The Huelva (Spain) tsunami-ready station and its interaction with Storm Emma (March 2018)

J. Lin-Ye

Laboratorio de Ingeniería Marítima, Universitat Politècnica de Catalunya, Barcelona, Spain.

Jordi Girona, 1-3, 08034, Barcelona, Spain.

Puertos del Estado, Madrid, Spain.

(Corresponding author) Av. Partenón, 10, 28042, Madrid, Spain. jln_externo@puertos.es .

ORCID: 0000-0003-4970-1345

B. Pérez-Gómez

Puertos del Estado, Madrid, Spain.

Av. Partenón, 10, 28042, Madrid

E. Álvarez-Fanjul

Puertos del Estado, Madrid, Spain.

Av. Partenón, 10, 28042, Madrid

J. García-Valdecasas

Oritia & Boreas, Madrid, Spain.

Ojos del Salado, 100, 18008, Granada, Spain

The Huelva (Spain) tsunami-ready station and its interaction with Storm Emma (March 2018)

Abstract: The sea level station operating since 1996 at Mazagón (Huelva, Spain) has been progressively upgraded to fit tsunami warning requirements, due to its location in one of the main regions at risk. Its radar water level sensor was complemented in 2017, with the addition of a pressure sensor. The performance of both sea level sensors and their response to sea level oscillations, at different frequencies, is assessed. Particular emphasis is put on the effect of extreme events, such as Storm Emma, when alternative methods to obtain 1-min data are tested, in contrast to the one based on arithmetic means. The overall differences are small, for the whole period of study (centered-root-mean-square-error below 1cm, for 5-min and hourly data; similar tidal parameters and sea level oscillations with periods between 30s and 5 min). However, during Storm Emma, the pressure sensor presents sensibly lower readings than the radar, with the centered-root-mean-square-error rising to 80mm on the 2nd of March of 2018. A new method to compute 1-min data, based on medians, reduced this value to 10 mm for the same day.

Keywords sea level; tide gauge; wave storm; copula

1 Introduction

Sea level changes in a coastal area can have significant social, environmental and economical implications (Pugh (1987), Yin et al. (2010); Curtis et al. (2011); Sánchez-Arcilla et al. (2016)). In particular, a rise in mean sea levels can potentially reduce the area of usable land in low-lying regions, increase the impact of storms and affect the load capacity of a vessel within a harbor. Moreover, sudden water surface elevation changes (as generated by meteotsunamis or infragravity waves) can further affect these vessels. Tsunamis can definitely cause devastating floods and loss of human lives and properties. Therefore, it is a priority in any coastal country to build a robust coastal sea level network that monitors the ocean extreme conditions and long-term sea level changes, as well as to develop models and guidelines that could better protect these regions (Colosi and Munk (2006)). In Spain, such a network exists under the name

REDMAR (Pérez-Gómez et al. (2014)), and has been operated by *Puertos del Estado* (Ports of Spain) since 1992.

Radar water level sensors have been more commonly used in recent times as the primary tide gauges to measure the total sea level (SL), due to their high accuracy and easy maintenance (Woodworth and Smith (2003); Pérez-Gómez et al. (2014); IOC (2016)). The REDMAR network is presently composed of 39 stations located at the main Spanish harbors of the Iberian Peninsula, the Balearic Islands, the Canary Islands and the city of Melilla (north Africa). Its primary instrumentation is the MIROS Range Finder radar water level sensor, a Frequency Modulated - Continuous Wave (FMCW) type of radar (“radar” or “radar sensor”), that transmits a continuous frequency modulated microwave chirp signal and receives the echo from the water surface. The signal propagation delay given by the distance from the antenna to the water surface causes a beat signal in the receiver. An accurate air gap distance is obtained by means of advanced signal processing applied to the beat frequency. The height of the sea water surface (η), in the case of a radar (η_{radar}), is the distance from antenna to the zero of reference (Datum) minus the air gap distance (Y_{radar}). The raw sampling rate of a MIROS radar sensor in the REDMAR network is 0.5s (2-Hz). These raw data, η_{radar} , are stored and processed at a computer at the station to provide wave parameters (every 20-min) and 1-min averaged sea levels (SL_{radar}), that are transmitted in real time to *Puertos del Estado*. The measurement error is 1cm, for individual measurements, and 1mm, for averaged measurements. More details about technical specifications of the MIROS Range Finder (in particular, of the model SM-094) can be found in:

http://www.oceandata.co.kr/groupware/home/product/brochure/miros/sm094_datasheet.

The radar sensor is mounted on top of a structure over the sea water surface (Fig. 1). Some REDMAR stations provide, as well, atmospheric pressure (1-min) and wind data. Although 1-min sampling and latency has allowed integration of all the stations in the national and regional tsunami warning systems, several additional recommended requirements (IOC (2006, 2009, 2016)), especially focused on ensuring data availability and station survival in case of a large tsunami wave, were not yet fulfilled: e.g. installation of a redundant (pressure) sensor and redundant satellite data

transmission. This requirement was particularly recommended for areas close to tsunamigenic sources, such as the Gulf of Cadiz, in Spain.

[Fig. 1 near here]

In 2017, in the framework of the national project SOPRANO (Pérez-Gómez et al. (2019)), a new tsunami-ready pilot station was established in Huelva (Gulf of Cádiz), by upgrading the existing radar based water level radar sensor with a redundant pressure sensor (Fig. 1), a co-located GNSS (Global Navigation Satellite System) and satellite redundant communication link. The pressure sensor installed in Huelva is a RBR Virtuoso Single Channel Recorder (Pressure Sensor) (accuracy: 0.05% full scale, resolution < 0.001% full scale, drift approx 0.1%/year). More detailed specifications can be found in: <https://www.geometius.nl/wp-content/uploads/2015/05/Brochure-Virtuoso.pdf>.

Pressure sensors are installed at a certain depth, well below the lowest low water level. In this case, it is located at a Y_{pres} equal to 2.671m with respect to the Spanish National Mean Sea Level. They record the height of the sea water surface ($\eta_{pressure}$) as the increment of water level due to changes in pressure: the height of sea water surface above the bleed hole (Y_{pres}) plus the distance from the pressure sensor to the zero of reference (h_0 minus $Datum_p$). According to IOC (2006):

$$Y_{pres} = (p_{measured} - p_{atm}) / (\rho_{water} \cdot g) = (p_{measured} - p_{atm}) / \gamma_{water} \quad (1).$$

Here, $p_{measured}$ = measured pressure, p_{atm} = atmospheric pressure, ρ_{water} = sea water density, g = gravitational acceleration and γ_{water} = specific weight of the sea water. The raw sampling rate of the pressure sensor is 1-Hz. The raw data are then averaged every 1-min and transmitted, as well ($SL_{pressure}$).

In order to ensure the chances of survival of, at least, the pressure sensor, in an extreme event, careful attention was paid to its installation (inside the quay) and the move of key electronic equipment to the top of a nearby building. The new station has been in operation since June 2017, and both channels of sea level data (from the radar and pressure sensors) are transmitted every 1-min to the *Instituto Geográfico Nacional*

(National Geographic Institute), that holds since 2013 the National Tsunami Warning Center.

Radar and pressure sensors each have their own weaknesses (IOC (2006)): the former, in most cases, are more prompt to disappear and stop working during extreme sea level events (e.g. a big tsunami) and there is some concern about the influence of waves on the final water levels during a storm; the latter, on the other hand, are affected by changes in density of the water, depth of installation and long-term drifts. Their combined use at the same location has often been recommended and widely implemented (IOC (2009)).

As previously mentioned, the primary objective of the Huelva REDMAR station renovation was to improve its suitability for tsunami warning purposes. However, tsunamis are rare events and coastal sea level stations are used on a daily basis for many practical and scientific applications, including monitoring tides and the mean sea level evolution. The new pressure sensor will be of benefit as well, to these applications, by allowing the implementation of new quality control procedures. For example, a buddy checking or comparison between two sea level channels can be particularly useful to detect datum changes. On the other hand, this redundant sensor will allow for a new intercomparison analysis of historical simultaneous time series, to improve our understanding of the differences between the two types of tide gauges at different frequencies or time scales. Several metrics, such as the moving mean or the centered-root-mean-square-error, are used to compare the several components of the total sea level measured by the two sensors. It is especially relevant to analyze the higher frequency range of the spectra, which is possible, thanks to the raw η of 1-2Hz.

Furthermore, there is a special stress on understanding the different responses from each sensor to external environmental conditions. As the effect of extreme wave conditions on tide gauges is not well known, a possible association between larger deviations between radar and pressure sensors and extreme environmental conditions is sought. The target extreme event is the Storm Emma, which occurred in late February to early March of 2018. Moreover, as the 1-min processed data might be flawed by the asymmetry of the SL , alternatives, such as a low-pass filter and a median, are proposed, in hopes of improving the quality of the measured data.

2 Study area and Storm Emma (March of 2018)

The station involved in this study is located in the Mazagón Sports Harbor, in Huelva (Morales et al. (2004)), southern Spain (see Fig. 2), and has been operating since 1996. The area is bathed by the North Atlantic Ocean and has a mean tidal range of 2m, which makes it meso-tidal (Davies (1964); Pugh (1987)). The mean spring tidal range in the study area is 3.06m. It decreases accordingly when approaching the Strait of Gibraltar, where it is scarcely 1.22m. The energy of the tides gradually dissipates in this direction (Benavente et al. (2015)). The theoretical maximum tidal range in the Gulf of Cádiz during equinoctial spring tides is 3.74m, but wind and atmospheric pressure during storms may add as much as 50cm to the astronomical high tide in the case of severe storms (Del Río et al. (2012)).

[Fig. 2 near here]

Tsunami events have been reported along the Atlantic coast of Europe since prehistoric times, while the Southern coasts of Portugal and Spain (Gulf of Cádiz) have experienced some of the largest tsunamis generated by earthquakes in the area. The most important and well known tsunami is the one following the great earthquake on November 1st, 1755, the same one that destroyed the city of Lisbon (Solares and López-Arroyo (2004); Baptista and Miranda (2009)).

The environmental conditions surrounding the station can be queried at the website of *Puertos del Estado*: www.puertos.es. Based on data provided by the buoy of Cádiz (Fig. 2), between 2017 and 2018, the air temperature fluctuated between 13.52°C and 23.42°C and the water temperature varied between 15.76°C and 23.32°C. According to the data provided by the *Puertos del Estado* wave model (Gómez and Carretero (2005)), at a node shown in Fig. 2, the monthly averaged wind velocities (monthly averaged V_w) ranged between 3.53m/s and 9.16m/s. The main wave directions were west and south-west. The monthly average significant wave height (H_s) was about 0.37m to 1.78m. The monthly averaged peak wave period (T_p) was between 3.86s and 9.45s. The 20-min local H_{m0} measured at the station by the radar sensor ($H_{m0,agit}$) ranged between 0.19m and 0.59m.

The Storm Emma took place by the end of February and the beginning of March of 2018. The March 2nd of 2018 was characterized by a region of deep low pressure located west of the Iberian Peninsula, channeling strong southwesterly winds over the Gulf of Cádiz. It coincided with elevated tides and wave heights (Ferreira et al. (2019)), and it led to severe damages in the Atlantic coast of southwestern Spain (Plomaritis et al. (2019)).

The V_w on March 1st at 13:00 was 15.23m/s in the buoy of Cádiz. It more than doubled the monthly averaged V_w , which was 7.00m/s. It was also significantly larger than the monthly averaged V_w in February and April, which were 5.17m/s and 5.66m/s, respectively. The hourly averaged wind direction at 00:00 was 233° (west-south-west), which was within the range of the most common wind directions in the area. The air temperature in the Gulf of Cádiz was 16.25° C (the monthly average value was 14.5°C), the sea temperature was 15.79°C (the monthly average value was 15.75°C). The air pressure was 991.90mbar (the monthly average value was 1015.00mbar).

On March 1st of 2018, at 13:00, an extreme hourly value of H_s was registered at the buoy. The value was $H_s = 7.22$ m, whereas the monthly average value was 1.85m. It was the highest H_s ever recorded in the buoy, ever since it started recording in 1996 (de Alfonso et al. (Submitted)). Meanwhile, the $H_{m0,agit}$ measured in the station within the Mazagón Harbor, during the same hour, was 0.50m. It was close to the maximum 20-min $H_{m0,agit}$ during 2017-2018, but record values were not reached in the station as it did in the buoy. The hourly averaged wave direction was 232° (south-west). It was clearly a consequence of the approaching Storm Emma. It also coincided with the most frequent wave direction. T_p was 12.11s, which cannot be considered an extreme T_p in 2017-2018, as it is not much larger than the maximum month averaged T_p in this period of time.

3 Methodology

The $SL_{pressure}$ is compared to the SL_{radar} at different time samplings (different frequencies of sea level variations). The latter is considered as the reference for the

comparisons. The values of SL are given with respect to the zero of the REDMAR network (Fig. 1).

The pressure sensor used in this study had a burn-in period that lasted from its installation in June of 2017 to August of the same year. The SL registered in this period of time is of lower quality and is thus discarded. Therefore, the selected period of study will be from September of 2017 to September of 2018, except in one case, which will be described later.

The four different sampling rates are: 1-Hz (raw data), 1-min, 5-min, and hourly data. These are standard sea level products derived from the REDMAR tide gauge network, as part of the sea level quality control and data processing. The η_{radar} is converted from 2-Hz to 1-Hz by taking the secondly averages. In this manner, it becomes comparable to the 1-Hz $\eta_{pressure}$. These sets of data contain different sea level processes, and storms might affect, for example, one sampling rate more than another. As, depending on the applications, different users prefer different sampling rates, it is necessary to assess and compare all of them.

The pressure sensor computes $\eta_{pressure}$ with Eq. 1, by considering the 1-Hz $p_{measured}$ and a constant p_{atm} of 101325Pa at sea level. The 1-min $SL_{pressure}$ are computed by using a $p_{measured}$ equal to the average of the total pressure over 1-min and p_{atm} equal to the measured 1-min atmospheric pressure. Moreover, in both the $\eta_{pressure}$ and the processed $SL_{pressure}$, the pressure sensor considers a constant water density, equal to $1028.1 kg/m^3$. The 5-min data is obtained by averaging 5-min long of 1-min data, whereas the hourly data is obtained from the 5-min data by applying the standard filter for tidal time series that is described in Pugh (1987). It is expected that the data filtering process could reduce the variance of the η .

Several statistics are applied to analyze the data (see Table 1) and the most relevant results will be presented. The moving absolute difference of SL is

$$dSL = SL_{pressure} - SL_{radar},$$

and the relative difference of SL is

$$rdSL = (SL_{pressure} - SL_{radar})/SL_{radar}.$$

Some moving statistics have been applied to study the dSL and $rdSL$: the moving average, the 5th and the 95th percentile. Also, the following moving error metrics are applied to compare the SL of the pair of sensors: bias and centered-root-mean-square-error (CRMSE). The definition of these two metrics can be found in Bennet et al. (2013), where the reference value is set to be the SL_{radar} . The moving bias can, among other applications, help identify a drift between the two time series. The CRMSE is a detrended value, and thus is not contaminated by a possible drift of one sensor with respect to the other. All moving error metrics and statistics for the 1-min, 5-min and hourly data are computed on a moving window of 2days. The length of these moving windows is chosen for presenting both statistical significance of the data and a time series that was not overly smooth. The 1-Hz η are computed on a moving window of 1h. This complies with the same criteria as the other sampling rates, but adapted to the fact that the η are only available for one single day.

[Table 1 near here]

A linear regression is carried out at each moving window, relating the SL from the two sensors. Hence, the following are given and examined: intercept and slope of each linear regression of the $SL_{pressure}$ with respect to the SL_{radar} , the corresponding R-value and the scale error ($eesc$). In practice, these can be considered *moving* intercept, slope, R-value and $eesc$. The expression for $eesc$ is $eesc = (slope - 1) \cdot 100$ (Martín-Míguez et al. (2008)). Stationary error metrics, such as the stationary bias and CRMSE, are also analyzed. The extreme dSL are presented, and a single linear regression is applied to the SL of both sensors, resulting in what is hereafter called the *stationary* intercept, slope, R-value and $eesc$.

The spectra of the SL of different sampling rates are evaluated, while the conformity of $rdSL$ to a Gaussian probability distribution function is examined through a normality test (D'Agostino and Pearson (1973)). A $rdSL$ that followed a Gaussian probability distribution function could be considered to be total white noise. Moreover, the dSL has been related to SL_{radar} , by means of the Van den Castelee plot (Martín-Míguez et al. (2008)). Several consecutive Van den Castelee plots (e.g. Fig. 5) are confectioned for different values of $H_{m0,agit}$: below 0.2m, from 0.2m to 0.4m, and

above 0.4m. Besides these graphs, the CRMSE of the two sensors are computed, for bins of 10mm of SL_{radar} . These values of CRMSE can serve to see how the points are spread across the Van den Castele graph. The larger the CRMSE, the larger is the range of the dSL at a given SL_{radar} . Another type of technique is used to improve the expressivity of the Van den Castele graphs (Fig. 6). They are hexagonal bins, which can serve to cluster points in a scatter plot and show darker colors for increasingly dense regions of points.

Atmospheric and wave parameters from the HARMONIE-AROME atmospheric model, the buoy of Cádiz (Fig. 2), a wave model output in the vicinity of the station and the radar sensor itself (local waves or port agitation) are examined in those periods where $Bias \leq -10mm$ and for periods with a CRMSE over 15mm. These parameters are: the local spectral significant wave height ($H_{m0,agit}$), the local mean absolute wave period ($T_{m02,agit}$); the spectral significant wave height ($H_{m0,buoy}$) and the mean maximum wave period from the buoy of Cádiz ($T_{m02,buoy}$); the spectral significant wave height $H_{m0,mod}$, the mean maximum wave period ($T_{m02,mod}$) and the wave direction (Dir_{mod}) from the *Puertos del Estado* wave model. Please note that the spectral wave models, like the one here, tend to underestimate wave heights. The average value of the bias is given for $Bias \leq -10mm$, whereas the average value of CRMSE is given for CRMSE over 15mm. $Bias \leq -10mm$ and $CRMSE \geq 15mm$ are thresholds that are perceived here as enough to separate extreme and average values of the bias and the CRMSE. The bias of the SL , by comparing the pressure sensor to the radar, is negative, hence the negative sign of the threshold. The resulting periods of time are not homogeneous and are only meant to roughly identify the most problematic periods of time with the largest differences between the sensors. When one such period is identified, the relationship of the atmospheric and wave parameters to the bias and the CRMSE of η are studied.

HARMONIE-AROME (Bengtsson et al. (2017)) uses the non-hydrostatic dynamical core developed within the ALADIN-HIRLAM Numerical Weather Prediction System (Bénard et al. (2010)). It is based on the fully compressible Euler equations. The horizontal, spatial resolution is 2.5km, and the model time step is 75s. Sixty-five levels are used in the vertical, with a model top at approximately 10hPa and

the lowest level at 12m. The wind data is taken at the point 37.1°N, 6.825°W, which is close to the station. Both model and buoy wave data have an hourly sampling rate, while local wave data from the radar sensor are integrated every 20min.

Apart from the basic intercomparison methods described above, two additional tests have been applied. Infrequent ocean parameters, of great return periods, are likely to coincide in time, especially during storms (Lin-Ye et al. (2016)). Their relationship can be non-linear (not possible to be described by a linear regression). It is the case of the bias or the CRMSE of SL and the environmental conditions. Their degrees of dependence can be measured by an indicator called Kendall's τ ((Abdi (2007))). In this study we have computed this indicator for the hourly data of all the storms in 2017-2018. The Kendalls τ ranges between 0 and 1 ($\tau \in [0,1] \in \mathbb{R}$), where 0 corresponds to complete independence and 1 corresponds to complete dependence.

On the other hand, a technique called copula (Sklar (1959)) can provide further details on the dependence between them, by providing the probable combinations of a pair of ocean parameters. The dependence between the bias or the CRMSE of SL and the environmental conditions can be mathematically modelled by asymmetric copulas (Khoudraji (1995)). Again, the hourly data of the whole period of study are used. The statistical parameters for the copula chosen in this case (the independence copula and the Clayton type Archimedean copula) are given, along with the goodness-of-fit parameters of the asymmetric copula. The details of the theory of copulas are given in the Appendix.

The methodology followed for the SL data, except for the evaluation of the drift, was repeated on the storm surge or non-tidal data (*Surge*, sea level minus tide) and the sea level oscillations with periods T_p ranging between 30s and 5min (SL_{LF}), which are both components of the total SL . The SL_{LF} is a type of oscillation that can significantly damage ships stationed within a harbor. It is hardly observable by the naked eye, despite it can induce significant movement of the vessels. Please note that “LF” stands, here, for “low frequency”, in opposition to the frequencies of wind waves, employed in Coastal Engineering. SL_{LF} is a product computed automatically for the REDMAR network, by applying a high-pass finite impulse response (FIR) filter to 1-

min data, with a 15 order Kaiser window function, to remove low frequency signals such as tides and oscillations with periods larger than 3h (Pérez-Gómez et al., 2013). SL_{LF} has a sampling rate of 1-min. The same filter was applied to obtain SL_{LF} from the pressure sensor, from February to September of 2018. The *Surge* is available in the same period as SL , but its sole sampling rate is hourly. Lastly, the tidal constants to be compared are single values obtained for the whole study period.

Alternative methodologies to compute 1-min SL_{radar} , such as a low-pass filter or the computation of medians, is explored here for March 2nd of 2018, during Emma storm. The selected low-pass filter is a combination of the Chebyshev type II digital filter and the two-pass linear digital filter, and it is applied to η_{radar} . 1-min SL_{radar} obtained through these other methodologies have been examined, and compared to the original 1-min SL_{radar} and $SL_{pressure}$. Thereafter, the alternatives are used to compute $SL_{pressure}$ from $\eta_{pressure}$. The results are compared to the original 1-min SL_{radar} and $SL_{pressure}$, as well. The alternative $\eta_{pressure}$ is computed by interpolating the 1-min actual p_{atm} for each second. When computing the moving bias and CRMSE, a time window of 1h is used.

A particularity of the pressure sensors is that, if located too deeply, the water column surrounding it can hamper its accuracy in measuring water levels of higher frequencies. Its measurements could present an attenuation of 10% with respect to the actual value (Cavaleri (1980); Pérez-Gómez et al. (2014)). In the present case, supposing a monochromatic and unidirectional water movement, an irrotational fluid, and that the pressure sensor is only sensitive to the horizontal component of the water motion, the percentage contribution due to the extra pressure of the water can be computed through the expression:

$$\Phi' = \frac{a}{2} \kappa e^{-\kappa z} (2).$$

Here, a is the amplitude of the water level, κ is the wave number and z is the depth. κ can be related to the frequency (f) as:

$$\kappa = (2\pi f)^2 / g \quad (3),$$

where g is the constant of gravity. Eq. (2) can be rewritten in the form:

$$\Phi'' = \Phi' \frac{2z}{a} = \kappa z e^{-\kappa z} \quad (4).$$

Φ forms a curve (Fig. 3). Given an a and a z , it is straightforward to estimate Φ'' and Φ' . The maximum sensibility of a pressure sensor to the dynamic component, for a water level of given a, f and z occurs when Φ' is maximum.

[Fig. 3 near here]

4 Results

4.1 1-min SL

Generally, the radar and the pressure sensors present fairly similar recordings of the SL , although the pressure sensor slightly under-estimates with respect to the radar (Fig. 4). The SL of higher frequencies is filtered in the processing from 1-2Hz to 1-min data, and that explains why the part of the spectrum corresponding to these frequencies has almost no energy at all. The maximum and minimum 1-min dSL are $154.00mm$ and $-182.00mm$, respectively (Table 2).

[Fig. 4 near here]

[Table 2 near here]

The stationary $eesc$ in the whole period of study is -0.34%, whereas the moving $eesc$ is close to zero, ranging from -2.5% to 0.1%. The physical explanation for these negative $eesc$ is that the pressure sensor undermeasures SL , in comparison to the radar.

The van den Castele plot of SL (Fig. 5) shows a dSL that ranges between $-100mm$ and $100mm$, although the greatest density of points is for a dSL from $-10mm$ to $10mm$ (Fig. 6). According to Fig. 5, for a $H_{m0,agit}$ below $0.2m$, the dSL is generally largest for middle values of SL_{radar} . However, as $H_{m0,agit}$ increases, the CRMSE increases as well, especially for high SL_{radar} . Also, for a SL_{radar} greater or equal to $1000mm$, and a $H_{m0,agit}$ below $0.4m$, there is a slight under-estimation of the pressure sensor with respect to the radar. The normality test on the $rdSL$ yields a p-value equal to zero. Hence, $rdSL$ does not follow a Gaussian probability distribution function.

[Fig. 5 near here]

[Fig. 6 near here]

The stationary bias and CRMSE of SL of the whole time period of study are $-4.35mm$ and $11.52mm$, respectively. At a first glance, it seems that the moving bias and CRMSE could be affected by the passage of vessels. This is due to the fact that the two sensors do not measure the extra fluctuations induced by the vessels in the same manner. However, the passage of vessels are not correlated to the bias nor the CRMSE, and thus this effect can be ignored.

Higher moving bias and CRMSE of SL are observed between February 28th and March 12th of 2018 (Fig. 7), during Storm Emma. The moving bias and the CRMSE of SL are particularly large in the March 12th and the 2nd of 2018 respectively, reaching values equal to $-39.0mm$ and $25.0mm$, each. These peaks are within an acceptable order of magnitude, compared to historical data. They coincide with the maximum V_w and $H_{m0,agit}$. The concurrence of the maximum CRMSE during Storm Emma with the peaks in $H_{m0,agit}$ is especially illustrative. Therefore, more conclusions can be drawn if the dependence during storms of these two parameters are mathematically modelled by a copula.

[Fig. 7 near here]

Table 3 shows, for periods of time with $bias \leq -10mm$: the average value of the bias of SL , $H_{m0,agit}$, $T_{m02,agit}$, $H_{m0,buoy}$, $T_{m02,buoy}$, $H_{m0,mod}$, $T_{m02,mod}$ and Dir_{mod} . Table 4 shows, for $CRMSE \geq 15mm$, the average value of CRMSE of SL , as well as the same oceanographic parameters as in table 3, except for the bias.

[Table 3 near here]

[Table 4 near here]

In Tables 3 and 4, bold characters are assigned to a certain period, which concurred with Storm Emma. There was not evident damage to the sensors that led to any additional drift in the relative values (de Alfonso et al. (Submitted)). No incidences were registered by the maintenance of the sensors.

Furthermore, the extreme bias and CRMSE of SL in this case are accompanied by significantly high $H_{m0,agit}$: noteworthy, $H_{m0,agit}$, $H_{m0,buoy}$ and $H_{m0,mod}$ are close to the maximum values of each parameter, respectively, in the period 2017-2018 (see Section 2 and hourly data in Fig. 7b). At the same time, $T_{m02,agit}$, $T_{m02,buoy}$ and $T_{m02,mod}$ do not show extreme values during Storm Emma nor any other periods of relevant values of the bias and the CRMSE of SL . The wave direction during Storm Emma is the southwest, but it is not repeated for the other periods with extreme biases of the SL . It should be noted that the moving bias of SL increased by $-0.71mm$ per month (Fig. 7c). This indicates a certain long-term drift in the measurement of SL between the two sensors.

4.2 Differences and relevant similarities in 5-min and hourly SL

It is intended to tell whether the 5-min and hourly SL change considerably with respect to the 1-min data. The maximum and minimum dSL for the 5-min and hourly SL are more reduced than for the 1-min SL , as expected, due to the smaller variance of SL in these cases. The maximum and minimum differences for the 5-min SL are $120mm$ and $-95mm$, whereas those for the hourly SL are $85mm$ and $-68mm$, respectively (Table 2).

From the Van den Castele graphs, the range of the dSL in 5-min and hourly data are $[-50,50]$ and $[-20,20]$, respectively. The underestimation of the pressure sensor is also present in these sampling rates, and has not been significantly reduced by the further processing of the data.

The normality test on the $rdSL$ is spared for these sampling rates, as it would provide similar results than for 1-min SL . The stationary bias and CRMSE of SL are reduced for the 5-min and the hourly SL . The stationary bias for the 5-min and hourly

SL are $-4.56mm$ and $-4.57mm$, respectively, whereas the stationary CRMSE for the 5-min and hourly SL are $9.51mm$ and $9.15mm$, respectively. The larger bias and CRMSE of SL due to the Storm Emma are persistent in the 5-min and hourly data (Fig. 7).

4.3 Effect of the peak of Storm Emma on the measurement of η by each sensor

The quality of the measurement in the pressure sensor (Fig. 8) is acceptable, since the stationary $eesc$ is -0.34% , whereas the moving $eesc$ ranges from -78.0% to 10.0% . These negative $eesc$ reflect its under-measurement of η . The spectra of the η measured by the two sensors differ for higher frequencies, especially for those above $0.045Hz$.

[Fig. 8 near here]

Most environmental conditions were within the range of the usual values for this season, and could not lead to the higher bias or CRMSE of the η . However, the wave height and direction seem to have some critical effect on the bias and CRMSE of η . The V_w does as well, possibly due to its indirect effect on the wave height. These relationships will be studied, except the role of the wave direction, which would require directional statistics and is left for future work.

4.4 V_w , $H_{m0,agit}$ and their effect on the CRMSE of SL (hourly data)

A strong wind might increase the bias and the CRMSE of SL . Therefore, these two statistical parameters might correlate with two oceanographic parameters related to the V_w : the drag coefficient of the surface of the sea water (Large and Pond (1981)) and the wave set-up. The drag coefficient is computed where the wind data from the HARMONIE-AROMA atmospheric model is available. The wave set-up is computed through the methodologies given in Melet et al. (2018) and in Roelvink (2011), respectively, with data from the buoy of Cádiz.

It is revealed that, from March 1st at 1:00 to March 3rd at 00:00, the CRMSE of SL is strongly correlated with the drag coefficient near the station, where the Pearsons

correlation coefficient is equal to 0.48. The CRMSE of SL is also strongly correlated to the wave set-up computed using the formula in Melet et al. (2018); here, the correlation coefficient is equal to 0.42. As for the wave set-up computed using the formula Roelvink (2011), its correlation to the CRMSE of SL is moderate, being 0.25.

For a drag coefficient larger than 0.0012, which corresponds to a V_w above and equal to 11m/s, and during 2017-2018, the dependence for the drag coefficient vs. the CRMSE of the SL (both as hourly data) is moderate, $\tau = 0.21$. Even small, this dependence is not null, so the two parameters are not totally independent from each other. The copula of the two oceanographic parameters is shown in Fig. 9. This figure displays the parameter of the Clayton copula, as well as the two shape parameters of the Khoudraji asymmetric copula. The log-likelihood is the one for convergence to the asymmetric copula, when the definitive shape parameters are reached. The figure shows that the empirical dependence structure of the CRMSE of the SL and the drag coefficient (Fig. 9a) does fit into an asymmetric copula (Fig. 9b). This denotes a certain dependence structure that is not always present in all relationships between ocean parameters. The resulting copula model (Fig. 9b) shows a higher density below the diagonal of the graph. This translates into, for example, that a drag coefficient corresponding to the 75% strongest winds are related to the 50-75% largest CRMSE in the measurement of SL of the pressure sensor with respect to the radar.

[Fig. 9 near here]

Extreme H_s at the buoy of Cádiz (thus, the $H_{m0,buoy}$) and $H_{m0,agit}$ occur shortly before the extreme values of CRMSE of SL on March 2nd (Fig. 7d). On the other hand, not surprisingly, higher values of CRMSE of SL correspond to those periods with higher values of $H_{m0,agit}$, especially during Storm Emma. No other physical conditions are different during this period, so the effect of local waves is evident. For a $H_{m0,agit}$ larger than 0.08m (90th percentile of $H_{m0,agit}$), during 2017-2018, the Kendall's τ coefficient for $H_{m0,agit}$ vs. the CRMSE of the SL (both as hourly data) is 0.22. This Kendall's τ is equally moderate, in this case, than for the drag coefficient vs. the CRMSE of the SL . Similarly to Fig. 9, the asymmetric copula in Fig. 10 shows the non-linear relationship, the dependence structure between $H_{m0,agit}$ and CRMSE of SL . It confirms previous

experience by displaying that, for example, the 25% largest $H_{m0,agit}$ corresponds to the 0-25% largest CRMSE of the SL .

[Fig. 10 near here]

4.5 Other considerations: Surge (hourly data), SL_{LF} (1min-data) and the tidal constants

Being the *Surge* related to the V_w , its maximum value occurs during the Emma Storm, when the V_w is highest (Fig. 11). Its absolute difference ($dSurge$) is significantly correlated with its value, *Surge* (see Fig. 12), and there is more under-estimation in the pressure sensor with greater values of *Surge*.

[Fig. 11 near here]

[Fig. 12 near here]

The spectra of the SL_{LF} measured in the two sensors match significantly (Fig. 13). Therefore, the pressure sensor can correctly measure the SL where $30s \leq T_p \leq 5min$. The dSL_{LF} (Fig. 14) is much centered in a range between $-5mm$ and $5mm$. Seemingly, the increase of the water column does not increase the error of the measurement of the pressure sensor of these frequencies.

[Fig. 13 near here]

[Fig. 14 near here]

The tidal constants are listed on Table 5. The S2 principal solar semidiurnal, the M2 principal lunar semidiurnal, the K1 lunar diurnal, the S1 solar diurnal, the P1 solar diurnal, the MM lunar monthly and the SA solar annual constituents from the pressure sensor differ in less than $0.1mm$ in all cases except in S1 solar diurnal and M2 principal lunar semidiurnal, where it is about $0.3mm$. In all the cases except in S1, the difference is less than 1% of the amplitude of the tidal constant. Given that the mean residuals of

the tidal constants measured by the radar sensor in 2018 is of -0.04 mm , this difference is acceptably low. The difference in phase is less than 1° , except in the S1 solar diurnal and MM lunar monthly constituents. These are small differences that reflect how the two sensors are essentially measuring the tide in the same manner.

[Table 5 near here]

4.6 Alternative methods to obtain 1-min SL

First, 1-min SL_{radar} , computed by applying the low-pass filter, are compared to the original 1-min SL_{radar} and the original 1-min $SL_{pressure}$. The improvement is shown on Table 6. Both alternatives reduce the bias of the new SL_{radar} with respect to the original SL_{radar} , especially the method of the medians. Furthermore, both alternatives reduce the CRMSE of SL_{radar} with respect to $SL_{pressure}$. Again, the method of the medians reduces this value more than the low-pass filter. Therefore, the method of the medians is the one that might provide the most reliable 1-min SL_{radar} , compared to the original method and the low-pass filter.

[Table 6 near here]

Secondly, 1-min $SL_{pressure}$, computed by applying the low-pass filter, are compared to the original 1-min SL_{radar} and the original 1-min $SL_{pressure}$. The low-pass filter does not reduce the bias of the new $SL_{pressure}$ compared to the original $SL_{pressure}$ or the CRMSE of the new $SL_{pressure}$ and the SL_{radar} . However, the method of the medians does reduce both the bias of the new $SL_{pressure}$ compared to the original $SL_{pressure}$ or the CRMSE of the new $SL_{pressure}$ and the original SL_{radar} . Therefore, the method of the medians is also the best option of the three to produce the 1-min $SL_{pressure}$.

5 Discussion

Some details need further comments, below. For instance, the $rdSL$ does not follow a Gaussian probability distribution function (Table 2). Thus, $rdSL$ is not a white noise and there can be a physical source for it. There is a certain under-estimation from the pressure sensor, which is more prominent for $SL > 1000mm$ (Fig. 5). This might be due to the fact that high tides increase the length of the water column above the pressure sensor, thus hampering its precision. It is also caused by the eventual drift of the pressure sensor with respect to the radar. As pointed out above, there is also a greater $eesc$ in this case, and the reason can be an under-estimation from the pressure sensor for the larger water columns, which related to higher SL . The long-term drift between the two sensors (Fig. 7c) could be explained by the typical loss of references and configuration of the pressure sensor over time (Martín-Míguez et al. (2012)). This can be solved by a more regular maintenance of this type of sensor.

The wave number κ is simplified as $\kappa = (2\pi f)^2 / g$ (Eq. 3). z is equal to Y_{pres} . According to these conditions, the maximum precision and accuracy in the measurement of the dynamic component are for frequencies between $0.010Hz$ and $0.331Hz$ (Fig. 3, Eq. 4). Below $0.010Hz$, the dynamic component is not present, whereas above $0.331Hz$ the pressure sensor does not effectively catch the whole energy of the dynamic component.

It is clear for $\eta_{pressure}$ that water levels of frequencies over $0.331Hz$ should be greatly attenuated (see Fig. 8), and should be the main cause for the under-estimation of the pressure sensor. In practice, the reduction of the energy in the $\eta_{pressure}$ is up to 10% of the radar, as described in Cavaleri (1980), but there is not a clear range of frequencies where the sensitivity of the pressure sensor to the water motion was greater.

The spectra of the 1-min SL_{LF} on both sensors coincide totally (Fig. 13). This is consistent with the nature of this component of the SL . The SL_{LF} is correctly measured by the pressure sensor because it is entirely represented by hydrostatic pressures. That is, the dynamic effect due to the relative motion of the water particles surrounding the pressure sensor is not measured, in the SL_{LF} .

Thanks to Storm Emma, it has been observed that the V_w is strongly correlated to the CRMSE of the SL . That is, high wind velocities induce large CRMSE of SL . This has a series of implications, such as the equally enhanced correlation between the CRMSE of the SL and the drag coefficient of the water surface or the wave set-up. It is possible that the radar registered more accurately the changes in the water surface produced by the wind, but that the pressure sensor could not measure all the extraordinary energy due to this water motion. Similarly, the CRMSE of SL is strongly correlated to $H_{m0,agit}$.

The moving bias is maximum several days after the peak of $H_{m0,agit}$, during Storm Emma (Fig. 7c). There is some indirect effect of waves outside the harbour, $H_{m0,buoy}$, which might affect the instruments because of the diffraction in the harbor. This could also explain why in some cases of extreme $H_{m0,agit}$, but with lower waves offshore, the CRMSE of SL was not as high as during Storm Emma (Fig. 7). The overall relationship between the CRMSE of SL and environmental conditions during storms can thus be better described by Kendall's τ and copulas.

The dependence between V_w and the CRMSE of SL (when V_w above 11m/s) is moderate, but not null. The dependence between $H_{m0,agit}$ (when $H_{m0,agit}$ above 0.08m) vs. the CRMSE of SL is also not null. The asymmetric copulas in Figs. 9 and 10 show that the pressure sensor works better as a support sensor in milder storms. These copulas also bound the degree of deviance of one sensor to another, during storms. Fortunately, supposing that what fails is the pressure sensor, it only does so proportionally to the V_w and $H_{m0,agit}$ (e.g. 50% largest CRMSE of SL for the 50% largest V_w or $H_{m0,agit}$), or it can also perform better than expected (e.g. 25% largest CRMSE of SL for the 50% largest V_w or $H_{m0,agit}$), but it does never perform worse than expected (e.g. 75% largest CRMSE of SL for the 50% largest V_w or $H_{m0,agit}$).

Finally, the alternative method of the medians seems to provide the 1-min SL for each sensor that reduce the respective deviances from the original SL_{radar} and $SL_{pressure}$ (Table 6). This advantage has a logical explanation: when a certain η reaches the pier of the Mazagón Harbor, the relationship between its crests and troughs is asymmetrical (Ursell (1953)). This is especially prominent in stormy events. A method

based on averaging the η in a 1-min time window, which is the case of the original approach, might ignore this phenomenon and force data to present symmetry. However, the alternative method of the medians embraces the asymmetry of the free water surface. Hence, this method can be proposed to substitute the original one in the sensors and provide a more realistic η .

6 Conclusions

There are currently two sensors to measure the SL in the station of Huelva (Spain). A radar has been operating for more than ten years, and a pressure sensor was introduced in 2017 as a redundant sensor, for tsunami warning requirements. These two should provide statistically almost identical measurements of the SL , whereas the effect of extreme wave events, such as the Storm Emma, are studied. It is also intended to improve the methodology that is automatically executed by the sensors to compute 1-min SL . The original approach averaged the η in a 1-min time window.

Error measures, such as the stationary and moving bias and CRMSE of SL , and techniques such as the van den Castele plot, are used, among others, to compare SL , $Surge$ and SL_{LF} from the two sensors. The spectral aspects of each oceanographic parameter are also analyzed. Apart from this, the tidal constants registered by the sensors are contrasted.

The SL has been analyzed in depth in its various sampling rates (raw 1-Hz data, 1-min, 5-min, hourly) and have been related to V_w and $H_{m0,agit}$, among other parameters. The results show that the two sensors provide highly similar measurements of the SL , the $Surge$, the SL_{LF} and the tidal constants, although there is a drift of the SL between the sensors, due to inherent limitations of the pressure sensor. Also, the extreme wind and wave conditions during Storm Emma significantly hampered the functioning of the pressure sensor, which already presents an under-estimation, especially for SL of frequencies above $0.045Hz$. Moreover, the radar is more efficient in registering the water motion produced by the wind.

Alternative methods to obtain the 1-min processed data are: a method based on a low-pass filter and another one based on medians. The method of the medians applied on the *SL* from the radar and the pressure sensor provides the 1-min data that can better reflect the asymmetrical nature of the *SL* measured in the station. Therefore, it is recommended for this methodology to substitute the existing one.

Aknowledgements. This research has been possible with the support of the Spanish national project SOPRANO (Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016 (Convocatoria 2014), Ministerio de Economía y Competitividad).

Declaration of interests. There is no conflict of interests related to this research.

References

Abdi H (2007) The Kendall rank correlation coefficient. Encyclopedia of Measurement and Statistics Sage, Thousand Oaks, CA pp 508–510

Baptista MA, Miranda JM (2009) Revision of the Portuguese catalog of tsunamis. Nat Hazards Earth Syst Sci 9:25–42

Bénard P, Vivoda J, Mašek J, Smolíková P, Yessad K, Smith C, Brožková R, Geleyn JF (2010) Dynamical kernel of the Aladin–NH spectral limited-area model: Revised formulation and sensitivity experiments. Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography 136(646):155–169

Benavente J, Gracia FJ, Del Río L, Anfuso G, Rodríguez-Ramírez A (2015) Morphodynamic characterization of the Spanish beaches of the Gulf of Cadiz. Boletín Geológico y Minero 126(2–3):409–426

Bengtsson L, Andrae U, Aspelien T, Batrak Y, Calvo J, de Rooy W, Gleeson E, Hansen-Sass B, Homleid M, Hortal M, et al. (2017) The HARMONIE–AROME model configuration in the ALADIN–HIRLAM NWP system. Monthly Weather Review 145(5):1919–1935

Bennet N, Croke BF, Guariso G, Guillaume JH, Hamilton SH; Jakeman A, Marsili-Libelli S, Newham LT, Norton J, Perrin C, Pierce S, Robson B, Seppelt R, Voinov A, Fath B, Andreassian V (2013) Characterising performance of environmental models. *Environmental Modelling and Software* **40**: 1-20.

Cavaleri L (1980) Wave measurement using pressure transducer. *Oceanologica Acta* 3(3):339–346

Colosi JA, Munk W (2006) Tales of the venerable Honolulu tide gauge. *Journal of Physical Oceanography* 36(6):967–996

Curtis S, Long P, Arrigo J (2011) Climate, weather, and tourism: Issues and opportunities. *Bulletin of the American Meteorological Society* 92(3):361–363

D’Agostino R, Pearson ES (1973) Test for departure from normality. *Biometrika* 60(3):613–622

Davies JL (1964) A morphogenic approach to world shorelines. *Zeitschrift fur Geomorphologie* 8:127–142

de Alfonso M, García-Valdecasas JM, Aznar R, Pérez-Gómez B, Rodríguez P, de los Santos FJ, Álvarez-Fanjul E (Submitted) Record wave storm in the gulf of Cádiz over the past 20 years and its impact on harbours. Unknown

de Amo E, Díaz-Carrillo M, Fernández-Sánchez J (2011) Measure-preserving functions and the independence copula. *Mediterranean journal of mathematics* 8(3):431–450

Del Río L, Plomaritis TA, Benavente J, Valladares M, Ribera P (2012) Establishing storm thresholds for the Spanish Gulf of Cádiz coast. *Geomorphology* 143–144:13 – 23, DOI <https://doi.org/10.1016/j.geomorph.2011.04.048>

Ferreira Ó, Plomaritis TA, Costas S (2019) Effectiveness assessment of risk reduction measures at coastal areas using a decision support system: Findings from Emma storm. *Science of the Total Environment* 657:124–135

- Gómez M, Carretero JC (2005) Wave forecasting at the spanish coasts. *Journal of Atmospheric and Ocean Science* 10(4):389–405
- Hofert M, Vriens F (2013) Sibuya copulas. *Journal of Multivariate Analysis* 114:318 – 337, DOI <https://doi.org/10.1016/j.jmva.2012.08.007>
- Hürlimann W (2004) Fitting bivariate cumulative returns with copulas. *Computational statistics & data analysis* 45(2):355–372
- IOC (2006) Manual on sea level measurement and interpretation, volume IV: An update to 2006. Tech. rep., Intergovernmental Oceanographic Commission, Paris, UNESCO
- IOC (2009) Tsunami early warning and mitigation system in the North Eastern Atlantic, the Mediterranean and connected seas, NEAMTWS, Implementation Plan. Tech. rep., Intergovernmental Oceanographic Commission, Paris: UNESCO
- IOC (2016) Manual on sea-level measurements and interpretation, volume V: Radar gauges. Tech. rep., Intergovernmental Oceanographic Commission, Paris, UNESCO
- Khoudraji A (1995) Contribution à l'étude des copulas et à la modélisation des valeurs extrêmes bivariées. PhD thesis, Université Laval, Québec, Canada
- Lagarias JC, Reeds JA, Wright MH, Wright PE (1998) Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM Journal on optimization* 9(1):112–147
- Large WG, Pond S (1981) Open ocean momentum flux measurements in moderate to strong winds. *Journal of physical oceanography* 11(3):324–336
- Lin-Ye J, García-León M, Gràcia V, Sánchez-Arcilla A (2016) A multivariate statistical model of extreme events: An application to the Catalan coast. *Coastal Engineering* 117:138 – 156

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

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

Melet A, Meyssignac B, Almar R, Le Cozannet G (2018) Under-estimated wave contribution to coastal sea-level rise. *Nature Climate Change* 8(3):234

Morales JA, Borrego J, Ballesta M (2004) Influence of harbour constructions on morphosedimentary changes in the Tinto-Odiel estuary mouth (South-West Spain). *Environmental Geology* 46(2):151–164

Pérez-Gómez B, Álvarez Fanjul E, Pérez S, de Alfonso M, Vela J (2013) Use of tide gauge data in operational oceanography and sea level hazard warning systems, *Journal of Operational Oceanography*, 6:2, 1-18, DOI: 10.1080/1755876X.2013.11020147

Pérez-Gómez B, Payo A, López D, Woodworth PL, Álvarez-Fanjul E (2014) Overlapping sea level time series measured using different technologies: an example from the REDMAR Spanish network. *Natural Hazards and Earth System Science* 14(3):589–610

Pérez-Gómez B, Pérez-González I, García-Sotillo M, Álvarez-Fanjul E (2019) Retos de los sistemas de observación y predicción en los riesgos asociados al nivel del mar. *Ribagua*

Plomaritis TA, Ferreira Ó, Costas S (2019) Validation of a Bayesian based early warning system for coastal hazards: The Emma storm impact at Faro Beach (South Portugal). *Coastal Sediments*

Pugh DT (1987) *Tides, surges and mean sea-level: a handbook for engineers and scientists* Roelvink D (2011) *A guide to modeling coastal morphology*, vol 12. world scientific

Sánchez-Arcilla A, García-León M, Gracia V, Devoy R, Stanica A, Gault J (2016) Managing coastal environments under climate change: Pathways to adaptation. *Science of the total environment* 572:1336–1352

Sklar A (1959) Fonctions de répartition à n dimension et leurs marges. Université Paris 8

Solares JM, López-Arroyo A (2004) The great historical 1755 earthquake. Effects and damage in Spain. *Journal of Seismology* 8:275–294, DOI 10.1023/B:JOSE.0000021365.94606.03

Ursell F (1953) Short surface waves due to an oscillating immersed body. *Proc R Soc Lond A* 220(1140):90–103

Vanem E (2016) Joint statistical models for significant wave height and wave period in a changing climate. *Marine Structures* 49:180–205

Woodworth PL, Smith DE (2003) A one-year comparison of radar and bubbler tide gauges at Liverpool. *International Hydrographic Review* 4(3):42–49

Yin J, Griffies SM, Stouffer RJ (2010) Spatial variability of sea level rise in twenty-first century projections. *Journal of Climate* 23(17):4585–4607

Zhang Y, Kim CW, Beer M, Dai H, Guedes Soares C (2018) Modeling multivariate ocean data using asymmetric copulas. *Coastal Engineering* 135:91–111

A Theoretical background of an applied statistical technique: an asymmetric copula

A copula (Sklar (1959)), C , for any bivariate distribution function H with margins

$F_i(x_i) = u_i, x \in \mathbb{R}, j \in \{1,2\}$ can be defined such that:

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)) = C(u_1, u_2), \quad x \in \mathbb{R}.$$

A type of standard copula is the independence copula (de Amo et al. (2011)). It has the form:

$$C(u_1, u_2) = u_1 u_2.$$

Another standard type is the Clayton type Archimedean copula (Hürlimann (2004)). It has the form:

$$C(u_1, u_2) = \max\left([u_1^{-\theta} + u_2^{-\theta} - 1]^{-1/\theta}, 0\right), \quad \theta \in [-1, \infty) \setminus \{0\},$$

where θ is a parameter that defines the shape of the Clayton type Archimedean copula.

A method to obtain not-so-standard copulas is the asymmetric copula (Zhang et al. (2018); Hofert and Vrin (2013)). It is called in this manner because, in the presence of oceanographic parameters which present asymmetric relationships between them, this type of copula can mix different types of standard copulas in order to mathematically model the dependence between the oceanographic parameters (Vanem (2016)). It can allow for the existence of asymmetric relationships between the u_i by combining different types of standard copulas. The Khoudraji device (Khoudraji (1995)) to build an asymmetric bivariate copula has the form:

$$C_1(u_1^{1-a_1}, u_2^{1-a_2}) C_2(u_1^{a_1}, u_2^{a_2}),$$

where C_1 and C_2 are two different copulas. a_1 and a_2 are the shape parameters for the Khoudraji device and range between 0 and 1 ($a_1, a_2 \in [0,1] \in \mathbb{R}$).

In order to determine the initial values for the parameters in the fitted asymmetric copulas, a sensitivity test is carried out by estimating the initial values that provided the best fitting theoretical asymmetric copula on the first 100h of empirical data. Then, these initial values are used to compute the asymmetric copula on the whole dataset. The Nelder-Mead optimization method (Lagarias et al. (1998)) has been employed in both the sensitivity test and the final computation of the asymmetric copula.

Table 1 Summary of the atmospheric, wave and statistical parameters that appear in the text

Source	Variables	Derived parameters
Radar and pressure sensor	total sea level (SL)	Moving average, 5 th and 95 th percentile of absolute difference of total sea level (dSL) relative difference of total sea level ($rdSL$) ; moving bias, centered-root-mean-square-error (CRMSE); scale error ($eesc$) of a moving linear regression applied to SL of both sensors; stationary dSL , $rdSL$, bias, CRMSE; intercept, slope, R-value, $eesc$ of a stationary linear regression applied to SL of both sensors
Radar	spectral significant wave height ($H_{m0,agit}$), mean absolute wave period in the harbor ($T_{m02,agit}$)	
Model	$H_{m0,mod}$, $T_{m02,mod}$, wind velocity (V_w)	
Buoy	$H_{m0,buoy}$, $T_{m02,buoy}$	

Table 2 Stationary statistical results for total sea level (*SL*) of different sampling rates.

	1-min	5-min	Hourly
Maximum difference in <i>SL</i> (<i>dSL</i>) [mm]	154.00	120	85
Minimum <i>dSL</i> [mm]	-182.00	-95	-68
Scale error (<i>eesc</i>) [%]	-0.34	--	--
p-value of the normality test on the relative difference in <i>SL</i> (<i>rdSL</i>)	0 (not a Gaussian distribution)	--	--
Bias of <i>SL</i> [mm]	-4.35	-4.56	-4.57
Centered-root-mean-square-error of <i>SL</i> [mm]	11.52	9.51	9.15

Table 3 The following variables are given for periods of bias $\leq -10\text{mm}$, in 2018. The wave parameters are: the maximum spectral significant wave height and the average mean absolute wave period within the Mazagón Harbor ($H_{m0,agit}$ and $T_{m02,agit}$); the same parameters from the nearby buoy of Cádiz ($H_{m0,buoy}$ and $T_{m02,buoy}$), and a node from the Puertos del Estado wave model database ($H_{m0,mod}$ and $T_{m02,mod}$) (Fig. 2). The wave direction from the model (Dir_{mod}) is also given. Storm Emma is highlighted.

<i>Period</i>	<i>Mean bias</i>	<i>H_{m0,agit}</i>	<i>T_{m02,agit}</i>	<i>H_{m0,buoy}</i>	<i>T_{m02,buoy}</i>
<i>month-day hour:minute</i>	<i>[mm]</i>	<i>[m]</i>	<i>[s]</i>	<i>[m]</i>	<i>[s]</i>
Mar-01 10:00 to Mar-18 01:40	-17.61	0.09	3.95	03.03	6.16
Apr-18 13:40 to Apr-22 02:40	-11.11	0.08	4.37	2.22	5.32
Aug-19 00:20 to Aug-22 05:40	-10.12	0.05	3.67	1.20	4.41
Aug-25 04:40 to Aug-26 13:00	-10.13	0.03	3.23	0.48	04.05
<i>Period</i>	<i>H_{m0,mod}</i>	<i>T_{m02,mod}</i>	<i>Dir_{mod}</i>		
<i>month-day hour:minute</i>	<i>[m]</i>	<i>[s]</i>	<i>[°]</i>		
Mar-01 10:00 to Mar-18 01:40	1.96	06.01	235.61 (south-west)		
Apr-18 13:40 to Apr-22 02:40:00	0.69	4.67	185.13 (south)		
Aug-19 00:20 to Aug-22 05:40	0.29	3.32	168.21 (south)		
Aug-25 04:40 to Aug-26 13:00	0.17	2.41	248.58 (west)		

Table 4 Same as table 3, but the parameters are given for a CRMSE surpassing 15mm. Storm Emma is highlighted.

<i>Period</i>	<i>Mean CRMSE</i>	<i>H_{m0,agit}</i>	<i>T_{m02,agit}</i>	<i>H_{m0,buoy}</i>	<i>T_{m02,buoy}</i>
<i>month-day hour:minute</i>	<i>[mm]</i>	<i>[m]</i>	<i>[s]</i>	<i>[m]</i>	<i>[s]</i>
Feb-26 20:00 to Mar-26 00:00	18.54	0.09	3.94	2.94	5.95
<i>Period</i>	<i>H_{m0,mod}</i>	<i>T_{m02,mod}</i>	<i>Dir_{mod}</i>		
<i>month-day hour:minute</i>	<i>[m]</i>	<i>[s]</i>	<i>[°]</i>		
Feb-26 20:00 to Mar-26 00:00	1.91	5.79	237.37 (south-west)		

Table 5 Main tidal constants at the study area, in 2017-2018.

Name	Description	Difference between amplitude of pressure sensor and radar [mm]	Difference between phase of pressure sensor and radar [°]
S2	Principal solar semidiurnal	-0.0656	-0.31
M2	Principal lunar semidiurnal	-0.3894	-0.28
K1	Lunar diurnal	0.0284	-0.52
S1	Solar diurnal	0.3556	1.25
MM	Lunar monthly	-0.1658	-13.01
SA	Solar annual	-0.0382	0.49

Table 6 Comparison of total sea level (SL) in a sensor (after applying alternative methods to it) with the original SL in the same sensor and the original SL in the other sensor. The used statistics are the moving bias and the moving centered-root-mean-square-error (CRMSE). At the end of the table, the original radar SL (SL_{radar}) and the original pressure sensor SL ($SL_{pressure}$) are compared to each other, in order to see the improvement reached with the alternative methods. The time period is March 2nd of 2018.

	Original SL_{radar}	Original $SL_{pressure}$
Low-pass filter applied to SL_{radar}	Bias ranging between - 7.0mm and 7.0 mm	CRMSE ranging between 7.5mm and 20.0 mm
Low-pass filter applied to $SL_{pressure}$	CRMSE ranging between 10.0mm and 75.0mm	Bias ranging between - 30mm and 75mm
Method of medians applied to SL_{radar}	Bias ranging between - 1.5mm and 2.0mm	CRMSE ranging between 3.0mm and 10.5mm
Method of medians applied to $SL_{pressure}$	CRMSE ranging between 3.0mm and 10.2mm	Bias ranging between - 1.9mm and 1.1mm
Original SL_{radar}	--	Bias ranging between 0mm and 40.0mm. CRMSE ranging between 10.0mm and 80.2mm.

Fig. 1

Fig. 2

Fig. 3

Fig. 4

(a)

(b)

(c)

Fig. 5

Fig. 6

(a1)

(a2)

(b1)

(b2)

(c1)

(c2)

(d1)

(d2)

Fig. 7

Fig. 8

(a)

(b)

Fig. 9

(a)

(b)

Fig. 10

(a)

(b)

Fig. 11

Fig. 12

Fig. 13

Fig. 14

Fig. 1 Sketch of the sea level station operating at Mazagón (Huelva, Spain). The acronyms and the variables in the image are, among others: tidal gauge benchmark (TGBM), free water surface (η), η of the radar sensor (η_{radar}), η of the pressure sensor ($\eta_{pressure}$), height of the radar sensor above the sea water surface (Y_{radar}), depth of pressure sensor below sea water surface (Y_{pres}), distance between TGBM and zero of REDMAR tide gauge (h_0), distance between pressure sensor and TGBM (Distance_p), pressure measured in the pressure sensor ($p_{measured}$), atmospheric pressure (p_{atm}) and specific weight of the water (γ_{water}).

Fig. 2 Maps of the study area. The figure above indicates the location of the station (radar and pressure sensor) and the buoy. The figure below is a detailed view of the Mazagón Harbor, and it provides the location of the point (coded 5033022) where the Puertos del estado wave model data are given.

Fig. 3 Extra contribution by moving waters to total sea level measured by pressure sensor (Φ''), depending on the wave number (κ) and the depth of the sensor (z).

Fig. 4 Comparison of the spectra of the total sea level (SL) measured in the radar and the pressure sensor. 1-min data from 2017-2018.

Fig. 5 The left part of the figures are the van den Castelee plot of the total sea level (SL), separated for different spectral significant wave heights near the station ($H_{m0,agit}$): a) from 0 to 0.2m, b) from 0.2m to 0.4m and c) above 0.4m. 1-min data from 2017-2018. The right part of the figure are the corresponding centered-root-mean-square-error (CRMSE). Note: full color, adapted for color vision deficiency, is available online.

Fig. 6 Van den Castelee plot of the entire sample of total sea levels (SL). The points in each region of the graph are grouped into hexagons, which display darker colors for denser points. 1-min data from 2017-2018.

Fig. 7 a1) Hourly local wind velocity (V_w) and b1) hourly local spectral significant wave height ($H_{m0,agit}$) at the station. c1) 1-min/5-min/hourly moving bias (plus trend of 1-

min moving bias). d1) 1-min/5-min/hourly CRMSE of the total sea level (SL) measured by the radar vs. the pressure sensor (plus trend of 1-min moving CRMSE, which is close to zero, as the CRMSE in each time-window is detrended). a2, b2, c2 and d2 are the corresponding parameters during Storm Emma. The dates are given in "year-month" in the left column and in month-day in the right column. Note: full color, adapted for color vision deficiency, is available online.

Fig. 8 Comparison of the spectra of the level of the free water surface (η) measured by the radar and the pressure sensor. March 2nd of 2018, during Storm Emma.

Fig. 9 a) Empirical cumulative probability function (ECDF) of drag coefficient vs. centered-root-mean-square-error (CRMSE) of the total sea level (SL), b) joint probability density of the fitted asymmetrical copula, the parameters of the copula and the log-likelihood of the fitting. Hourly data for a wind velocity above 11m/s (drag coefficient larger than 0.0012), for the period 2017-2018, are used. The counts represent the number elements within each cell of the heatmap. The diagonal red line is a visual reference and has a slope equal to 1. Note: full color, adapted for color vision deficiency, is available online.

Fig. 10 Same as Fig. 9, but considering $H_{m0,agit}$ instead of drag coefficient. Hourly data for a $H_{m0,agit}$ larger than 0.08m, for the period 2017-2018, are used.

Fig. 11 a) Storm surge (hourly data) and b) storm surge (hourly data) during Storm Emma. The dates are given in "year-month" in a), and month-day in b). Note: full color, adapted for color vision deficiency, is available online.

Fig. 12 Van den Castele plot of storm surge ($Surge$) in 2017-2018. $dSurge$ is the difference between the $Surge$ of the pressure sensor and the radar. Hourly data. The points in each region of the graph are grouped into hexagons, which display darker colors for denser points.

Fig. 13 Comparison of the spectra of the sea level due to low frequency oscillations (SL_{LF}) measured by the radar and the pressure sensor. 1-min data from 2018.

Fig. 14 Van den Castele plot of sea level due to low frequency oscillations (SL_{LF}). 1-min data from 2017-2018. The points in each region of the graph are grouped into hexagons, which display darker colors for denser points.