Solar passive distiller with high productivity and Marangoni effect-driven salt rejection

# On the inference of the Peak Wave Period using Satellite Altimetry measurements

1st Leonardo Gambarelli
*DIMEAS*
*Politecnico di Torino*
Torino, Italy
leonardo.gambarelli@polito.it

2nd Edoardo Pasta
*DIMEAS*
*Politecnico di Torino*
Torino, Italy
edoardo.pasta@polito.it

3rd Claudia Cecioni
*DICITA*
*Università degli Studi Roma Tre*
Rome, Italy
claudia.cecioni@uniroma3.it

4nd Paolo Brandimarte
*DISMA*
*Politecnico di Torino*
Torino, Italy
edoardo.pasta@polito.it

5th Giuseppe Giorgi
*DIMEAS*
*Politecnico di Torino*
Torino, Italy
giuseppe.giorgi@polito.it

***Abstract*—Satellite altimeters provide valuable measurements of ocean surface parameters such as significant wave height and wind speed but do not directly observe wave period, a key variable for sea state characterization and offshore applications. This work presents a machine learning-based methodology to estimate the peak wave period using altimeter-derived data. Leveraging a large dataset of numerical wave simulations and in-situ observations, we develop a two-step approach: first, a classification model discriminates between locally generated wind-sea and remotely generated swell conditions; second, separate regression models are trained for each wave type to predict the peak wave period. The models are implemented using Random Forests and incorporate spatial and temporal features to enhance accuracy. Results show high predictive performance, with mean absolute errors around 0.3 seconds and coefficients of determination exceeding 0.93 in test sets. The methodology enables enhanced sea state reconstruction from satellite observations, supporting improved ocean monitoring and offshore planning.**

***Index Terms*—Wave Period, Satellite Altrimetry, Random Forest, Numerical Simulation**

## I. Introduction

Wave energy is gaining traction as a promising and underexploited renewable energy source with the potential to significantly contribute to the decarbonization of the energy sector [1], [2]. Unlike wind and solar, wave energy is characterized by a high degree of predictability and temporal availability [3], making it an attractive candidate for reliable and dispatchable renewable power generation. This potential has been acknowledged at the policy level, with the European Commission setting ambitious deployment targets of at least 1 GW of installed ocean energy capacity—including wave and tidal—by 2030, and 40 GW by 2050 [4].

This work is partially part of the project NODES which has received funding from the MUR – M4C21.5 of PNRR with grant agreement no.ECS00000036 (mainly the first and fifth authors). The authors are also part of the project AIMS: Artificial Intelligence to Monitor our Seas – funded by the European Union - NextGeneration EU - Ministry of University and Research - under the PRIN 2022 PNRR Call for Proposals (D.D.1409 of 14/09/2022).

Despite these prospects, wave energy technologies have yet to achieve commercial viability, mainly due to their elevated Levelised Cost of Energy (LCoE) [5]. Several factors contribute to this challenge, for example the lack of a single device architecture that has proven optimal across the broad range of sea states encountered globally [6], necessitating site-specific designs that complicate standardization and scaling.

Another critical barrier is the limited availability of comprehensive, high-quality wave datasets on a global scale. Accurate knowledge of wave parameters such as significant wave height, period, and direction is essential for the design, optimization, and siting of wave energy converters (WECs) [7]. The current scarcity of such data restricts the ability to assess resource potential and hampers the development of robust simulation and forecasting models.

One way to alleviate this data limitation is through satellite altimetry, which offers wide spatial coverage and repeated observations across the globe. Unlike fixed buoys, altimeters can capture wave conditions in multiple, remote locations. However, a key drawback is that they do not directly measure wave period [8], which is an essential parameter for wave energy applications. This paper addresses this limitation by presenting a Random Forest (RF) based methodology to infer the peak wave period from altimeter-derived measurements, leveraging the different dynamics of sea states and swell states. This methodology is then tested on numerical data coming from ERA5.

## II. Sea State Parameters Description

Sea states are typically characterized by three main parameters: wave height, wave period, and wave direction [9]. For irregular wave conditions, these are expressed through statistical measures—most commonly the significant wave height ($H_{m0}$), peak wave period ($T_p$), and mean direction ($D_m$), which are usually derived from sea surface elevation time series, recorded over intervals of several minutes at 1–2 Hz sampling rates.

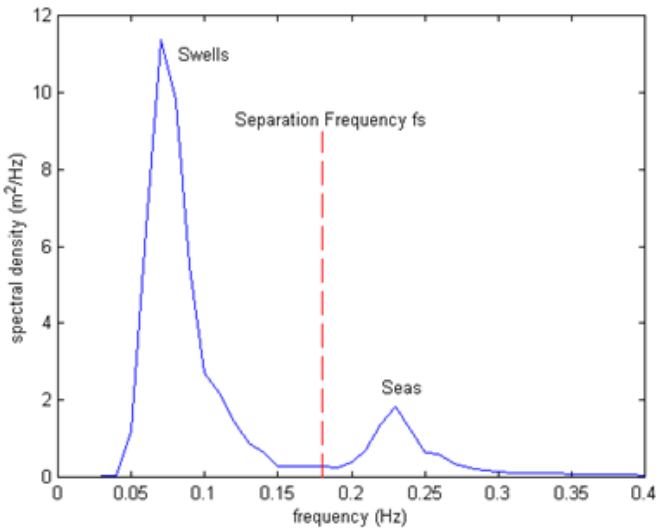


Fig. 1. A typical wave spectrum with dominant swell waves, showing the separation frequency and the distribution of swell and wind-seas energy with frequency.

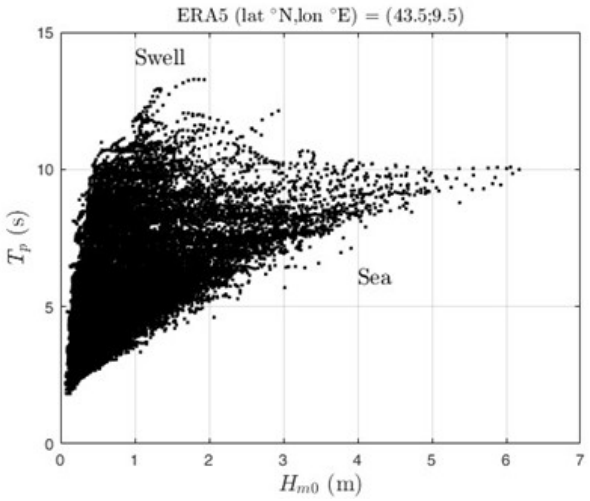


Fig. 2. Scatter diagram of peak wave period vs significant wave height. Total of 43824 wave data from ERA5 database at node 43.5°N, 9.5°E.

In the frequency domain, the significant wave height is given by

$$H_{m_0} = 4\sqrt{m_0} \tag{1}$$

where $m_0$ is the zero-th moment of the energy density spectrum. The peak period $T_p$ corresponds to the inverse of the frequency at which the energy spectrum peaks. In the time domain, the significant wave height, often noted as $H_s$, is defined as the average height of the highest third of observed waves and is generally equivalent to $H_{m0}$.

Wind-generated waves become taller and longer with stronger wind speeds and longer fetch (the distance over which wind blows steadily). Waves generated within the fetch are known as sea waves: they are irregular, short-crested, and travel roughly in the wind direction. As they move away from their generation area, they evolve into more regular, longer-period swell waves, characterized by lower height and more uniform directionality [10].

Typically, sea waves have periods of 5–12 seconds, while swell can reach 15–20 seconds [11]. Mixed sea states may present similar bulk parameters ($H_{m0}$, $T_p$) despite different compositions, making it difficult to distinguish between wind sea and swell. In such cases, spectral analysis is required to resolve the energy distribution across frequencies, with swell generally appearing as a peak at lower frequencies (Figure 1).

## III. Considered Case Study

The proposed methodology is based on an analysis of wave data extracted from the ERA5 reanalysis [12] at the location 43.5°N, 9.5°E, covering the period from January 1, 2018, to December 31, 2022. Located in the central Ligurian Sea, the study site experiences a mixed wave climate: short, fetch-limited wind-seas driven by local mistral and tramontane bursts alternate with long-period swells that propagate from the western Mediterranean and North Atlantic. This combination of energetic but steep sea states and lower-steepness swells offers an ideal test bed for the two-step classification-and-regression framework presented here.

Figure 2 shows a scatter plot illustrating the relationship between peak wave period $T_p$ and significant wave height $H_{m0}$.

As expected, $T_p$ generally increases with $H_{m0}$. Two distinct sea state regimes can be observed: swell conditions, characterized by low wave heights (typically below 2.5 m) and long wave periods, and wind-sea conditions, where both $H_{m0}$ and $T_p$ increase together, following a power-law trend.

To differentiate swell from wind-sea more rigorously, spectral analysis can be performed to identify a separation frequency (like in Figure 1). Additionally, the wave steepness $\epsilon$, defined as the ratio between wave height and wavelength (the latter being proportional to the square of $T_p$), is used as a discriminant [13]. A threshold value of $\epsilon = 0.015$ was found to effectively separate the two regimes: swell waves ($\epsilon < 0.015$) and sea waves ($\epsilon < 0.015$), as shown in Figure 3.

Since wind speed is also available from altimeter data, Figure 4 presents a scatter plot of $T_p$ versus wind speed.

While higher wind speeds are generally associated with higher $T_p$ values, longer wave periods are also observed at low wind speeds—typical of swell propagating from distant sources. However, significant overlap exists, particularly for low $H_{m0}$ and wind speed values, where both sea and swell states may occur, introducing uncertainty in wave period inference from wind and wave height alone. This ambiguity highlights the need for more advanced methods to reliably infer wave period from satellite-observed parameters.

## IV. Wave Period Inference Methodology

The first step in estimating the wave period involves distinguishing whether a given sea state corresponds to wind

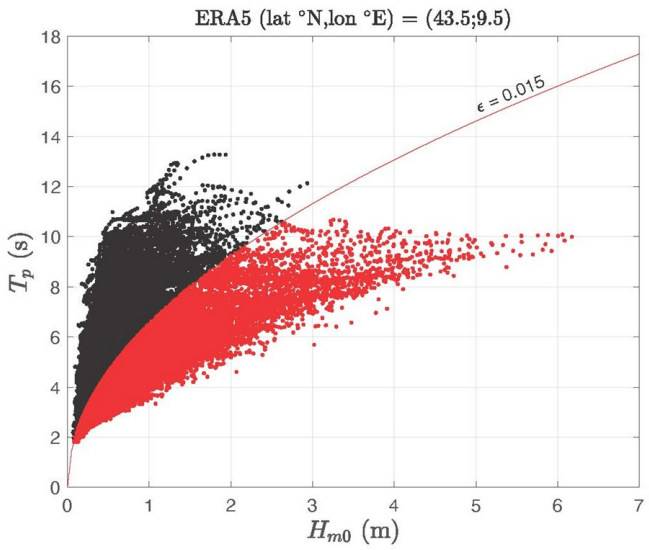


Fig. 3. Scatter diagram of peak wave period vs significant wave height. Red dots are relative to sea state with a wave steepness higher than 0.015, while black dots refer to wave steepness lower than 0.015. Total of 43824 wave data from ERA5 database at node 43.5°N, 9.5°E.

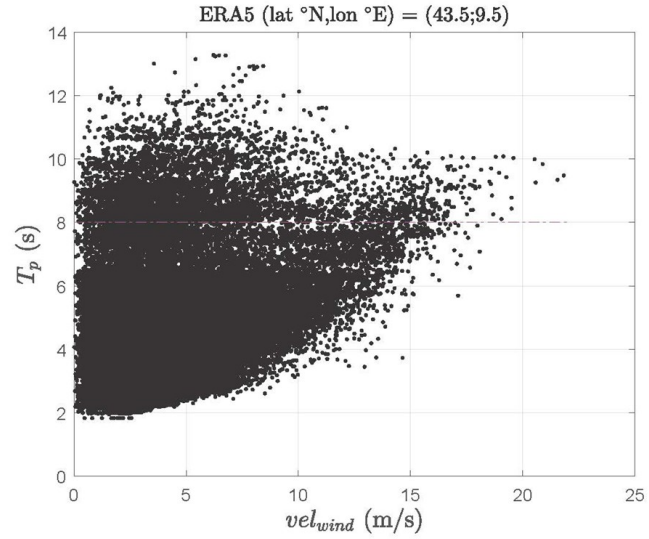


Fig. 4. Scatter diagram of peak wave periods versus wind velocities. Total of 43824 wave data from ERA5 database at node 43.5°N, 9.5°E.

sea or swell, as these regimes exhibit different dynamics and must be treated separately in the subsequent modelling phase. Specifically, two dedicated Random Forest (RF) regressors are trained: one for estimating the peak wave period $T_p$ for sea conditions, and another for swell conditions. Accurate classification is therefore crucial for reliable period inference.

Although this task is naturally framed as a classification problem, it was implemented as a regression problem, where the model is trained to predict the probability of belonging to a swell state ($Prob_{swell}$). Since the sea and swell probabilities are complementary, the probability of sea state occurrence can be simply computed as:

$$Prob_{sea} = 1 - Prob_{swell} \tag{2}$$

The complete dataset, consisting of 43,824 states, was labelled as either sea or swell based on wave steepness criteria (see Figure 3). The data was then randomly divided into training, validation, and test sets using an $80\% - 10\% - 10\%$ split to train and evaluate the Random Forest (RF) model for sea state classification.

The output of this RF was a continuous value between 0 and 1, representing the estimated probability of a swell condition. Predictions above 0.5 were classified as swell, while those below 0.5 were classified as sea.

Following the classification step, the subsets corresponding to sea and swell conditions were further split—using the same $80\% - 10\% - 10\%$ partitioning—into separate training, validation, and test sets. These subsets were used to independently train two additional RF regressors, each designed to predict the peak wave period $T_p$ from altimeter-derived parameters, tailored to their respective wave regime.

All Random Forest models described above use the following input features:

1) $H_{m0}$;
2) Wind Speed;
3) Spatial coordinates latitude and longitude;
4) Temporal features, encoded as four separate scalars (year, month, day, and hour) to capture both seasonal trends and diurnal variations.

All the described RFs were implemented with the 'RandomForestRegressor' command from scikit-learn using Python 3.12.7.

## V. Obtained Results

### A. State Classification Results

The confusion matrices of the validation and test set for the state classification problem are reported in Figure 5 and Figure 6. The results presented in these Figures are showing the classification made on the same location of the results provided within Figure 3, keeping the same wave steepness value to classify the wave condition.

As shown in Figure 5 and Figure 6, the trained classification model successfully identifies the majority of unseen sea states, with a misclassification rate of approximately $15\%$ on both the validation and test sets. The similar performance across these sets also indicates that the model is not overfitting.

### B. Wave Period Prediction Results

The scatterplots of the real $T_p$ against the predicted ones, on the validaiton and test sets, are reported for the swell states in Figure 7 and Figure 8, and for the sea states in Figure 9 and Figure 10.

Table 1 and Table 2 present the performance of the algorithms for the swell and sea states, reporting the achieved Mean Absolute Error ($MAE$), Root Mean Square Error ($RMSE$), and coefficient of determination ($R^2$). The formulas for these performance metrics are reported below.

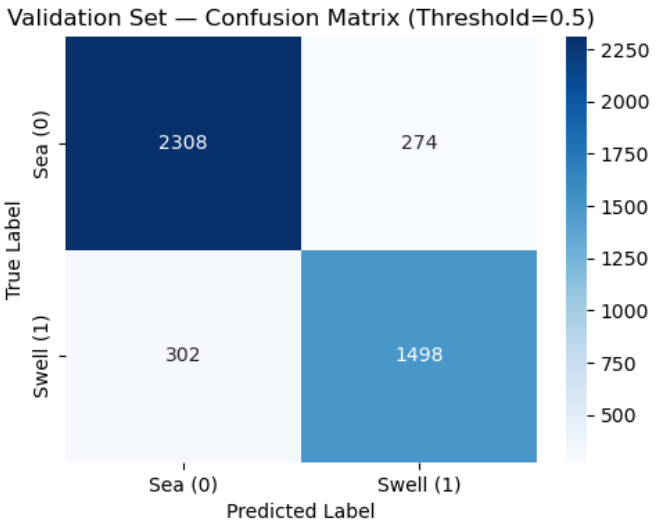


Fig. 5. Confusion matrix of the classification tree on the validation set.

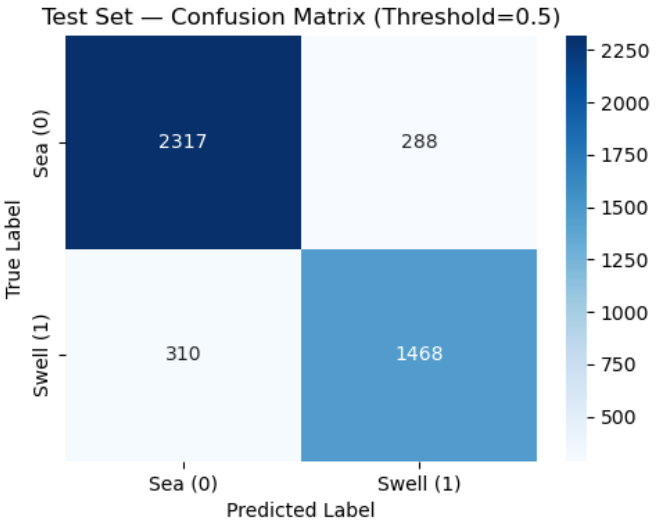


Fig. 6. Confusion matrix of the classification tree on the test set.

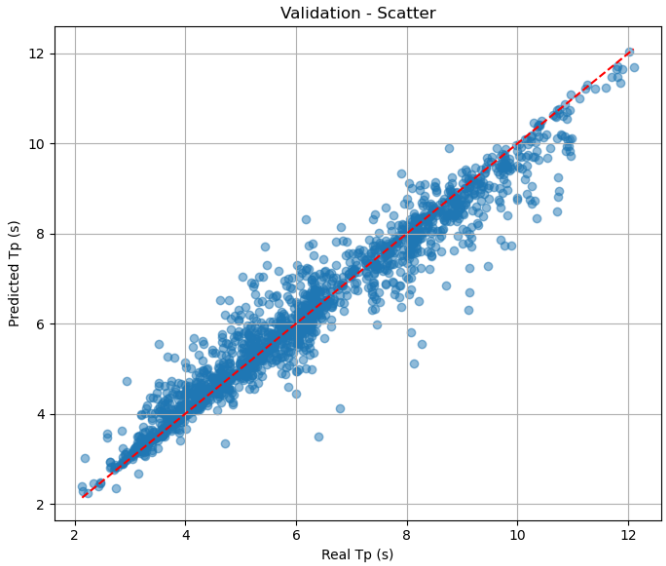


Fig. 7. Validation scatterplot for the swell states

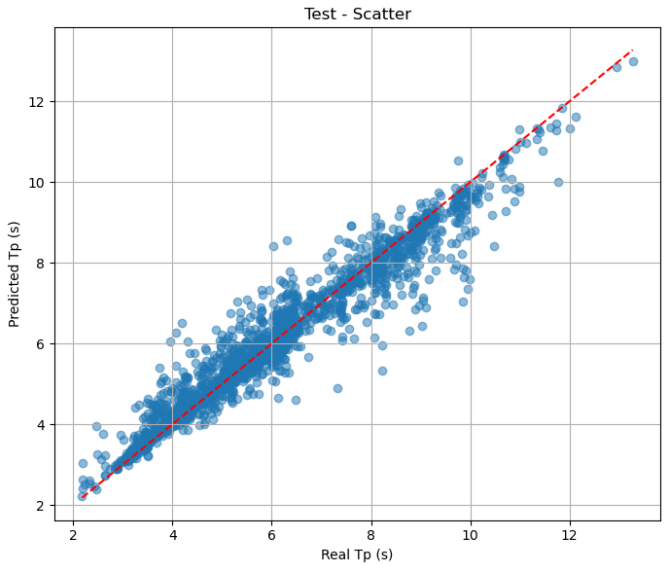


Fig. 8. Test scatterplot for the swell states

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{3}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{4}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{5}$$

where $y_i$ and $\hat{y}_i$ are the observed and predicted values, respectively; $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ is the sample mean of the observations; and $n$ is the number of data points.

The reported figures and tables show that the regression trees accurately estimate the peak period using spatial, temporal, and altimeter data, with low errors ( 0.3s) and high explained variance across validation and test sets. Performance is slightly lower for swell states, indicating greater difficulty in predicting wave periods in those conditions.

## VI. Conclusions

This study presents a preliminary yet effective approach for estimating the peak wave period $T_p$ from satellite altimetry data, using only variables accessible from satellite altimetry measurements, such as significant wave height and wind speed. By adopting a two-step Random Forest (RF) framework—first

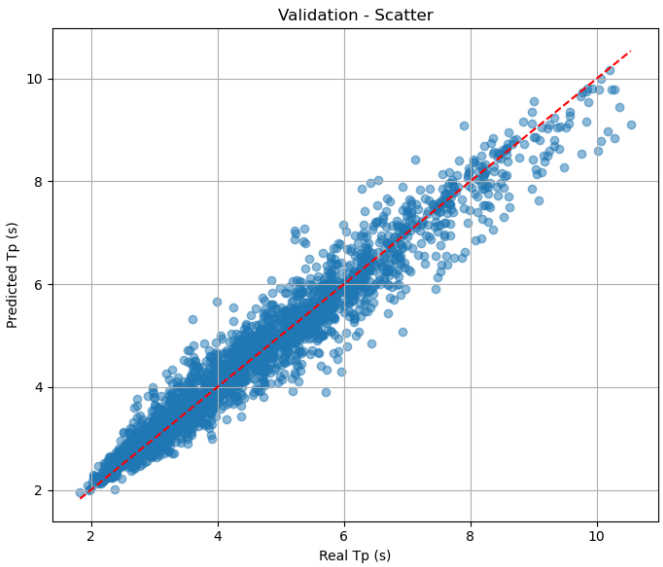


Fig. 9. Validation scatterplot for the sea states

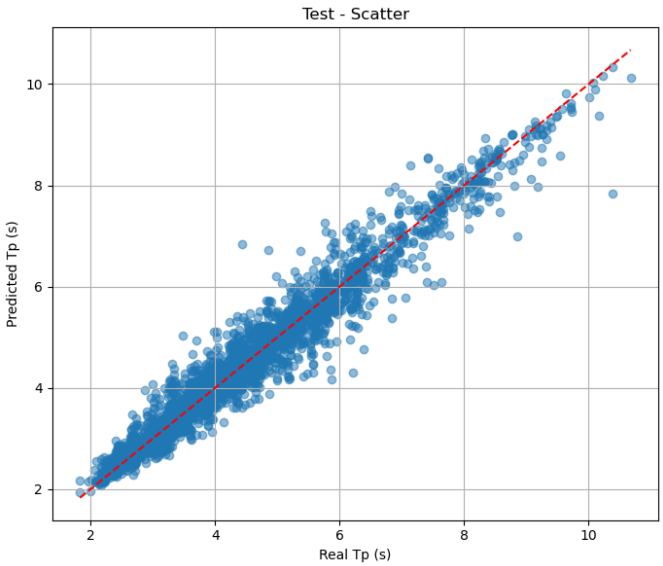


Fig. 10. Test scatterplot for the sea states

TABLE I
OBTAINED METRIC PERFORMANCES OF THE REGRESSION TREE FOR THE SWELL STATES

| Metric | Dataset | | |
|---|---|---|---|
| | ***Training*** | ***Validation*** | ***Test*** |
| MAE (s) | 0.1272 | 0.3425 | 0.3521 |
| RMSE (s) | 0.1987 | 0.5236 | 0.5349 |
| $R^2$ | 0.9905 | 0.9355 | 0.9285 |

MAE: Mean Absolute Error, RMSE: Root Mean Square Error.

TABLE II
OBTAINED METRIC PERFORMANCES OF THE REGRESSION TREE FOR THE SEA STATES

| Metric | Dataset | | |
|---|---|---|---|
| | ***Training*** | ***Validation*** | ***Test*** |
| MAE (s) | 0.1003 | 0.2828 | 0.2635 |
| RMSE (s) | 0.1430 | 0.3964 | 0.3689 |
| $R^2$ | 0.9925 | 0.9435 | 0.9479 |

MAE: Mean Absolute Error, RMSE: Root Mean Square Error.

classifying sea states as either wind-sea or swell, and then applying dedicated regressors for each regime—the methodology leverages the distinct physical dynamics of the two wave types.

The results demonstrate that Random Forest models are capable of accurately inferring the wave period, achieving mean absolute errors around 0.3 seconds and high coefficients of determination ($R^2 > 0.93$) across both validation and test datasets. Notably, the performance is slightly better for wind-sea conditions, while swell states exhibit greater variability and remain more challenging to predict.

This work highlights the potential of machine learning to enhance the value of altimeter missions by inferring otherwise unmeasured ocean parameters. Future developments may involve incorporating additional satellite-derived features, extending the method to different geographical regions, and exploring deep learning architectures for further performance gains.

## REFERENCES


[1] A. Terrero González, P. Dunning, I. Howard, K. McKee, and M. Wiercigroch. (2021). "Is wave energy untapped potential?" International Journal of Mechanical Sciences, 205, 106544.

[2] K. von Schuckmann, E. Holland, P. Haugan, and P. Thomson. (2020). "Ocean science, data, and services for the UN 2030 Sustainable Development Goals." Marine Policy, 121, 104154.

[3] E. B. Mackay, A. S. Bahaj, and P. G. Challenor. (2010). "Uncertainty in wave energy resource assessment. Part 2: Variability and predictability." Renewable Energy, 35(8), 1809–1819.

[4] European Commission, An EU Strategy to harness the potential of offshore renewable energy for a climate neutral future, COM(2020) 741 final, Nov. 2020. [Online].

[5] E. Giglio, E. Petracca, B. Paduano, C. Moscoloni, G. Giorgi, and S. A. Sirigu. (2023). "Estimating the Cost of Wave Energy Converters at an Early Design Stage: A Bottom-Up Approach." Sustainability, 15(8).

[6] A. Trueworthy and B. DuPont. (2020). "The Wave Energy Converter Design Process: Methods Applied in Industry and Shortcomings of Current Practices." Journal of Marine Science and Engineering, 8(11).

[7] O. Yaakob, F. E. Hashim, K. Mohd Omar, A. H. Md Din, and K. K. Koh. (2016). "Satellite-based wave data and wave energy resource assessment for South China Sea." Renewable Energy, 88, 359–371.

[8] J. Wang, L. Aouf, S. Badulin, "Retrieval of wave period from altimetry: Deep learning accounting for random wave field dynamics," Remote Sensing of Environment, vol. 265, 112629, 2021.

[9] L. H. Holthuijsen. (2007). Waves in Oceanic and Coastal Waters. Cambridge University Press.

[10] F. Ardhuin, B. Chapron, and F. Collard. (2009). "Observation of swell dissipation across oceans." Geophysical Research Letters, 36(6).

[11] J. H. G. M. Alves. (2006). "Numerical modeling of ocean swell contributions to the global wind-wave climate." Ocean Modelling, 11(1–2), 98–122.

[12] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger,

S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J.-N. Thépaut. (2020). "The ERA5 global reanalysis." Quarterly Journal of the Royal Meteorological Society, 146(730), 1999–2049.

[13] E. M. Bitner-Gregersen and A. K. Magnusson. (2014). "On the importance of wave steepness in the modelling of extreme wave conditions." Ocean Engineering, 86, 113–128.