Review 1

"WAPOSAL: A Multi-Regional Wave Dataset from Satellite Altimetry for Significant Wave Height, Period Estimation, and Wave Power Density".

Submitted by S. Ponce de León et al to ESSD

Reviewer #1

General comments

This manuscript presents a new data set which contains in particular, an estimation of ocean wave power density in various coastal sites, at high spatial resolution. It is based on satellite radar altimeter measurements functioning in the "Synthetic Aperture Radar" (SAR) mode. The wave power density estimation relies on two variables classically provided by the satellite altimeter measurements (the normalized radar cross-section and the significant wave height H_S), which are estimated here using a processing specifically designed for SAR-altimeters and coastal regions (SAMOSA+ , see Dinardo et al, 2018), combined with an empirical function which relates these two variables to the zero-crossing wave period T_Z. This empirical function to estimate the zero-crossing period was already proposed and assessed by Gommenginger et al (2003), but the authors present here a new estimation of the relationship and a validation based on their satellite data set and buoy measurements.

The data set presented by Ponce de Leon et al. was obtained from two satellite data sets (Sentinel-3 A/B and Cryosat-2) and covers the Atlantic coasts of Europe, Madeira, the French Polynesian archipelagos, the Azores, the Canary Islands, the Mediterranean and Baltic Seas, and the coastal zone of French Guiana, over a period of 13 years (2011-2023).

This data set is new, and should interest users involved in the development of energy production from the waves. It is potentially an interesting downstream product from satellite observations. However, at the moment it lacks information on the possible biases in wave power density induced by the method developed by the authors.

As the data set also contains significant wave height and zero-crossing period with a high spatial resolution and in various coastal zones, it may also serve the validation of wave models specifically designed for these coastal zones and other oceanic and coastal applications.

The manuscript suffers however from a lack of details regarding i) the method to estimate the zero-crossing period and its validation, ii) sources of errors and error estimates concerning the wave power density , iii) the practical usefulness of the data set for wave energy applications, taking into account the time sampling at each site, which seems not very high.

After revision, I think that the manuscript could be published.

Specific comments

Section 2.2

- In Gommenginger, the empirical relationship between T_Z and X is proposed based on dimensional arguments, which implicitly assume that the wave dispersion relationship is the one valid for deep water. Here, you apply it in coastal environment with various depths. So, justify why this approach may still be used even if applied in shallow water.

Reply:

We thank the reviewer for this insightful comment. We agree that the empirical relationship proposed by Gommenginger et al. (2003) is derived from dimensional arguments that implicitly assume deep-water wave dispersion conditions.

In the revised version of the manuscript, we added this paragraph in section 2.2,

after this sentence: The samples with a 'misfit' greater than 4 counts are eliminated.

- Concerning the evaluation of the a and b coefficients of the T_Z equation: please be clearer on how it was done: (a,b) values estimated for each buoy location? Or mixing all locations? According to line 92, it seems to be done for each location, but improve the clarity of the earlier text. And explain in the manuscript why this has to be done for each location.

I guess that by doing so, the result of the fit (coefficients a,b) is adapted for each encountered condition of sea-state and bathymetry.

But please, comment in the text. Also, do you mix all satellites data when you perform the regression or do you separate the data sets between Cryosat, S3-A, S3-B?

Please also comment in the manuscript and justify (if you mix all the satellites, did you check for potential H_S or s_0 biases between the three satellites?)

Reply:

1. We thank the reviewer for pointing out the need for clarification.

The regression coefficients a and b in the formulation proposed by Gommenginger et al. (2003) were estimated independently for each wave buoy location, using only the collocated altimeter-buoy pairs for that site.

We acknowledge that this was not sufficiently clear in the original manuscript and have now revised the text to explicitly state this point.

We estimated the regression coefficients a and b for each wave buoy separately. Yes, we separate the datasets for both missions.

In the revised version of the manuscript, we added more information in response to your suggestions. Please, see below:

“The regression coefficients a and b were estimated independently for each buoy location using collocated altimeter and in situ observations. This site-specific calibration accounts for regional differences in wave climate, spectral characteristics, and coastal effects, which can influence the relationship between the altimeter-derived parameter X and the mean zero-crossing period Tz.”

Once the regression coefficients (a and b) are derived from the relationship between X (computed from altimeter measurements) and the zero-crossing wave period (Tz) from buoy observations, the two datasets must be collocated to form paired (X, Tz) samples.

For each buoy, the collocation procedure is performed independently for each satellite mission. For a given buoy record, all altimeter measurements within a 40 km radius are averaged to obtain representative values of significant wave height (Hs) and normalized radar cross section (σ_0). This approach produces mission-specific time series of satellite-derived parameters that are consistently matched to in situ observations at each buoy location. The temporal spacing between collocated pairs depends on buoy position, in situ sampling frequency, and each satellite mission's revisit characteristics.

The temporal spacing between collocated pairs varies with the buoy's geographic location and the dataset's sampling frequency.

The main advantage of this approach is that, once the wave energy period (Te) is estimated, the wave power density can be readily computed from satellite altimetry data. This enables large-scale assessments of wave energy resources in regions where in situ observations are not available.”

The estimation of the linear regression coefficients (a,b) is performed separately for each location, motivated by the fact that the relationship between the altimeter-derived parameter X and the mean zero-crossing period Tz is not strictly universal but depends on local sea state characteristics and environmental conditions, including:

- *Variations in wave climate (wind-sea vs. swell dominance)*
- *Differences in spectral shape and bandwidth*
- *Bathymetric effects influencing wave transformation in coastal regions*
- *Potential regional biases in altimeter measurements, especially in the nearshore zone*

As a result, a single global calibration may introduce systematic biases when applied across heterogeneous environments. Performing the regression at each buoy location allows the *empirical relationship to better adapt to local conditions and improves the accuracy of the retrieved Tz."*

2. *Regarding the joint use of three satellites, we did not combine data from all three when estimating the regression coefficients, nor did we assess potential inter-satellite biases in Hs or σ_0.*
However, for both CryoSat-2 and Sentinel-3 (with Sentinel-3A and Sentinel-3B processed jointly), we estimated bias, scatter index, and correlation coefficient relative to buoy measurements. The results for buoys around the British Isles show good agreement for both missions. Although this is not a direct inter-mission comparison, it suggests that the measurements from the two missions are broadly consistent; however, a dedicated cross-comparison would be required to quantify any systematic differences between them.

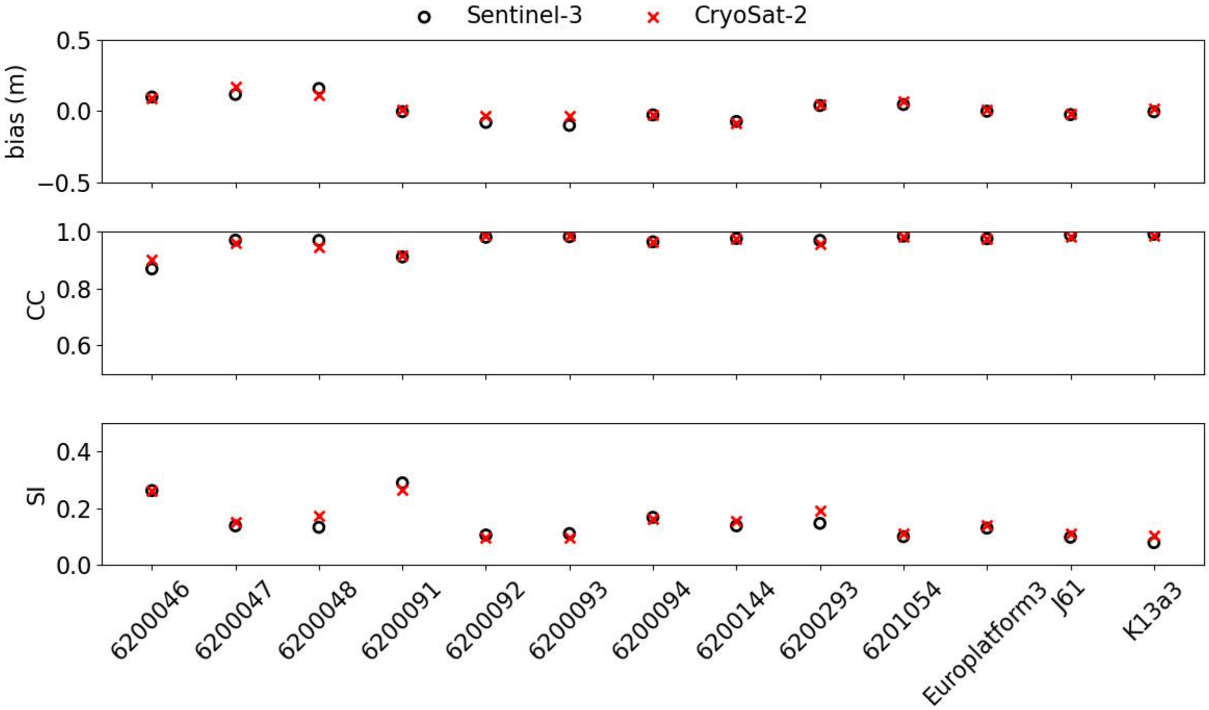


Fig. 1 Comparison of statistical metrics for Hs. Red crosses – Cryosat-2, circles – Sentinel-3A/B.

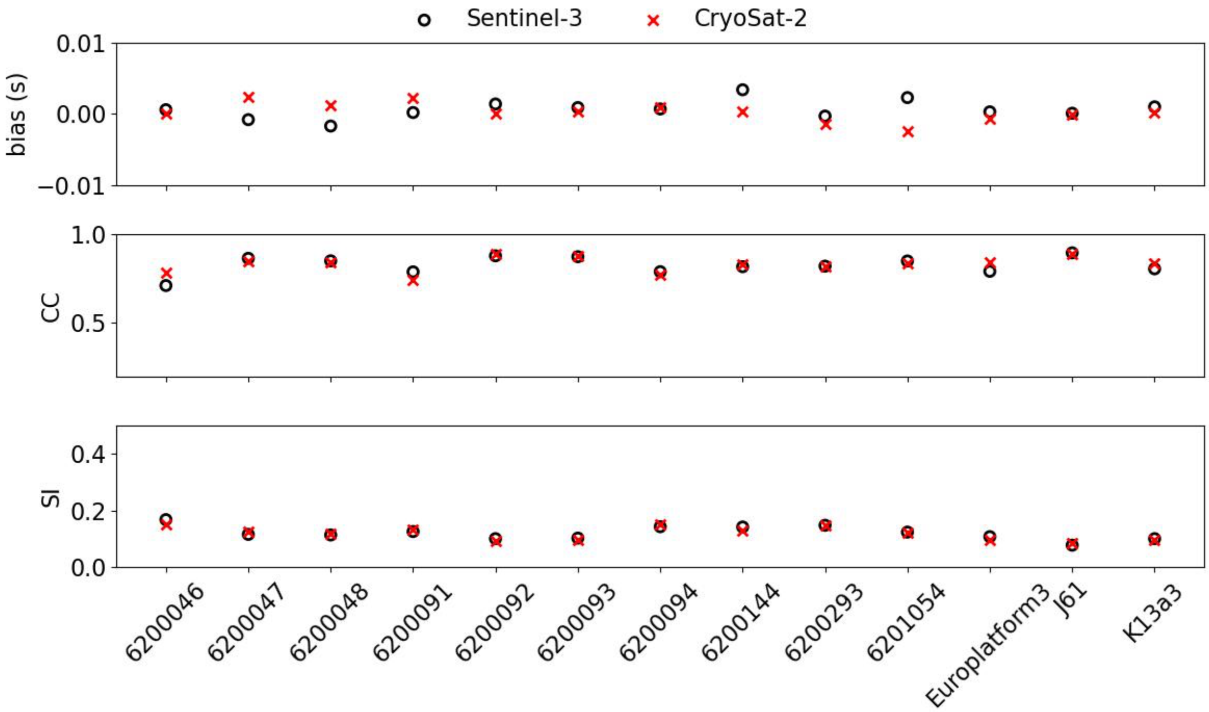


Fig. 2 Comparison of statistical metrics for Tz. Red crosses – Cryosat-2, circles – Sentinel-3A/B.

- training and validation sets: please give more details about how data sets have been separated: all buoys used in both steps? Different time periods for training and validation? Random selection?

Reply:

The dataset was randomly partitioned into two equal subsets: a training set and a validation set. The training set was used to estimate the regression coefficients, while the validation set was employed to evaluate model performance in terms of bias, scatter index, and correlation coefficient. This procedure was repeated 1000 times, and the resulting ensemble was used to compute averaged estimates of both the regression coefficients and the statistical metrics.

We have included this information to the text.

- about results of the regression: can you comment on the range of values obtained for (a,b) from the different sites? And also, is the quality of the regression the same for all sites? More generally what is the quality of the fit when estimating (a,b)?

Reply:

The quality of the fit when estimated (a,b) was validated for 82 wave buoys locations across the 11 different regions of the WAPOSAL project database.

The regression coefficients vary across regions where the WAPOSAL project applied the Gommenginger et al. 2003 method. This variation is directly related to the basin, the number of wave buoys (which varies widely from place to place), the quality and length of the records, and the sea state.

For the Mediterranean Sea (21 wave buoys), the regression coefficients were derived from CryoSat-2 data, with "a" ranging from 0.8801 to 1.6321 and "b" from 0.5204 to 1.7894. For Sentinel-3 A/B, the linear regression coefficients for "a" ranged from 1.1788 to 2.9321, and for "b" from 0.45892 to 1.5063.

In general, for the Mediterranean Sea, the statistical fits between wave data derived from satellite observations and those measured in situ by buoys are good. This indicates that the satellite data represent the field measurements quite well. As in the other study areas, the buoys with the best statistical fits are those farthest from the coast, which have a larger fetch.

In the North Atlantic, for example, for the French Façade, the CryoSat-2-derived linear regression coefficients ranged from 1.4446 to 2.116 for "a" and from 0.71117 to 2.6025 for "b." Using Sentinel-3 A/B, the linear regression coefficients for "a" ranged from 1.3648 to 2.2453 and from 1.0178 to 3.0135 for "b".

The quality of the regression varies across sites depending on the quality and length of the records. The correlation coefficient is high for the Hs and wave-period comparisons, indicating a strong linear relationship between the satellite series and the buoy data, with greater variability in the period data.

For the French Façade in general, the statistical fits between the periods and wave heights determined by satellite and in situ measurements are good. The best statistical fits generally occur at those buoys that are relatively far from the coast and are not surrounded by islands.

- Comparison of H_S and T_Z between satellite and buoy measurements (Fig.2): The results seem really good, but it would be useful to know whether all sites give equivalent quality or, regarding T_Z, whether some of them (in particular those more exposed to swell) give results of lower quality. This is an important information to provide to the future users of the data set.

Reply: Thank you for the comment. We performed a careful assessment using 82 wave buoys distributed across the 11 regions of the WAPOSAL project. The results are very interesting. We agree with the Reviewer that quality varies. Now provide a buoy-by-buoy assessment highlighting that not always swell-exposed sites perform worse for Tz.

We agree that the overall comparison shown in Fig. 2 may mask performance differences among buoy locations, and that this information is important for future users of the dataset.

Indeed, the correlation for Hs is usually higher than for Tz. In areas where swell is dominant, the correlation for Tz is lower at sheltered sites, such as wave buoy 6200046, than in areas dominated by local wind sea states, such as the Mediterranean. However, our results showed that there are also swell-dominated areas, such as the French Façade, where the fit for Tz shows high correlation.

The quality of the estimated wave power density is almost always high for the wave buoys in the Mediterranean and the North Atlantic because it depends primarily on Hs.

In the new version of the manuscript, we added section 2.3.1, containing Fig. 4 showing the cases of 2 buoys from North Atlantic and Mediterranean and Table 1 and the description for them:

"Two locations were selected in the North Atlantic (wave buoys 6200048 and 6200046) and one in the Mediterranean Sea (wave buoy 6100196). The positions of the buoys are shown in Figure 3. Table 1 presents the statistical metrics for Tz, Hs, and wave power density, along with the regression coefficients for the selected locations."

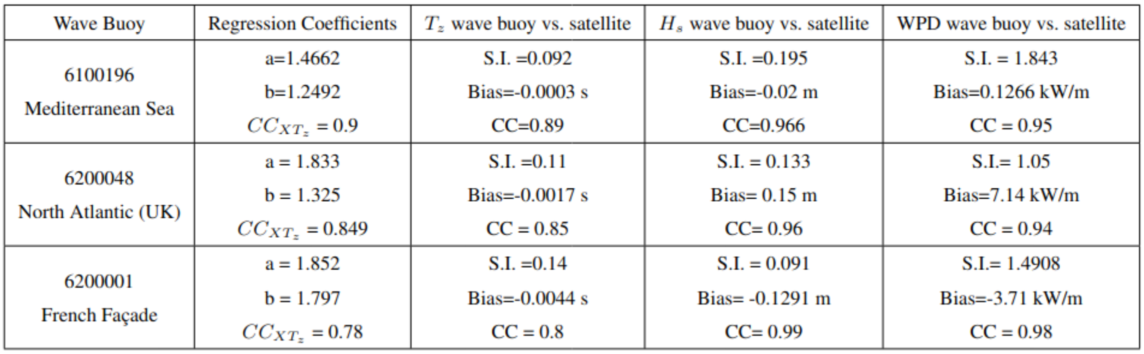
In addition, our results show that in some swell-dominated locations, the correlation is also good for Tz. As shown in Table 1, for the North Atlantic locations, the correlation is good for the French Facade wave buoy (0.8 for Tz, 0.99 for Hs, and 0.98 for WPD). For the wave buoy selected in the North of the UK (6200048), the correlation (using Sentinel-3) is also high for Tz (0.85), Hs (0.96), and wave power density (0.94). For the location shown for the Mediterranean (61000196) in deep waters of the Gulf of Leon, using Sentinel-3 A/B, the correlation for Tz is high (0.89), as are those for Tz (0.97) and WPD (0.95).

In Figure 4, an example comparing buoy and altimeter wave periods is shown for buoys in the Atlantic and Mediterranean seas. The root mean square error (RMSE) for Tz for buoy 61000196 in the Mediterranean Sea is 0.4 seconds, while for buoy 6200048 in the North Atlantic, the RMSE is higher, reaching 0.79 seconds.

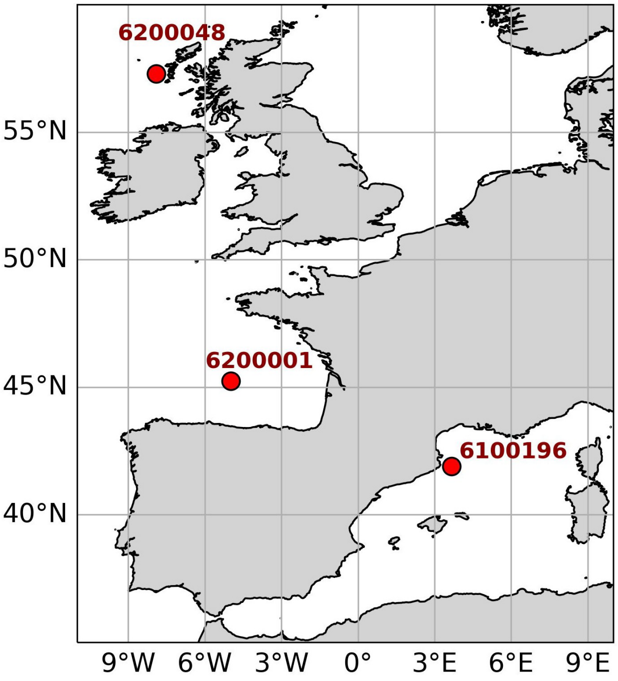
A site-by-site analysis shows that the agreement between satellite and buoy measurements varies with the local wave climate. Mediterranean locations, dominated by fetch-limited wind seas, generally show higher skill, particularly for Tz. In contrast, North Atlantic sites, more exposed to long-period swell and bimodal spectra, show slightly reduced performance in Tz, reflecting the empirical formulation's sensitivity to spectral shape. This spatial variability should be considered when using the dataset, especially in swell-dominated environments. This result highlights that the accuracy of satellite-derived wave period is not *uniform and depends on sea-state conditions, with better performance in wind-sea-dominated environments than in swell-dominated regimes.*

This behavior is consistent with the known limitation that the empirical relationship used to estimate Tz (Gommenginger et al. 2003) depends on the underlying spectral shape. Under swell-dominated or multi-peaked conditions, the relationship between the altimeter-derived parameter X and Tz becomes less robust.

Table 1. Statistics from the collocation of Sentinel-3AB and CMENS wave buoys. CC - correlation coefficient; S.I. - scatter index.



Wave Buoy	Regression Coefficients	T_z wave buoy vs. satellite	H_s wave buoy vs. satellite	WPD wave buoy vs. satellite
6100196 Mediterranean Sea	a=1.4662 b=1.2492 CC_{XT_z} = 0.9	S.I. =0.092 Bias=-0.0003 s CC=0.89	S.I. =0.195 Bias=-0.02 m CC=0.966	S.I. = 1.843 Bias=0.1266 kW/m CC = 0.95
6200048 North Atlantic (UK)	a = 1.833 b = 1.325 CC_{XT_z} = 0.849	S.I. =0.11 Bias=-0.0017 s CC = 0.85	S.I. = 0.133 Bias= 0.15 m CC= 0.96	S.I.= 1.05 Bias=7.14 kW/m CC = 0.94
6200001 French Façade	a = 1.852 b = 1.797 CC_{XT_z} = 0.78	S.I. =0.14 Bias=-0.0044 s CC = 0.8	S.I. = 0.091 Bias= -0.1291 m CC= 0.99	S.I.= 1.4908 Bias=-3.71 kW/m CC = 0.98



”

Figure 3. Locations of wave buoys for which data are presented in Table 1.

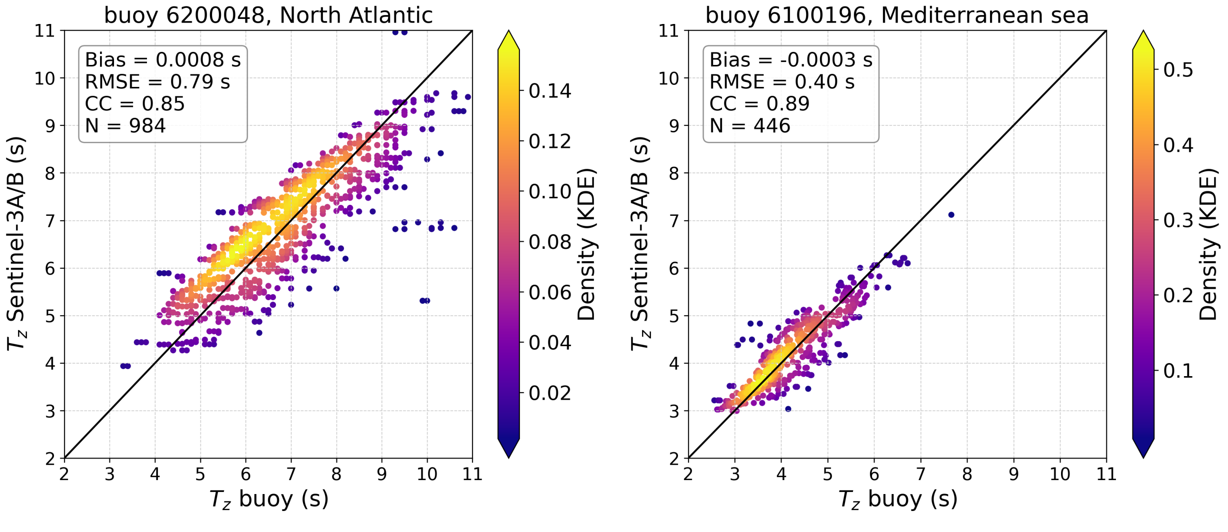


Figure 4. Scatter plots of wave period for the CMENS wave buoys 6200048 and 6100196, respectively, from Sentinel-3 A/B and the wave buoy at an Atlantic location (left) and a Mediterranean location (right).

- estimation of wave power density *P*: you provide an expression which relates the “energy period” T_e to the mean zero-crossing period T_z. However, it is known that the relation between T_e and T_z depends on the wave spectrum shape (see Cahill and Lewis, 2014 – same paper as in your reference list). So, you must model the underlying assumption on the wave spectrum shape associated with this relationship and estimate the possible bias on the wave power density due to other sea state conditions (or shape of wave spectrum). At least give a range of error in the estimate on *P*. Also, a statistical comparison of *P* from buoy and satellites using, for the buoys, T_e directly estimated from the buoy spectral measurements would give more confidence in the data wave power density set.

Reply:

We thank the reviewer for this important comment. Please see the statistics for wave period and wave power density in the new Table above. We fully agree that the relationship between the energy period Te and the mean zero-crossing period Tz is not universal and depends on the shape of the wave spectrum, as discussed in Cahill and Lewis (2014).

1) *Assumption on spectral shape*

In this study, the conversion from Tz to Te implicitly assumes a unimodal wave spectrum representative of wind-sea-dominated or moderately mixed sea states (e.g., JONSWAP-type spectra). Under these conditions, previous studies (including Cahill and Lewis, 2014) indicate that the Te/Tz ratio typically falls within a relatively narrow range.

2) *Expected uncertainty and impact on wave power density*

Based on the range of spectral shapes reported in the literature, the ratio Te/Tz can vary typically between ~1.05 and ~1.20 depending on sea state conditions (e.g., narrowband wind sea vs. broader or bimodal spectra). Since wave power density scales linearly with Te, this translates directly into an uncertainty in PP of approximately ±10–15% under typical conditions, and up to ~20% in more complex (e.g., swell-dominated or bimodal) sea states.

3) *Implications for coastal applications*

We note that in the study region, wave conditions are predominantly short- to intermediate-period wind seas that tend to exhibit relatively stable spectral shapes. Therefore, the variability in the Te/Tz ratio—and thus in P—is expected to remain within the lower end of the uncertainty range for most cases.

4) *Validation with buoy-derived energy period*

To the best of our knowledge, the wave buoys that measure wave spectrum belong to NDBC network and they are located near the coast of North America. In the areas of WAPOSAL project unfortunately we have not found buoy data on the wave spectrum or Te directly.

In the revised manuscript we added this information:

"Estimating wave power density relies on converting the mean zero-crossing period Tz to the energy period Te, which depends on the spectral shape (Cahill and Lewis, 2014). In this study, we implicitly assume a unimodal spectrum typical of wind-sea or mixed conditions. Under these assumptions, Te/Tz typically ranges from 1.05 to 1.20, introducing an uncertainty in wave power density of about 10–15%, and up to ~20% for more complex spectral conditions (e.g., bimodal sea states). This limitation is inherent to the use of bulk parameters and should be considered when interpreting the results."

Section 4 (example of data application)

Line 158-160 and Figure 3: please mention what is the time period of averaging (annual mean or over all the data set period ?)

Reply:

The figure caption was replaced with the following: Figure 4. Along-track mean wave power density for 11 regions, derived from Sentinel-3A/B altimeter measurements and averaged over the period 2016–2023.

In the revised paper, we updated the information in line 160 as follows:

"The along-track mean wave power density derived from Sentinel-3A/B data is shown in Figure 4 for the 11 study regions. Although these data can be further processed through interpolation, binning, or smoothing, the original along-track product preserves the measurements' intrinsic high spatial resolution. This high-resolution information is particularly valuable in coastal zones, where wave conditions exhibit strong spatial variability and may not be adequately captured by gridded products."

Lines 162-163 and figure 4: please mention what is the sampling time for this estimation of wave power density. Are the values plotted in Fig.4 averaged over a given time period? Or is it the minimum sampling time that is shown in Fig.4 ? If the sampling period is more than 1 or 2 days, discuss the usefulness for wave energy applications with such long gaps.

Reply: The time series presents collocated altimeter and buoy data. The time series has gaps, because the revisit time of the altimeter is limited. For Cryosat-2 revisit time is approximately 10 days for the area around the buoy with the radius of 50 km. We have also added the time series for Sentinel-3A/B. For the data from the two missions the revisit time is less: around 3 days. The time series indicates seasonal variations of wave power density. However, due to limitations *of sampling, some intense events are captured by Cryosat-2 and Sentinel-3A/B at different moments.*

In Fig.4- bottom panel, I suggest that you overlay the results from S-3 1/B and discuss their agreement with the estimations from Cryosat.

Reply: We do apologize because the time series from S3 A/B is missed. In the revised version this is fixed.
We decided to show the plots for two sensors on separate panels.

"The time series for this location, derived from 11 years of CryoSat-2 data and 7 years of Sentinel-3A/B data, are shown in Figure 5."

For comparison, the time series for the same location is also presented using ERA5 data. Seasonal variability in CryoSat-2 and Sentinel-3A/B is similar; however, due to differences in temporal sampling between the two missions, some intense wave events may be captured by one mission and missed by the other. For the 2016–2022 period, data from the two satellites complement each other and provide a more complete time series.

Suggested technical corrections

- Everywhere: please number the equations

Reply: Fixed.

- Lines 74-76 = Suggestion to make the sentence and reference to Gommenginger more clear: " In (Gommenginger et al., 2003), it was **suggested**, **based on dimensional arguments**, that T_Z is linearly related to **a variable named X, which is** a combination of significant wave height $\mathbf{H_S}$ and of the normalized radar cross section **σ0**, with X = $(H_S^2\ \sigma0)^{0.25}$, and σ0 expressed in natural units (non-dB)."

Reply: Thanks for the suggestion.

- Suggestion for Line 77: Therefore, following Gommenginger et al, 2003, using collocated Tz in situ wave buoy data and altimeter Hs and σ0 data, we have re-evaluated for each collocated data set the coefficients (a and b) of the following equation:

Tz= a X + b

Reply: Thanks for the suggestion.

- Before line 80: add a sentence to explain that the method was validated by Gommenginger et al. 2003, using Topex altimeter measurements and a set of NBDC buoys, and why it is necessary to establish this relationship in your analysis again.

Reply:

The empirical method was originally validated by Gommenginger et al. (2003) using TOPEX altimeter measurements and a set of NDBC buoys. However, the relationship must be re-established in the present study, as calibration can depend on several factors that differ from the original work, including the use of coastal-retracked altimetry (SAMOSA+), the nearshore application (down to ~5 km from the coast), and the regional wave climate (Mediterranean vs. North Atlantic, for example). These differences may affect both the altimeter-derived parameter X and its relationship with Tz, thus justifying a new, dataset-specific calibration.

We find that:

- *The agreement on significant wave height (Hs) remains consistently high across all locations.*
- *The variability in performance is mainly observed in Tz, as expected from its stronger sensitivity to spectral characteristics.*

We have now included buoy-specific statistical metrics and added a discussion of these regional differences. This provides a clearer indication of where the dataset performs best and where additional caution is required (see Table 1).

Text added to the manuscript:

"Although the relationship proposed by Gommenginger et al. (2003) was originally calibrated using TOPEX altimeter data and NDBC buoys, a recalibration is required here due to differences in the altimeter processing (use of SAMOSA+ retracker), the proximity to the coast, and the regional wave climate. These factors can influence the altimeter-derived parameter and its relationship with Tz, justifying a site-specific calibration."

- Line 93: suggestion; **Using the triplets** (σ0, Hs from altimeters Tz from buoy), the regression coefficients a and b were determined **for each buoy location**

Reply: Fixed.

- Line 109: inconsistency between the reference given here and the reference in the list at the end of the manuscript. I think that the reference here is Cahill and Lewis, 2014, not Cahill, 2012.

Reply: Fixed.

- in the reference list: the doi and doi link for "Kasiulis et al, 2015" is not correct.

Reply: Fixed. Thank you for the comment.

Data description and data availability

Following the link given in the manuscript, I checked that I could access to the ncdf files (sorted per country, satellite, date, and satellite tracks). But I did not find the .zarr files mentioned in section 3 and in appendix A.

Reply: Thanks for the comment. Yes, .zarr files are available for download only via the Python environment or the GDAL command line. We have tested this option: when the python script is run, the content of .zarr file is directly attributed to the variable ds as described in the Appendix.

There is a small Inconsistency between the content of the manuscript and the general description found when following the link indicated in the abstract and at line 147 or 204 (https://doi.org/10.57780/ESA- 1AB8CF3). Indeed, in the landing page of https://doi.org/10.57780/ESA-1AB8CF3, the products that are mentioned are H_S and T_Z but not the wave power. However, the product handbook is consistent with the current manuscript. (see screen copy below).

Reply: We have notified the team responsible for hosting the dataset to correct the description