Advances in Renewable Energies Offshore – Guedes Soares (Ed.)

Wave height error estimation with the triple collocation method for the Canary Islands

R.B. Fonseca, M. Gonçalves & C. Guedes Soares
Centre for Marine Technology and Ocean Engineering (CENTEC), Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

ABSTRACT: The error associated with three datasets of significant wave height is estimated using the triple collocation method. The method requires three independent data sources, so data from a spectral wave model (SWAN), satellite altimeter and wave buoy are used. The three datasets are located in the Canary Islands area, in the North Atlantic Ocean, and include a 10 years period of data. The triple collocation results show differences when using a different referential, and the causes of those differences are discussed in the paper. The lowest error was obtained when using the satellite as referential (0.16 and 0.26 m for the wave buoy and SWAN model, respectively). The error results from the triple collocation method for satellite and SWAN are also compared to the ones obtained using the least square error method with the wave buoy as the true value. The results from the two methods show a difference of 0.08 and 0.16 for the satellite and SWAN, respectively.

1 INTRODUCTION

The renewable energy sector is in constant grow, and study by the industry and scientific community. The study of wind and wave resources is essential for this sector, weather it is for the energy resource or the structural loads applied to the energy convertor. These studies require quality data, and its assessment can be performed with several methods.

Regarding waves, wave height is one of the most important parameters for the design of maritime structures, both coastal and offshore, and for the planning of maritime routes. In this context, significant wave height (SWH) is one of the most used parameters by both the industrial and scientific communities to characterize the wave loads on structures. The SWH can be obtained through several techniques, namely through in situ instrument measurements (wave buoys, ADCP, etc.), remote sensing (high frequency radar, altimeter) and physical and numerical model (SWAN, WW3, WAM, etc).

In situ instrument measurements are generally used as the true value of the parameter. In particular for SWH, the wave buoys are commonly used to validate other type of instruments, as for example the remote sensing technique and the numerical model results. This is usually done using statistical parameters such as bias, root mean square error, correlation coefficient, etc.

Although the buoy error is usually disregarded, it is not zero. On this subject, Liu et al. (2015) compared measurements from wave buoys in laboratory conditions to wave gauges, obtaining correlation coefficients above 95% between them.

The triple collocation method allows for the estimation of the root mean square error of all the datasets involved in its method. Stoffelen (1998) used the triple collocation method to estimate the error of wind datasets from three sources: buoy, satellite and model analysis. Caires & Sterl (2003) also used the triple collocation method to validate the ERA – 40 winds and wave height against buoy and satellite measurements. They found that the high wind speeds are underestimated and that ERA 40 has a larger variance of the error than both the buoy and satellite measurements.

Wang et al. (2014) also used the triple collocation method to estimate the wave height error from buoy, model and satellite. They observed a significant difference between the error results associated with coastal and open shore buoys. In fact, measurements from satellite may have an associated error related to its proximity to the coast, which may lead to misreading by the altimeter.

For this study, three data sources were used: satellite altimetry measurements, wave buoy measurements and wave model results from the numerical wave model SWAN.

SWAN has been widely used by the scientific community for the modeling of wave transformation in coastal areas (Gonçalves et al., 2014; Bento et al., 2014; Rusu & Guedes Soares, 2013; Fonseca et al., 2017). It is a third-generation spectral wave model based on the evolution of the wave

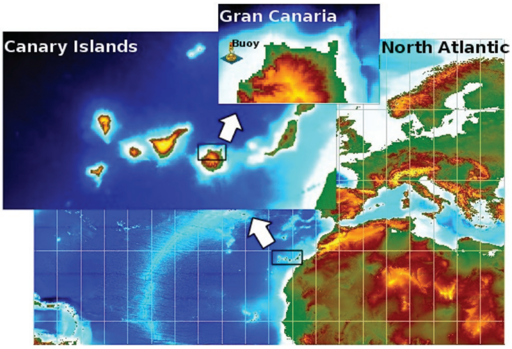

Figure 1. Location of the wave buoy used for the triple collocation method on the Canary Islands.

action density spectrum in time, geographical and spectral domains (Holthuijsen, 2007).

Satellite altimeter and wave buoy are two of several techniques to obtain wave parameter data. Pandian et al (2010) studied the advantages of some of the existent techniques for wave measuring, concluding that wave buoys are simple to install and cost effective, and that satellite have the advantage of a large spatial distribution.

For the application of the triple collocation method, a location near the Gran Canary Island was chosen, where the wave buoy is located (28.20ºN, 15.78ºW) as shown in Figure 1. The time period of the study is between 2000 and 2008.

2 DATASETS

Three sources were chosen for the datasets: wave buoy measurements, satellite altimeter measurements and numerical wave model (SWAN) results.

The satellite altimetry data used in this study is produced and distributed by Aviso web site (http://www.aviso.altimetry.fr/), as part of the Ssalto ground processing segment. For this study, data from two satellite missions was used: the Topex-Poseidon and the GEOSAT Follow-on (GFO) missions. Both missions are equipped with radar altimeters working in the Ku band (13.5 GHz). The data used is the CorSSH Level 2 (L2P) along-track data, with SWH data corrected for instrumental errors and system bias, with a 1 Hz frequency. More information about the data corrections can be found in the CorrSSH Product Handbook on the Aviso web page.

Regarding the SWAN model, the input boundary conditions were obtained from other wave spectral numerical model, the WaveWatch 3 (WW3). For both models, WW3 (Tolman, 1991) and SWAN (Booij et al, 1999) model, the bathymetry is provided by the General Bathymetric Chart of the Oceans, GEBCO database and the wind input fields from ERA Interim database (Dee et al, 2011), produced by the European Centre for Medium-range Weather Forecast (ECMWF), with time steps of 6 h, provided over a grid of 1.5º × 1.5º and interpolated over a grid of 0.5º × 0.5º resolution. WW3 is used to generate waves for the entire North Atlantic basin; its outputs are then used as boundary conditions for the SWAN model.

WW3 uses the JONSWAP spectrum with 24 frequencies, logarithmically spaced, from 0.040 Hz, with increments of 1.12 Hz and 24 directions spaced 15º. Regarding SWAN's parameterization, a spectral grid of 30 frequencies is assumed, logarithmic spaced, between 0.050 Hz e 0.6 Hz, with 0.1 Hz intervals and 36 directions. It is also considered a Janssen parameterization with linear growth. The hindcast system has been validated in (Gonçalves et al. 2014), with measurements from the Gran Canaria buoy, obtained from Puertos del Estado (28.20ºN, 15.78ºW).

3 COLLOCATION METHOD

Before the triple collocation method can be applied, the data from the 3 datasets must be matched in both temporal and spatial domains. The wave buoy has measurements every hour and the model every 3 hours. The satellite data however has measurements with a frequency of 1 Hz, meaning that every time the satellite passes close to the buoy location, for each buoy measurement, there are several satellite SWH values. To deal with this difference, a methodology was applied as in Sepulveda et al. (2015):

- First the satellite measurements within 30 minutes and 0.5 degrees from the time and location of the buoy and model data were grouped.
- The mean and standard deviation are then calculated for this group of observations, and the observations outside the interval of the mean plus or minus the standard deviation are excluded. This way, possible outliers are excluded from the group of observations.
- The remaining observations are then averaged again and the resulting value is used in the triple collocation method.

The wave buoy used is located approximately 8 km off the shore, at 780 meters depth, which is too close to land for the wave conditions to be considered stationary in the 0,5º spatial interval. There was, however, no other buoy available further from shore in the Canary Islands area.

4 THE TRIPLE COLLOCATION METHOD

The triple collocation method, as used by Stoffelen (1998), requires 3 independent datasets and assumes a linear relation between the true value and its estimation.

Considering 3 estimates of one true value X, Y and Z and assuming a linear relation between them and the truth:

$$X = \beta_x T + e_x + \alpha_x \tag{1}$$

$$Y = \beta_y T + e_y + \alpha_y \tag{2}$$

$$Z = \beta_z T + e_z + \alpha_{xz} \tag{3}$$

where β_i represents the calibration constant, α_i represents the bias and e_i corresponds to the root mean square error (rmse) of the estimates. If the system is unbiased, and removing the calibration constants from the equation leads to:

$$X' = \frac{X}{\beta_x} \rightarrow e'_x = \frac{e_x}{\beta_x} \rightarrow X' = T + e'_x \tag{4}$$

with an equal analysis for Y and Z. Considering that the estimates are independent from each other and that, therefore, the errors are uncorrelated, with zero covariance:

$$\langle e_x e_y \rangle = \langle e_x e_z \rangle = \langle e_z e_y \rangle = 0 \tag{5}$$

where the brackets indicate the average, leading to:

$$\langle e'_x \rangle = \langle (X' - Y')(X' - Z') \rangle \tag{6}$$

$$\langle e'_y \rangle = \langle (X' - Y')(Z' - Y') \rangle \tag{7}$$

$$\langle e'_z \rangle = \langle (X' - Z')(Y' - Z') \rangle \tag{8}$$

As in Janssen et al. (2007), the X dataset was chosen as the referential in order to perform the calibration of the datasets. The system is symmetric, so the choice of the reference dataset should not affect the calibration constant values. Using a neutral regression (Marsden, 1998), the calibration constants can be obtained with:

$$\gamma_y = \frac{\langle e_x^2 \rangle}{\langle e_y^2 \rangle} \tag{9}$$

$$B_y = \langle X^2 \rangle - \gamma_y \langle Y^2 \rangle \tag{10}$$

$$A_y = \gamma_y \langle XY \rangle \tag{11}$$

$$C_y = -\langle XY \rangle \tag{12}$$

$$\beta_y = \frac{-B_y + \sqrt{B_y^{\,2} - 4A_y C_y}}{2A_y} \tag{13}$$

By substituting y with z, it is possible to determine the calibration constants for Z.

Regarding the assumption of independence of datasets, satellite altimetry and wave buoys use different techniques for measuring wave data, so it is plausible to assume their independence. The wave models do not use any data assimilation in the process, so it is also reasonable to consider the independence of the wave model dataset.

5 RESULTS

The triple collocation method was applied and the results are now presented. An annual average value of the obtained collocated points for each dataset is shown in Figure 2, corresponding to a total of 82 collocated points. Because the Topex-Poseidon mission data ends in 2005, there is a larger concentration of collocated points in the first part of the time period. The majority of collocated points are located between 1 and 2 SWH meters.

This is also possible to see in Figure 3, where the probability of occurrence of the collocated points

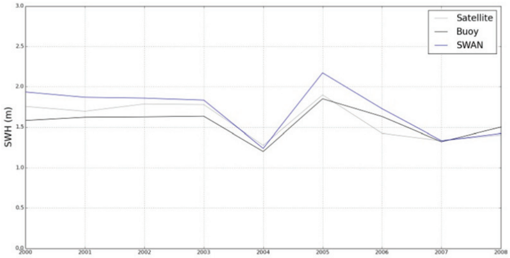

Figure 2. Annual average values of the collocated points for satellite, wave buoy and SWAN model for the time period between 2000 and 2008.

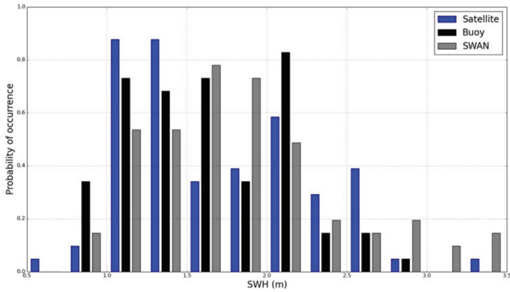

Figure 3. Probability of occurrence for each SWH interval of the collocated points.

better shows the differences between the measurements/results of the three wave data retrieving techniques. For the SWH values between 1.5 and 2 m the difference between the buoy's and the remaining datasets probability of occurrence is around 0.15. A smaller temporal and/or spatial interval for the collocation method could decrease these differences, although further study is required.

In Figure 4 it is possible to see the difference between the wave buoy data and the other two datasets. Compared to the buoy measurements, SWAN seems to generally overestimate the SWH values. The satellite data seems more in agreement with the buoy data, although there is still a larger amount of data that overestimates than that underestimates the SWH data.

Table 1 shows the results of the triple collocation method. Each column shows the error and calibration constant β when using a different referential dataset (satellite, wave buoy or SWAN model). There are some differences between the triple collocation results when considering the different datasets as the referential. These differences can be related to one of the validity of the method's assumptions, as for example the existence of a linear relation between the estimates and the true value. As previously mentioned, the wave buoy is located very close to the shore, in an area with depth variation, resulting in an expected non-stationary of the wave heights inside the spatial interval chosen for the collocation method. Further study is required in order to discover the origin of these differences.

The calibration constants obtained using the triple collocation method with the buoy as referential were then used to plot the linear regression lines of both satellite and model (Figure 5). They are compared with the satellite and model linear regression lines obtained using the buoy as true value in the least square error method. There is a clear difference between the two methods, which increases for high values of SWH. An explanation for this is that the errors might be larger for higher wave height values, leading to a larger discrepancy in the results.

This difference is also very clear in Table 2, where the results for the rmse from the triple collocation method are compared with the ones from the least square error method. For the Satellite case, the difference between the rmse from the two methods is around 0.16 meters.

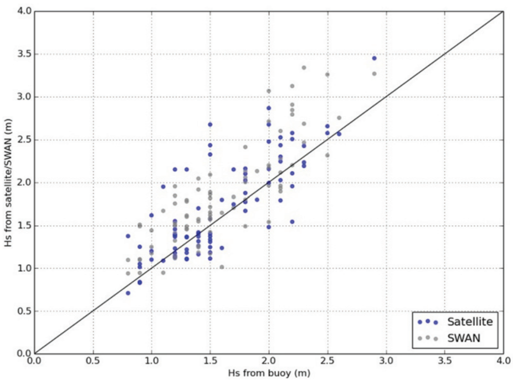

Figure 4. Satellite and SWAN scatter plots against the wave buoy data.

Table 1. Results for the triple collocation method using all 3 datasets as the referential. The columns show the error and calibration constant β estimated values for each dataset with different referentials.

	Buoy		Satellite		SWAN	
Referential	rmse	β	rmse	β	rmse	β
Buoy	–	–	0.293	1.073	0.243	1.147
Satellite	0.159	0.924	–	–	0.264	1.060
SWAN	0.170	0.865	0.338	0.928	–	–

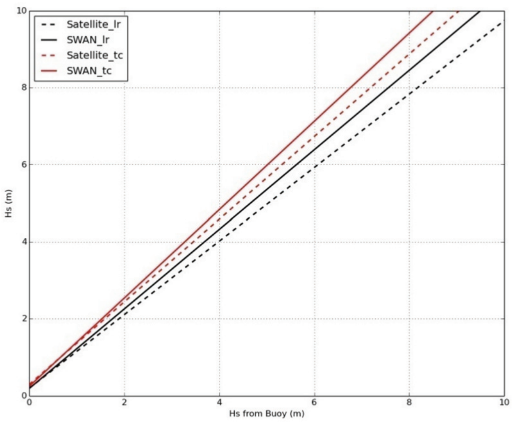

Figure 5. Comparison between the linear regression lines based on the triple collocation method and the least square error method.

Table 2. Results for the rmse using the triple collocation method and the least square error method (lsem), with the buoy as referential.

	Triple collocation	lsem
Satellite	0.293	0.369
SWAN	0.243	0.402

6 CONCLUSIONS

The results from the triple collocation method show a considerable difference to the ones obtained with the traditionally statistical comparison of the results against wave buoy data.

For the method to be applied, some assumptions were initially made. There is a difference between the calibration constants obtained with the different referential datasets that might indicate that some of those assumptions are not valid, as for example the linear relation between the true value and the 3 estimates.

The area considered for the collocation method, the wave buoy location, can be the source of some errors as well, since the closeness to shore may lead to non-stationary wave characteristics. Also, and despite the large time period considered, the number of collocated points obtained is relatively small for this sort of study. If a smaller interval had been chosen, for both spatial and temporal domains, the number of collocated points would be consequentially smaller, but the discrepancies between the datasets may have been smaller. For further works, a larger time period, additional data or different time and spatial interval should be considered.

ACKNOWLEDGEMENTS

This work was conducted within the project ARCWIND—Adaptation and implementation of floating wind energy conversion technology for the Atlantic region (EAPA 344/2016), which is co-financed by the European Regional Development Fund through the Interreg Atlantic Area Programme.

REFERENCES

Bento, A.R., Rusu, E., Martinho, P. & Guedes Soares, C., 2014. Assessment of the changes induced by a wave energy farm in the nearshore wave conditions, Computer & Geosciences, 71, 50–61.

Booij, N., Ris, R.C. & Holthuijsen, L.H., 1999. A third generation wave model for coastal regions. Part 1: Model description and validation. Journal of Geophysical Research, 104, 4, 7649–7666.

Caires, S. & Sterl, A., 2003. Validation of ocean wind and wave data using triple collocation. Journal of Geophysical Research, 108, C3, 3098.

Dee, DP, Uppala, S, Simmons, A, Berrisford, P, Poli, P, Kobayashi, S, Andrae, U, M. Balmaseda, A, Balsamo, G, Bauer, P, Bechtold, P, Beljaars, ACM, van de Berg, L, Bidlot, J-R, Bormann, N, Delsol, C, Dragani, R, Fuentes, M, Geer, A, Haimberger, L, Healy, S, Hersbach, H, Hólm, EV, Isaksen, L, Kållberg, PW, Köhler, M, Matricardi, M, McNally, A, Monge-Sanz, BM, Morcrette, J-J, Peubey, C, De Rosnay, P, Tavolato, C, Thepaut, J-J & Vitart, F., 2011.The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Q.J.R. Meteorol. Soc., 137, 553–597.

Fonseca, R.B., Gonçalves, M. & Guedes Soares, C., 2017. Comparing the performance of spectral wave models for coastal areas, Journal of Coastal Research, 32, 2, 331–346.

Gonçalves, M., Martinho, P. & Guedes Soares, C., 2014. Wave energy conditions in the western French coast, Renewable Energy, 62, 155.163.

Holthuijsen, H., 2007. Waves in oceanic and coastal waters. Cambridge, U.K.: Cambridge University Press, 387p.

Janssen, P.A.E.M., Abdalla, S., Hersbach, H. & Bidlot, J.R., 2007. Error estimation of buoy, satellite and model wave height data. Journal of Atmospheric and Ocean Technology, 24, 1665–1677.

Liu, Q., Lewis, T., Zhang, Y. & Sheng, W., 2015. Performance assessment of wave measurements of wave buoys, International Journal of Marine Energy, 12, 63–76.

Marsden, R.F., 1998. A proposal for a neutral regression. Journal of Atmospheric and Ocean Technology, 16, 876–883.

Pandian, P.K., Emmanuel, O., Ruscoe, J.P., Side, J.C., Harris, R.E., Kerr, S.A. & Bullen, C.R, 2010. An overview of recent technologies on wave and current measurement in coastal and marine applications. Journal of Oceanography and Marine Science, 1, 1, 001–010.

Rusu, L. & Guedes Soares, C., 2013. Evaluation of a high-resolution wave forecasting system for the approaches to ports, Ocean Engineering, 58, 224–28.

Sepulveda, H.H., Queffeulou & P., Ardhuin, F., 2015. Assessments of SARAL/AltiKa wave height measurements relative to buoy, Jason-2 and Cryosat-2 data, Marine Geodesy, 38, 1, 449–465.

Stoffelen, A., 1998. Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, Journal of Geophysical Research, 103, C4, 7755–7766.

Tolman, H.L., 1991. A third generation model for wind waves on slowly varying, unsteady, and in homogeneous depths and currents. Journal of Physical Oceanography, 21, 6, 782–797.

Wang, H., Zhu, J. & Yang, J., 2014. Error analysis on ESA's Envisat ASAR wave mode significant wave height retrievals using triple collocation model, Remote Sensing, 6, 12217–12233.