waters. This turns out to be particularly obvious in the test site of Brest, which is located in a region of great tidal range.

In Fig. 1f we depict a case where the Van de Casteele test serves to make a first evaluation of the performance of new equipment. It concerns the new Kerguelen station, installed in the South Indian Ocean as part of the Indian Ocean Tsunami Warning System (IOC 2005). This station is equipped with a pressure gauge and a waveguided pulse radar measuring the sea level with a 2-min sampling rate in the same stilling well. Time shift errors are not expected because both sensors are synchronized. Despite the low rms, the Van de Casteele diagram immediately alerts us about a problem with one of the sensors, with systematic errors appearing. This problem could be due to the interaction between the radar microwave signal and the stilling well, in other words, a problem related to the inadequacy of the installation. Radar sensors employ microwaves that, unlike acoustic sensors, are only minimally affected by the air conditions. Nevertheless, the signals are relatively sensitive to the interaction with certain materials and the installation of this type of gauges within a tube can bring about problems that are evident in the Van de Casteele diagram.

4. Discussion

The Van de Casteele test is an efficient way of bringing to surface the main tide gauge errors that can be contaminating the sea level records. Its use presents several advantages. To begin with, it is an on-site test, which avoids the problems associated with the reinstallation of the equipment after a laboratory check or the fact that laboratory checks ignore the possible influence of the environment on the internal constituents of the gauge. Second, a few tidal cycles of test (typically five in our examples, but one can be enough) prove sufficient to detect the most relevant problems. Moreover, displaying of the Van de Casteele diagram is extremely easy; no complicated preprocessing of the data is required, just their plotting in the correct way. The shape of the diagram provides a first characterization of the gauge faults in a straightforward way and different types of errors can be clearly identified and added. Its advantage with respect to classical scatterplots clearly appears in the context of large tidal ranges, where the errors are difficult to outline in the scatterplot from the straight line $y = x$.

All these aspects are advantageous in the context of the monitoring and maintenance of the tide gauge networks. Operators need neither long training nor expertise to perform and analyze the results of this type of test, which can be undertaken in a relatively reduced time. In short, the maintenance operations become more affordable.

The Van de Casteele test can also be useful with regard to the upgrading of the tide gauge networks and the checking of new equipment. According to GLOSS (IOC 1994), the new systems must be operated alongside the former ones for a minimum period of 1 yr to provide datum ties and data continuity with the historical time series. It is evident that the longer the period of comparison, the more confidence we can have in the reliability of the new equipment. It is also clear that there are certain aspects of the performance of the tide gauges that require long time series to be detected (the drifts of certain instruments, e.g.). However, there are many cases where long experiments with two instruments measuring simultaneously during a year are just not feasible. Under these circumstances, the Van de Casteele test can be the most cost-effective way of detecting the potential problems of the instruments. What is more, we have shown that despite its simplicity, the test also allows accurate quantification of the scale error and the time shift error. Finally, it is worth pointing out that although we restrict ourselves to tide gauge technologies, the test is likely to be of broader significance, and it may be applicable in other fields and datasets that include a relatively strong higher-frequency signal.

5. Conclusions

Developed under the era of mechanical tide gauges, the Van de Casteele test has been needlessly restricted in the collective mind to this type of tide gauge. The present study outlines the interest in bringing this test back into fashion to examine and further assess the accuracy of modern tide gauge measurements, whether they are acoustic, pressure, or radar. Our findings suggest that even the most recent radar gauges are subject to systematic errors that cannot be neglected depending upon the application. Moreover, the 1-cm accuracy level required by GLOSS for the individual sea level measurement cannot be evaluated on the sole rms estimate of the differences between the tide gauge measurement and the standard or reference gauge. In this context, the Van de Casteele test is a simple but helpful tool that immediately gives a qualitative illustration of the errors involved in the sea level measurement. It therefore provides a mean to investigate the accuracy—or the lack of accuracy—of the instrument by detecting systematic errors, which may then be studied and hopefully corrected. However, whereas this procedure has proven capable of sensing the presence of faults with whichever tide gauge technology is involved, further work should be undertaken to diagnose their cause. This will obviously require a better interpretation of the physics of the measurement techniques, in particular for the radar tide gauges.

Acknowledgments. The authors gratefully acknowledge the contributions of Philip Woodworth (POL, United Kingdom), Begoña Perez (Puertos del Estado, Spain), and Ronan Le Roy (SHOM, France) in kindly providing their tide gauge test sites data of Liverpool, Vilagarcia, and Brest. In addition, the text benefited from the valuable comments of Philip Woodworth. The work was partly funded by GRGS, which provided a postdoctoral contract for B. Martin Miguez. Special thanks are due to Pierre Exertier (GRGS/OCA) and Bernard Simon (GRGS/SHOM). We would also like to thank Frank Healy (University of La Rochelle) for his suggestions to find an appropriate title for the manuscript.

REFERENCES

IOC, 1985: Manual on sea-level measurement and interpretation. Volume I: Basic procedures. Intergovernmental Oceanographic Commission, Manuals & Guides 14, 84 pp.

——, 1994: Manual on sea-level measurement and interpretation. Volume II: Emerging technologies. Intergovernmental Oceanographic Commission, Manuals & Guides 14, 50 pp.

——, 1997: Global Sea Level Observing System (GLOSS)—Implementation plan. Intergovernmental Oceanographic Commission, Technical Series 50, 91 pp & annexes.

——, 2002: Manual on sea-level measurement and interpretation. Volume III: Reappraisals and recommendations as of the year 2000. Intergovernmental Oceanographic Commission, Manuals & Guides 14, 47 pp.

——, cited 2005: The group of experts on the Global Sea Level Observing System (GLOSS). GOOS Rep. 149, IOC/GE-GLOSS-IX/3. [Available online at http://unesdoc.unesco.org/images/0014/001400/140083e.pdf.]

Lennon, G. W., 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. *J. Atmos. Oceanic Technol.,* **10,** 238–245.

Le Roy, R., 2006: Evaluation des télémètres radar pour la mesure des hauteurs d'eau (Assesment of the radar sensors for the measurement of the sea level). EPSHOM, Rapport d'étude 001/06, 49 pp.

Martin Miguez, B., B. Perez Gomez, and E. Alvarez Fanjul, 2005: The ESEAS-RI sea level test station: Reliability and accuracy of different tide gauges. *Int. Hydrogr. Rev.,* **6** (1), 2–11.

——, R. Le Roy, and G. Wöppelmann, 2007: The use of radar tide gauges to measure the sea level along the French coast. *J. Coast. Res.*, doi:10.2112/06-0787.

Nerem, R. S., and G. Mitchum, 2001: Sea level change. *Satellite Altimetry and Earth Sciences: A Handbook of Techniques and Applications,* L.-L. Fu and A. Cazenave, Eds., Academic Press, 329–349.

Spencer, N. E., Ed., 1993: Joint IAPSO-IOC Workshop on sea level measurements and quality control, Paris, 12-13 October 1992. Intergovernmental Oceanographic Commission, Workshop Rep. 81, 176 pp.

Testut, L., G. Wöppelmann, B. Simon, and P. Téchiné, 2006: The sea level at Port-aux-Français, Kerguelen Island, from 1950 to the present. *Ocean Dyn.,* **56** (5–6), 464–472, doi:10.1007/s10236-005-0056-8.

Woodworth, P. L., and D. E. Smith, 2003: A one year comparison of radar and bubbler tide gauges at Liverpool. *Int. Hydrogr. Rev.,* **4** (3), 2–9.

Wöppelmann, G., and P. Pirazzoli, 2005: Tide gauges. *Encyclopedia of Coastal Science,* M. L. Schwartz, Ed., Springer, 984–986.