ADVANCES IN SPANISH PHYSICAL OCEANOGRAPHY
M. Espino, J. Font, J.L. Pelegrí, A. Sánchez-Arcilla (eds.)

SCIENTIA MARINA 76S1
September 2012, 221-228, Barcelona (Spain)
ISSN: 0214-8358
doi: 10.3989/scimar.03618.18A

Performance of modern tide gauges: towards mm-level accuracy

BELÉN MARTÍN MÍGUEZ [1], LAURENT TESTUT [2] and GUY WÖPPELMANN [3,4]

[1] Centro Tecnológico del Mar, Fundación CETMAR, Eduardo Cabello s/n, 36208 Vigo, Spain.
E-mail: bmiguez@gmail.com
[2] LEGOS, 14 av. Edouard Belin, 31400 Toulouse, France.
[3] LIENSs, CNRS/Université de La Rochelle, 2 rue Olympe de Gouges, 17000 La Rochelle, France.
[4] Fédération de Recherche en Environnement et Développement Durable FR 3097 CNRS/SONEL, 2 rue Olympe de Gouges, 17000 La Rochelle, France.

SUMMARY: Considerable efforts are being made worldwide to upgrade tide gauge networks using new technologies. Because of the unique location of the Kerguelen Islands, the measurement of sea level there has received particular attention, with up to four systems equipped with modern sensors functioning simultaneously (two pressure tide gauges, a radar tide gauge, and a GPS-equipped buoy). We analysed and compared the sea level data obtained with these systems from 2003 to 2010, together with a time series of tide pole observations. This is the first time that a multi-comparison study with tide gauges has been undertaken over such a long time span and that the stability of modern radar tide gauges has been examined. The multi-comparison enabled us to evaluate the performance of the tide gauges in several frequency ranges, identify errors and estimate their magnitude. The drift of the pressure sensors (up to 8.0 mm/yr) was found to be one of the most relevant sources of systematic error. Other sources of difference such as clock drift, scale error and different locations of the instruments were also detected. After correcting the time series of sea level for these errors we estimated an upper bound for the radar instrumental error in field condition at ~0.3 cm.

Keywords: radar, pressure gauges, sea level changes, sea level measurement, tide gauges, Kerguelen Islands, GPS buoy.

RESUMEN: FUNCIONAMIENTO DE MAREÓGRAFOS MODERNOS: HACIA LA PRECISIÓN MILIMÉTRICA. – Actualmente se están realizando muchos esfuerzos para renovar las redes maregráficas utilizando nuevas tecnologías. En este contexto, la monitorización del nivel del mar en las islas Kerguelen ha recibido una atención particular debido a su localización única. Hasta cuatro equipos han realizado medidas simultáneamente: dos mareógrafos de presión, un mareógrafo radar y una boya GPS. En este trabajo se analizan y comparan los datos obtenidos con dichos equipos desde 2003 hasta 2010, complementándolos con observaciones realizadas con una escala de marea. Es la primera vez que se plantea una comparación de estas características durante un periodo de tiempo tan largo, y que se aborda el estudio de la estabilidad de los mareógrafos radar a largo plazo. La comparación permitió evaluar el comportamiento de los mareógrafos en distintos rangos de frecuencia, identificar errores y estimar su magnitud. La deriva del sensor de presión apareció como la fuente de error más relevante (hasta 8 mm/año). También se detectaron otras fuentes de diferencias como derivas en el reloj, el error de escala o la diferente localización de los instrumentos. Tras corregir esos errores fue posible estimar un límite superior de ~0.3 cm para el error instrumental del radar en condiciones de campo.

Palabras claves: radar, mareógrafo de presión, cambios de nivel del mar, medida del nivel del mar, mareógrafos, Islas Kerguelen, boya GPS.

INTRODUCTION

The performance of tide gauges has been given considerable attention over the last decades, in particular in the context of the Global Sea Level Observing System (GLOSS) programme (IOC 1997, Merrifield *et al.* 2010). Interest in tide gauge data has increased because of concern about climate change and the rising sea level. Many tide gauges that are perfectly suitable for harbour operations provide data that are not accurate enough for long-term sea level studies. However, systems capable of storing and transmitting large quantities of data now allow tide gauges to survey phenomena in the supra-hourly range, with interesting applications such as the monitoring of storm surges or tsunamis. As a result of the increasing interest in greater-quality, higher-frequency sea level data, tide gauge networks have been upgraded and traditional mechanical float gauges have been progressively replaced by electronic tide gauges equipped with acoustic, pressure and (more recently) radar sensors. During this process, comparison experiments have been undertaken and the results have been published (Woodworth and Smith 2003, Martin *et al.* 2005, Martin Miguez *et al.* 2008b, Blasi 2009). Radar tides gauges in particular stood out as a promising option in terms of accuracy and ease of operation (Martin Miguez *et al.* 2008a). As a result, this technology was chosen for the upgrading of many national networks and in the GLOSS and ODINAfrica programmes (Woodworth *et al.* 2007, 2009, Martin Miguez *et al.* 2008a). This interest is also apparent on the UNESCO/IOC website (www.ioc-sealevelmonitoring.org), on which 25% of the displayed tide gauges use radar technology.

Nevertheless, due to the relatively recent incorporation of these modern sensors in the sea level networks, there is still a lack of information concerning their long-term behaviour and their suitability estimating sea-level trends. Most of the comparison experiments mentioned above used hourly data to determine whether the new tide gauges met the GLOSS 1-cm accuracy requirements (IOC 1997, 2002, 2006), but they did not assess the reliability of the high-frequency data (meaning a sampling interval of a few minutes) or the long-term, inter-annual stability of the sensors.

In this paper we will approach these issues, making use of a rich database comprising simultaneous sea-level data measured with two pressure gauges and one radar tide gauge, a tide pole and a GPS-equipped buoy. The database spans almost 8 years (2003-2010) and allows the performance of the equipment to be assessed at different frequency ranges, from long-term through hourly to high-frequency values. The experiment sheds light on the origin and relative importance of systematic errors and provides us with an estimation of the radar instrumental error in field conditions.

DATA AND METHODS

The sea level systems operating at the Kerguelen Islands (49°21'S, 70°13'E) are part of the ROSAME French Southern Ocean tide gauge network (http://www.legos.obs-mip.fr/en/observations/rosame/), which is a contribution to the GLOSS international network (http://www.gloss-sealevel.org/). The main scientific objectives of the ROSAME network are the study of ocean dynamics in the Kerguelen region and the validation of satellite altimetry missions. The Kerguelen site is currently composed of three tide gauges and a tide pole. A schematic diagram is provided in Figure 1. The first tide gauge (BP1), which is a WLR7 Aanderaa bottom pressure gauge, was installed in 1993 in a stilling well (Testut *et al*. 2006) and measures the bottom pressure, sea water temperature and conductivity. The sea level measurement from the bottom pressure gauges is derived from the difference between the bottom (P_b) and the atmospheric (P_a) pressure following the formula P_b-P_a/ρg, where g is the value of local gravity and the density ρ is computed from in situ ocean temperature and a constant value for salinity.

Though they are particularly suitable for operating in environmentally hostile or remote areas (IOC, 2006), bottom pressure gauges have some disadvantages, including difficulties in controlling their datum and sensor drift (Watts and Kontoyiannis 1990, Woodworth *et al.* 1996, Testut *et al.* 2006). To avoid the latter problem, manufacturers recommend that the sensors be recalibrated every six months. This can prove to be very challenging, particularly in remote places such as Kerguelen, which is generally visited only once a year.

Furthermore, the removal of the bottom pressure sensors for calibration will inevitably imply changes in its position, which are in turn difficult to control because they are underwater. As a consequence, their displacement is generally avoided and in fact the bottom pressure gauges at Kerguelen have not been recalibrated since 2003. Bearing this in mind, in 2003 a tide pole (TP) was installed as a means of having an external stable reference to relate tide gauge measurements and to control their stability and any sensor drift. Thanks to the thorough work of volunteers wintering at the island, tide pole observations have been carried out every month since then. The tide pole is located in a hut approximately 50 m from the BP1 station. The sampling interval is either 2 or 4 min centred on the hour; the volunteers take a total of 25 or 49 observations every 5 s within that interval and calculate an average. They do this every hour over a 12-hour period approximately once a month. The tide pole "monthly" values used in the analysis are the result of averaging those 12-hourly values.

In 2006 the infrastructure for a new tide gauge system was installed and incorporated in the network in the framework of the Indian Ocean Tsunami Warning System. The system is composed of a bottom pressure gauge (BP2, the same WLR7 Aanderaa model as BP1) and a radar gauge (RAD, a Krohne BM100). These two sensors are synchronized and installed in the same stilling well, located approximately 5 m from BP1. All three tide gauges transmit their data in real time via the ARGOS system and these data are available through the ROSAME website.

The whole site is visited on a yearly basis during the logistical rotation of the oceanographic research vessel *Marion Dufresne*. During these visits, infrastructure and equipment are maintained and fixed if necessary. Typical maintenance operations comprise changing the batteries, resetting the tide gauge clock and levelling

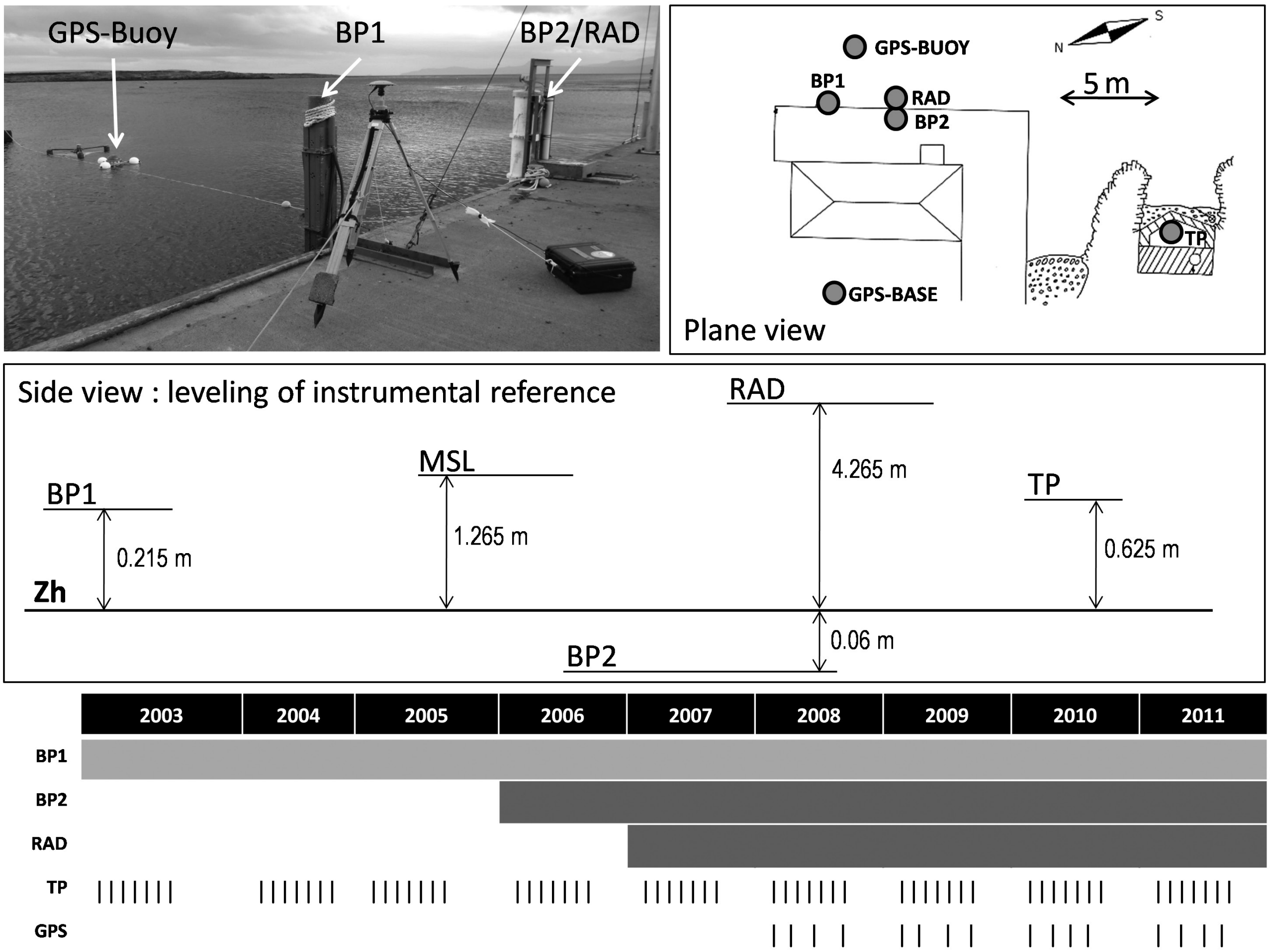

Fig. 1. – The upper right panel shows in plan view the location of the sea-level instruments used in this study (GPS BUOY, RAD, BP1, BP2 and TP) and the GPS permanent station. BP2 and RAD are installed in the same stilling well, approximately 5 m from the historic tide gauge station BP1 (also in a stilling well). The lower part of the figure shows the main levelling information with respect to Zh (local datum). Mean sea level (MSL) is also indicated as a reference.

the RAD sensor. Controlling the datum of BP1 and BP2 is much more difficult because they are underwater.

In addition to the tide gauges, which measure continuously, GPS-equipped buoy campaigns were undertaken in the open sea, 10 m in front of the tide gauge stations (see Fig. 1). The instrument used was a GPS TRIMBLE 5700 with a Zephyr antenna. GPS techniques have been successfully used to perform the in situ calibration of tide gauges in remote islands (Watson *et al.* 2008, Testut *et al.* 2010), their advantage being the absolute reference.

Furthermore, the high sampling rate of tide gauges (1 Hz) makes it possible to study supra-tidal oscillatory phenomena such as seiches and waves. Prior to the deployment, a full on-site calibration of the GPS buoy was performed to determine the distance between the antenna reference point and the sea surface. All the GPS sessions were processed with Total Trimble Control software using the carrier phase in a kinematic mode (on-the-fly mode). Because of the short baseline between the buoy and the base station (50 m) both ionospheric and tropospheric delay are highly correlated.

Table 1. – Sea level measurement at Kerguelen Islands: sea-level systems, periods of operation and sampling strategy.

Sea level system	Start	End	Integration period	Sampling interval
BP1	05/04/1993	Present	2 min or 4 min	1 h
BP2	01/01/2006	Present	40 s	2 min
RAD	01/01/2007	Present	40 s	1 min
TP	21/10/2003	Present	2 min or 4 min	1 h
GPS	16/01/2008	24/08/2010	1 s	1 s

Thus, differential GPS processing allowed accurate epoch-by-epoch (1 Hz) estimates of the sea surface height to be generated from the GPS-equipped buoy measurements.

Table 1 provides more details about the characteristics of the sea level data sets used in the paper. Sampling strategies are not identical for all sea level systems. Depending on the system, a very different number of individual samples are taken and averaged during the integration period, which can range from 1 s (GPS) to 4 min (BP1). This integrated value is then collected with a certain periodicity (sampling interval)

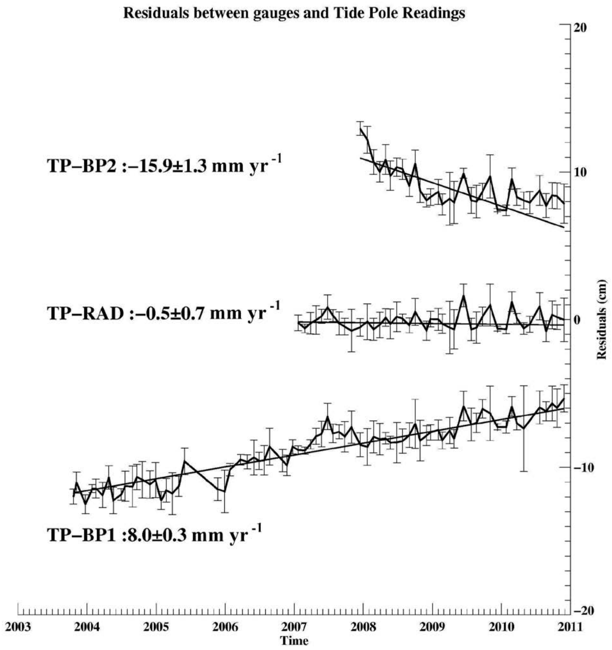

FIG. 2. – Time series of residuals between tide pole (TP) and tide gauge (BP1, BP2 and RAD) observations. Hourly values over a 12-hour period have been averaged to provide the mean values presented in the figure. Records have been offset for presentation purposes. The results of the weighted least-square fit are presented. The error bars correspond to the 1-sigma level of the residuals.

which can also vary. As we shall see in the following sections, time series have been re-processed in order to minimize differences due to those different strategies. Hereafter we will use the word "residuals" to refer to differences between the sea level data measured by the different systems presented in Table 1.

RESULTS

Long-term behaviour

Residuals between the tide pole hourly observations and tide gauge hourly data were calculated and are presented in Figure 2. The values shown in Figure 2 are the mean of the residuals calculated over a 12-hour period each month (i.e. the "monthly" residuals) for the 2003-2010 period of operation. The error bars represent the standard deviation of the residuals, in other words, their dispersion around the so-called monthly mean.

The evolution of the residuals indicates that the position of the instrumental reference of the BP1 and BP2 sensors with respect to the tide pole changed clearly over the 2003-2010 period. BP1 moved upwards with respect to the tide pole, whereas BP2 moved downwards. Linear trends were calculated and are presented in Figure 2 as a reference. We have used the linear approach for the sake of simplicity and comparison, but in fact the long-term behaviour of BP1 and BP2 is not necessarily linear. The position of BP2, in particular, fell sharply in the period 2007-2008, and stabilized after that.

Unlike the bottom pressure sensors, the relative position of the radar sensor with respect to the tide pole did not change clearly (the slope of the linear trend is not significantly different from zero), indicating relative stability.

Another feature worth noting is the similarity of the shape of the three time series of residuals at the intra-annual scale, suggesting that whenever there are high differences these are due to the tide pole: in fact, the accuracy of tide pole measurements can depend greatly on the observer's skills, expertise and the sea conditions.

Hourly data

Getting insight into the stability of the radar sensor and its performance in the hourly frequency range was possible thanks to the availability of another source of data: GPS-equipped buoys. A total of 17 GPS-equipped buoy measurement campaigns were undertaken between January 2008 and December 2010, with durations ranging from 1 to 15 days. BP1, BP2 and RAD measurements were also available during these periods, but no simultaneous tide pole observations were taken. As a result, there are a total of 1000 simultaneous hourly values per sensor, that are suitable to be used for comparison purposes. Hourly values were obtained by reproducing the BP1 sampling strategy (integration period =4 min) and were analysed campaign by campaign.

The GPS-equipped buoy was taken as the reference system. Time series of the residuals between GPS data and the other three tide gauges were computed for each GPS campaign. The mean and the standard deviation for each campaign are presented in Figure 3. As we see, mean values for the comparison between the GPS and the two bottom pressure systems changed over the test period: BP1 relative position moved upwards and BP2 relative position moved downwards. In contrast, RAD remained stable relative to GPS. This behaviour is similar to the one described in the previous section. Note, however, that the GPS campaigns were sporadic and the time span of the observations was shorter (2008-2010, in contrast with 2003-2010 in the previous section).

As mentioned above, the standard deviation indicates the dispersion of the residuals around the mean value. If the systems under comparison are affected by errors and their measurements differ, residuals will be greater and the standard deviation is likely to increase (unless both systems present exactly the same type of errors). As we can see, the magnitude of the bars changed over time and depending on the pair of sensors that we were comparing. The greatest values (and therefore the greatest differences between systems) were found when GPS and BP1 were compared. The differences for GPS-RAD and GPS-BP2 were similar,

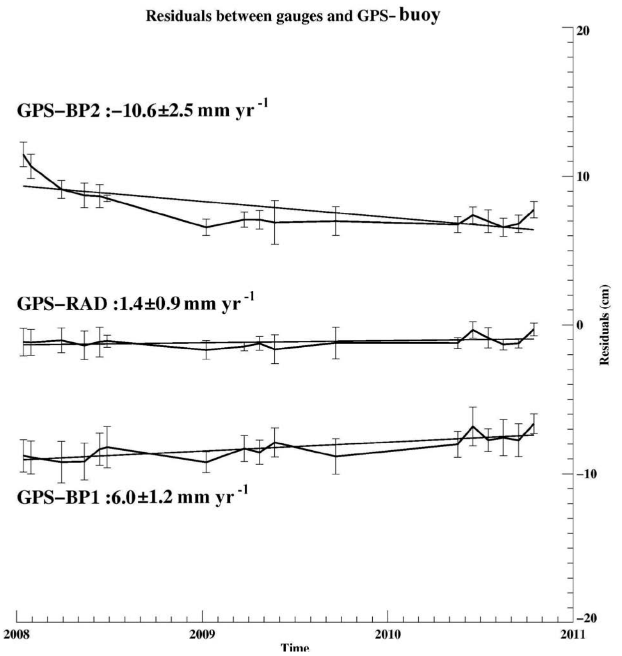

FIG. 3. – Time series of residuals between GPS-equipped buoy (GPS) and tide gauge (BP1, BP2 and RAD) observations. Hourly values over each GPS campaign have been averaged to provide the mean values presented in the figure. Records have been offset for presentation purposes. The results of the weighted least-square fit are presented. The error bars correspond to the 1-sigma level of the residuals.

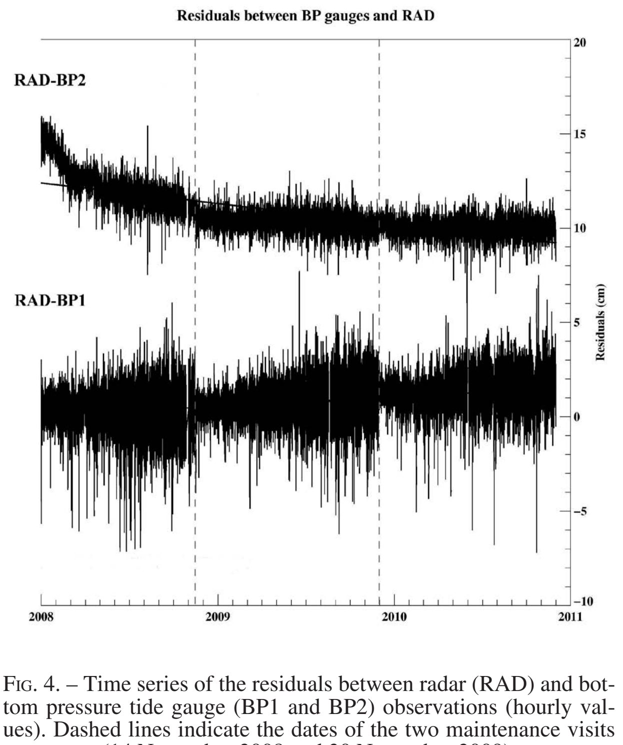

FIG. 4. – Time series of the residuals between radar (RAD) and bottom pressure tide gauge (BP1 and BP2) observations (hourly values). Dashed lines indicate the dates of the two maintenance visits (14 November 2008 and 30 November 2009).

generally <1 cm, lower than the ones found in the previous section when the tide pole was used as a reference system.

We can also compare the hourly time series of the radar and the two bottom pressure records from the time of installation of the radar sensor (Fig. 4). In this case we have a mostly continuous hourly time series of residuals spanning more than two years. We can notice the upwards (BP1) and downwards (BP2) displacement already detected when the tide pole and the GPS were used as reference systems. In addition to this, the plot shows that the residuals were smaller when RAD and BP2 were compared, than when RAD and BP1 were compared.

Furthermore, the RAD-BP2 residuals magnitude is almost homogeneous over the period of comparison, unlike RAD-BP1, in which sharp changes take place following maintenance visits (on 14 November 2008 and 30 November 2009).

High-frequency data

In order to get more information about the possible origin of the residuals, we analysed the data from the second GPS campaign undertaken between 28 January 2008 and 11 February 2008. This is the longest period when GPS, RAD and BP2 sensors were measuring simultaneously, so we can compare not only the hourly values but also their high-frequency time series (integration period =40 s, sampling interval =2 min). Because BP1 has a sampling interval of 1 h, this system will not be included in this part of the study.

We obtained the time series of residuals for each pair of systems (GPS-RAD, GPS-BP2 and RAD-BP2) and we calculated their power spectra (Fig. 5). The largest peak appears around the M_2 tidal frequency (2.2

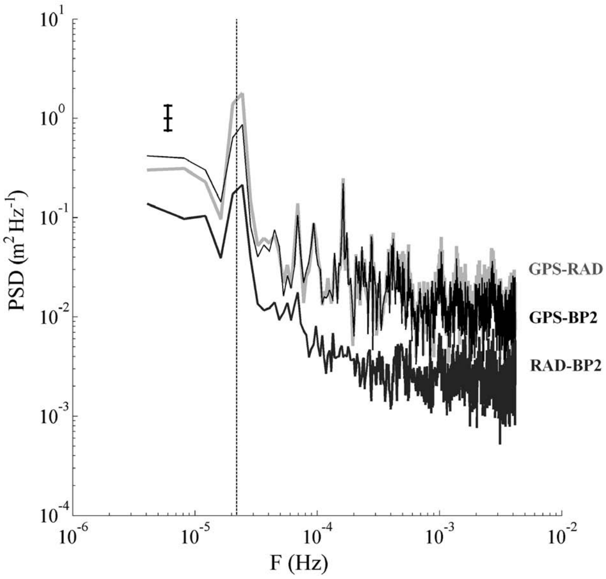

FIG. 5. – Power spectra of the time series of the residuals between GPS-RAD, GPS-BP2 and RAD-BP2 (sampling interval =2 min). The vertical dashed line indicates the M_2 tidal frequency (2.2 10^{-5} Hz). The upper left vertical line indicates the 95% confidence interval for the spectra.

Table 2. – Slope of the linear regression Y=*a*+*b*X between GPS, RAD and BP2

X/Y	GPS	RAD	BP2
GPS	1	1.0032±0.0002	1.0015±0.0002
RAD	0.9963±0.0002	1	0.9982±0.0001
BP2	0.9981±0.0002	1.0017±0.0001	1

10^{-5} Hz) for the three time series. The power spectra of the time series of residuals GPS-RAD and GPS-BP2 are very similar and contain more energy than the time series RAD-BP2. Some much less pronounced peaks can also be distinguished at supratidal frequencies.

One common source of error is related to the tide gauges measuring different tidal ranges (the so-called scale error). In order to investigate this, a linear regression analysis was done between the time series of GPS, RAD and BP2. The results presented in Table 2 show that the GPS sensor was measuring the greatest tidal range, and the RAD sensor was measuring the smallest range. However, the differences were very small. The slope of the linear trend *b* provides an estimation of the scale error ε_s, with $\varepsilon_s=b-1$. If we take the scale error between RAD and GPS (ε_s=3.2‰) and an M_2 amplitude of 53 cm in Kerguelen, this would involve a maximum residual of 0.17 cm.

DISCUSSION

Sources of error

The first focus of the study was to analyse the long-term behaviour of the tide gauges BP1, BP2 and RAD installed at Kerguelen using tide pole observations as an external, independent reference. Results consistently showed that the instrumental reference for the bottom pressure sensors BP1 and BP2 changed its position, moving upwards and downwards respectively. BP1 has not been relocated following maintenance visits since the sensor was last replaced in 2003. It seems unlikely that the sensor has moved physically upwards, so this anomalous change must be due to a drift of the sensor. In contrast, BP2 was located at the bottom of a tube on a soft substrate (probably mud). The substrate underneath might have compacted with time and this explains why the position descended at the beginning and then stabilized. This explanation does not preclude that BP2 may also be affected by a drift or that BP1 could sink following a similar process. In any event, it is obvious that given its magnitude, this change in the position of the sensors whether real or due to a sensor drift, should be taken into account if sea level trends were to be derived from the pressure sensor time series.

As we have seen, tide pole observations may be used to control and, if necessary, correct those drifts. However, the results must be interpreted with caution. The similarity of the shape of the three lines in Figure 2 at the intra-annual scale suggests that whenever there are high differences these are due to the tide pole. This indicates one of the limitations of this method: in addition to the tediousness of the task, tide pole observations can be doubtful when the sea conditions are not calm, limiting the usefulness of the measurements. We attempted to overcome the limitations by using the GPS data obtained in 17 campaigns between January 2008 and 2010 as a reference and studying the time series of the residuals between GPS data, BP1, BP2, and RAD. Despite their sporadic nature and limited duration, the GPS campaigns provided evidence that supported the previous results.

Complementary information was extracted by comparing the BP1, BP2 and RAD hourly values for the whole 2007-2009 period, as shown in Figure 4. The dispersion of the residuals was greater when RAD and BP1 were compared than when RAD and BP2 were compared. This finding could be related to their different location. Since the tide gauges BP2 and RAD were installed inside the same stilling well (Fig. 1), they could be actually measuring different oscillations from BP1 (located in a different stilling well 5 m away) or from the GPS-equipped buoy installed in the open sea. This also explains the lower energy content in the spectra of the time series of RAD-BP2 residuals depicted in Figure 5.

Another interesting feature has to do with the maintenance visits that took place on 14 November 2008 and 30 November 2009. The dispersion of data of the time series of residuals RAD-BP1 was clearly reduced after the visit. During these visits, the BP1 clock was reset so that differences due to a lack of synchronization were minimized just after visits, and progressively increased afterwards. In contrast, RAD and BP2 were always connected to the same clock and the dispersion of the residuals remained similar throughout the comparison period. This shows that even if drifts in the clock are not very important in absolute terms (maximum is 2 min according to Fichen and Tiphaneau [2008]), they can cause residuals of several cm depending on the tidal range. If we do not take the clock effect into account, we can interpret that one of the sensors is not measuring sea level accurately and thus overestimate its instrumental error.

The lack of synchronization between the sensors' clock also explains the peak found at the tidal frequency M_2 in Figure 5. In these cases, the greatest residuals occur systematically between low and high waters, coinciding with the moment of higher tidal velocities. This finding suggests that, though they were connected to the same clock, RAD and BP2 were not perfectly synchronized. In order to estimate the time shift between the two tide gauges, we performed the cross-spectrum analysis of both signals and obtained the phase difference at M_2 frequency (18 s). In addition to this, there were also smaller peaks at supra-hourly frequencies, possibly related to the occurrence of seiches whose signal was not present in the RAD and BP2 records because they were partially hindered by the stilling well.

Estimation of radar instrumental error

The results from the comparisons undertaken in the previous section can now be used to estimate an upper bound for the radar instrumental error in field conditions. We will call ΔH the residuals between the gauge to be examined (radar) and a reference gauge. We can express each residual as a sum of an experimental error ε_{exp} and an instrumental error ε_{inst} and thus write:

$$\Delta H = \varepsilon_{exp} + \varepsilon_{inst}$$

ε_{exp} comprises all possible sources of differences that we identified during the experiment, such as the scale error, the clock shift, the change of position of the sensor reference and the different location of the sensors. ε_{inst} is the term we want to estimate and is related to the instrumental noise, which we assume to be random and not subject to correction. Experimental (ε_{exp}) and instrumental (ε_{inst}) errors are assumed to be independent. Ideally, we would have a perfect reference gauge so that we could confidently assign all the error to the sensor under examination. In our case we used BP2 as a reference, so ε_{inst} will potentially also include BP2 instrumental error.

We calculated the variance of the time series of residuals $\sigma^2_{\Delta H}$ and used it as an indicator of the total error, that is to say, the instrumental error and the errors linked to the experiment. It is worth recalling here that using the variance implicitly assumes that white noise is a good description of the process, which may not always be the case. For instance, the sources of differences contributing to ε_{exp}, previously identified and described, were systematic rather than random. However, the calculation of variance at each step is a useful and simple way of assessing the relative importance of each type of experimental error in the final error budget. We applied this analysis to the period from 28 January 2008 to 11 February 2008, when we had high-frequency data for both the RAD and the BP2 sensors.

Using the results obtained in the previous sections, we progressively corrected the RAD time series for the effect of the scale error, the clock shift, the change of position of the sensor reference, and the different location of the sensors, and recalculated $\sigma^2_{\Delta H}$ at each step to examine how this quantity decreased by taking into account the different sources of experimental error. With this process we aimed to eliminate all errors contributing to ε_{exp} in order to appraise as far as possible the "purely" instrumental-noise-related error ε_{inst} for the radar in field conditions.

First of all, we applied the scale error presented in Table 2 to correct the original RAD time series. The variance of the time series of residuals RAD-BP2 is 0.201 cm^2, while the new time series (after correction of the scale error) amounts to 0.196 cm^2. The effect of this error was thus very small in terms of variance reduction.

However, the second type of error, the one due to the clock drift, appeared to be more relevant. Even though the sensors were set so that their measurements were synchronized, Figure 5 suggested the presence of a slight shift between the time of recording of BP2 and RAD, which we estimated as 18 s through cross-spectral analysis. We corrected this time shift in the RAD time series, obtained a new RAD time series and then recalculated the residuals. The total variance of the residuals RAD-BP2 decreased from 0.196 to 0.170 cm^2.

We also took into account the change of position of BP2 evidenced in the Results section. Since it was a relatively short period (14 days), we considered that the change was linear and performed a linear regression analysis expressing the time series of the RAD-BP2 residuals (corrected for scale error and clock shift) as a function of time. We detrended the time series accordingly, and the new variance further decreased from 0.170 to 0.116 cm^2.

The effect of the different location was negligible in this case. The fact that both RAD and BP2 were measuring exactly in the same stilling well served to cancel out all possible effects related to the place of installation.

Finally, we assumed that the final time series of residuals (i.e. after correction of RAD for all experimental errors) corresponded to the instrumental error, and that this instrumental error came only from the radar. Under all the above assumptions ε_{inst} for the radar system was obtained as the root mean square of the residuals, that is to say 0.341 cm. This value can be regarded as a conservative upper bound, as it might contain other sources of error (for instance, BP2 instrumental error).

CONCLUSIONS

Progress of knowledge and techniques requires a continuous revision of the tools employed to observe the sea level. Regular calibrations and comparisons with external data sets are mandatory exercises to appraise the quality of the data and to ensure their usefulness in present and future applications. This is particularly true in long-term sea level change studies from tide gauge records, which include intertwined climatic contributions and land motion that are challenging to detect because of their millimetre-per-year signature. In this context, the present study contributes to a better understanding of the performance of modern tide gauges and their long-term stability. Based on the comparison with tide pole visual observations, our results showed the superiority of the radar sensor in terms of stability, particularly when compared with bottom pressure sensors, which are prone to drifting. Furthermore, GPS-equipped buoys emerged as a promising technology for measuring the sea level and a valuable in situ metrological tool for the routine task of tide gauge calibration. The study also evidenced the importance of undertaking multi-comparison experiments with several tide gauges measuring simultaneously and of ensuring that the measurements are well synchronized. Only by making full use of all the sea level data acquired with different equipment at several frequency ranges did we manage to identify the sources of error, assess their magnitude, and in a further step obtain an upper bound for the radar instrumental error in field conditions. This error was estimated at ~0.3 cm, clearly below the GLOSS 1-cm requirements. However, the precision and accuracy of observations are never definitively resolved and comparison experiments of this type will continue to be important in the future.

ACKNOWLEDGEMENTS

The research presented in this paper was done in the framework of the French Polar Institute (IPEV) research project 688. It was further supported by the French National Space agency (CNES) in the framework of the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) and the CNES joint Ocean Surface Topography Science Team (OST-ST). The authors wish to express their sincere gratitude to the three anonymous reviewers, whose valuable comments greatly improved the manuscript. They would also like to thank the members of the LEGOS and LIENSs teams, who significantly contributed to the acquisition or processing of data used in this paper, and in particular Philippe Techine, Pascal Tiphaneau, Michel Calzas, Antoine Guillot, Christophe Guillerm and Nelly Rousseau. Useful communication with Francisco Hernandez from VLIZ is also gratefully acknowledged. This paper was mostly written during a fruitful research stay carried out at LIENSs within the framework of the SONEL observing infrastructure of INSU (http://www.sonel.org) and supported by a grant from the Dirección Xeral de Investigación, Desenvolvemento e Innovación, Consellería de Economía e Industria Xunta de Galicia, co-financed by the European Social Fund.

REFERENCES

Blasi C.J. 2009. A new technology for the measurement of the sea level and the sea state. *Environ. Geol.* 57: 331-336.

Fichen L., Tiphaneau P. 2008. Mission Nivmer 09, *NIVMER research cruises reports*. Retrieved January 10, 2011, from http://www.legos.obs-mip.fr/en/observations/rosame/communication/rapports/doc/nivmer09.pdf

IOC. 1997. *Global Sea Level Observing System (GLOSS) - Implementation plan.* Intergovernmental Oceanographic Commission of UNESCO, Technical Series No. 50, Paris.

IOC. 2002. *Manual on sea-level measurement and interpretation. Volume III: Reappraisals and recommendations as of the year 2000*. Intergovernmental Oceanographic Commission of UNESCO, IOC Manuals and Guides No.14, vol. III, Paris.

IOC. 2006. *Manual on Sea-level Measurements and Interpretation, Volume IV: An update to 2006*. Intergovernmental Oceanographic Commission of UNESCO, IOC Manuals and Guides No.14, vol. IV, Paris.

Martín B., Pérez B., Álvarez Fanjul E. 2005. The ESEAS-RI sea level test station: reliability and accuracy of different tide gauges. *Int. Hydrogr. Rev*. 6: 44-53.

Martín Míguez B., Le Roy R., Wöppelmann G. 2008a. The use of radar tide gauges to measure the sea level along the French coast. *J. Coast. Res.* 24: 61-68.

Martin Miguez B., Testut L., Wöppelmann G. 2008b. The Van de Casteele test revisited: an efficient approach to tide gauge error characterization. *J. Atmos. Ocean. Technol*. 25: 1238-1244.

Merrifield M., Aarup T., Allen A., Aman A., Bradshaw E., Caldwell P., Fernandes R.M., Hayashibara H., Hernandez F., Kilonsky B., Martin Miguez B., Mitchum G., Pérez Gómez B., Rickards L., Rosen D., Schöne T., Szabados M., Testut L., Woodworth P., Wöppelmann G., Zavala J. 2010. The Global Sea Level Observing System (GLOSS). In: J. Hall, D.E. Harrison and D. Stammer (eds.), *Proceedings of the "OceanObs'09: Sustained Ocean Observations and Information for Society" Conference (Vol. 2), Venice, Italy, 21-25 September 2009*. ESA Publication WPP-306 (in press).

Testut L., Wöppelmann G., Simon B., Téchiné P. 2006. The sea level at Port-aux-Français, Kerguelen Island, from 1949 to the present. *Ocean Dyn*. 56: 464-472.

Testut L., Miguez B.M., Wöppelmann G., Tiphaneau P., Pouvreau N., Karpytchev M. 2010. Sea level at Saint Paul Island, southern Indian Ocean, from 1874 to the present. *J. Geophys. Res*. 115: C12028.

Watson C., Coleman R., Handsworth R. 2008. Coastal tide gauge calibration: a case study at Macquarie Island using GPS buoy techniques. *J. Coast. Res.* 24: 1071-1079.

Watts D.R., Kontoyiannis H. 1990. Deep-Ocean Bottom Pressure Measurement: Drift Removal and Performance. *J. Atmos. Ocean. Technol.* 7: 296-306.

Woodworth P.L., Vassie J.M., Spencer R., Smith D.E. 1996. Precise datum control for pressure tide gauges. *Mar. Geod.* 19: 1-20.

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

Woodworth P.L., Aman A., Aarup T. 2007. Sea level monitoring in Africa. *Afr. J. Mar. Sci.*, 29: 321-330.

Woodworth P.L., Rickards L.J., Pérez B. 2009. A survey of European sea level infrastructure. *Nat. Hazards Earth Syst. Sci*. 9: 927-934.

Received January 31, 2011. Accepted June 28, 2011.
Published online August 6, 2012.