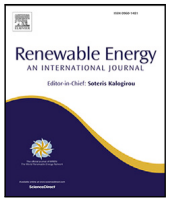
Bias correction techniques for uncertainty reduction of long-term metocean data for ocean renewable energy systems

Renewable Energy 219 (2023) 119404

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene



Bias correction techniques for uncertainty reduction of long-term metocean data for ocean renewable energy systems

Markel Penalba [a,b,*], Chao Guo [a,c], Ander Zarketa-Astigarraga [a], Giulia Cervelli [c], Giuseppe Giorgi [c], Bryson Robertson [d]

[a] Fluid Mechanics Department, Mondragon University, Loramendi 4, 20500 Arrasate, Spain
[b] Ikerbasque, Basque Foundation for Science, Euskadi Plaza 5, Bilbao, Spain
[c] Marine Offshore Renewable Energy Lab, Politecnico di Torino, 10129 Torino, Italy
[d] Civil and Construction Engineering, Oregon State University, Corvallis, OR 97331, USA

ARTICLE INFO

Keywords:
Offshore renewable energies
Resource assessment
Re-analysis datasets
Uncertainty reduction
Bias correction
Quantile-mapping

ABSTRACT

The design of effective and reliable solutions for Offshore Renewable Energy (ORE) technologies greatly relies on accurate metocean data. Uncertainties can significantly impact the design process. This paper presents a thorough evaluation of different bias-correction (BC) techniques applied to re-analysis datasets from diverse locations along the Spanish coast, presenting a multi-level evaluation procedure including novel, more appropriate statistical metrics beyond the traditional assessment techniques. First, the quality of raw ERA5 datasets is demonstrated to be rather poor, especially under extreme events, confirming the need for BC. Once the need is identified, results show that only the most sophisticated distribution-mapping BC techniques present the capacity to significantly improve the quality of the datasets, with the linearly-spaced Quantile-mapping (QM) showing the best overall performance for power production conditions (*PP*), reducing the bias by over an order of magnitude. In contrast, the Gumble Quantile-Mapping (GQM) outperforms the QM in survivability conditions (*Surv*). However, when computing overall performance of BC techniques, the predominance of *PP* hinders the relevance of *Surv*. Hence, adapted metrics with a more realistic balance between the *PP* and *Surv* regions are suggested, which show a better suitability of GQM providing more accurate estimations of the average power density and variability.

1. Introduction

A massive expansion of renewable energy sources is vital to transition from fossil fuels to clean energies and achieving a carbon neutral energy system worldwide. This transition is one of the key mitigation actions to fulfil the goals set out in the Paris Agreement [1] and the latest report from the International Panel for Climate Change (IPCC) [2] in order to avoid the most dramatic impacts of climate change [3]. Although mature and reliable renewable energy technologies already exist in the market, such as wind and solar energy, the size and speed of this transition will most likely require a substantial participation of other renewable technologies. The International Renewable Energy Agency (IRENA) estimates an increase of 14 TW on the total renewable energy installed capacity worldwide by 2050, meaning that current capacity will need to be multiplied by 5, which clearly illustrates the magnitude of the challenge [4]. Offshore renewable energy (ORE) systems are one of the alternatives that can assist the transition. In fact, The International Energy Agency predicts that about 45% of CO_2 emissions savings by 2050 will come from technologies that are still under development [5]. Offshore wind, for example, is expected to multiply the current worldwide installed capacity by 30 in the next 30 years [4]. Similarly, wave and tidal energy, although still immature, are also expected to contribute significantly to the future energy mix, with estimates showing a potential to cover about 10% of the future electricity demand [6,7].

Yet, wave energy converters (WECs), tidal energy converters and even floating offshore wind turbines (FOWTs), still need significant development in order to become competitive in the energy market. Key aspects for the development of these technologies include:

i. Optimising the designs of floating structures by reducing material use, while conserving their reliability and structural integrity,
ii. Increasing the energy generation capacity by considering non-linear hydrodynamics and designing control algorithms,

* Corresponding author at: Fluid Mechanics Department, Mondragon University, Loramendi 4, 20500 Arrasate, Spain.
E-mail address: mpenalba@mondragon.edu (M. Penalba).

https://doi.org/10.1016/j.renene.2023.119404
Received 17 April 2023; Received in revised form 6 September 2023; Accepted 2 October 2023
Available online 4 October 2023

iii. Enhancing the durability of key components, such as mooring lines and power take-off (PTO) systems via new materials and their design considering better adjusted operation regions, and
iv. Improving accessibility and availability by optimising the O&M strategies.

Despite their specific characteristics and requirements, the effective and reliable design of successful ORE technologies, including all the aspects mentioned above, relies on accurate metocean data. Design excellence comes from making confident decisions, which requires insight, not only data. The use of incomplete or inaccurate metocean data can result in misguided conclusions and, as a consequence, a poor understanding, incorporating a higher uncertainty to the design process where the uncertainty level is already significant [8]. In fact, these uncertainties lead to excessive conservatism in the design of ORE technologies, which, in turn, results in prohibitively large and expensive systems unable to compete in the current energy market [9]. As the resource assessment is the first link of the energy conversion chain of any ORE system, the uncertainty in metocean data affects the development of any stage in the chain, from the prediction of the system response to the estimation of the final energy generation [10,11]. Offshore Renewable Energy technologies are commonly designed with two particular operational modes, *i.e.* power production mode (*PP*) and survivability mode (*Surv*), dividing the metocean datasets into two significantly different regions. The former refers to the metocean conditions that enable the power production safely, while the latter refers to the metocean conditions under which the risk of damaging the ORE system is high and, thus, protecting the ORE device becomes the priority, interrupting power generation. Therefore, understanding and reducing the uncertainty of the metocean data across the whole operational domain including both regions, is of high relevance.

In addition, spatio-temporal variations (including inter- and intra-annual variability) [12], and the potential non-stationarity of the resource [13,14] suggest that long datasets are required for a better understanding of the resource. Accordingly, different international organisations recommend considering relatively long periods of data: the International Organization for Standardization (ISO) defines the required data as a function of the return period to be considered (25% of the return period of interest), meaning 10 years of data, for ORE systems with 20–30 years of lifespan [15]; while the Institute of Marine Engineering, Science & Technology (IMAREST) [16] suggests a significantly longer period of 30 years in order to accurately characterise extreme events.

The most common sources of metocean data in the ORE sector are observation buoys and re-analysis datasets. In addition, long-term hindcast data simulated by wind and wave models, such as Weather Research and Forecasting model (WRF), Simulating Waves Nearshore (SWAN) are also widely used. The latter are typically used for the assessment of relatively small areas, such as specific islands [17] or regional analyses [18–20], and commonly use re-analysis data as input [21]. Hence, observations and re-analysis datasets can be considered as the main sources of metocean data. However, such long periods of data are still difficult to cover with wave measurement buoys or other observation systems, and these systems tend to be complex and expensive. Furthermore, due to the harsh environment in the open ocean, the equipment installed in, or maintaining position for, measurement buoys often fail, resulting in relatively long periods for which no data is available. Therefore, re-analysis datasets and data from climate models are often used. These datasets cover very long time periods (can go back to 1900 in some cases [22,23]) and provide metocean data at any point in the world for virtually no cost. However, the main problem of re-analysis datasets is their limited accuracy under certain conditions. It is well-known that they exhibit systematic errors, particularly in extreme events. The calibration of re-analysis datasets enables the reduction of the differences (bias) between raw datasets and observations. That way, precise long-term metocean data that cannot be provided by pure observations can be generated for a more accurate assessment of ORE technologies.

1.1. Introduction to bias correction

Bias correction (BC) techniques enable scaling the values of the raw data to approach statistical properties of the observed data. These techniques are purely statistical tools and have become a common practice in climate and meteorological studies in the last two decades [24,25]. Therefore, the use of these techniques does not require any fundamental insight on the underlying physics of the models or data assimilation methods [26].

In principle, these techniques can be applied to any variable. Nevertheless, the quality of the corrected, or calibrated, datasets greatly depends on the quality of the reference dataset considered as the *ground truth*, which is usually represented by observations. In this sense, it should be noted that the objective of the different BC techniques should not be correcting or modifying the sensitivity of the climate models or data assimilation techniques. Therefore, BC should only be applied to the final results obtained from either approach. For further information about BC in climate and meteorological datasets, the reader is referred to [25]. Although these techniques can be used for correcting both re-analysis datasets and simulation results, this paper will focus only on the former, and, for the sake of clarity, re-analysis datasets will be referred to as assimilated data.

Among the different BC techniques suggested in the literature, the simplest approach is the *Delta* technique, which consists in merely subtracting from the assimilated data the mean difference between the assimilated and observed data over the whole period covered by the dataset. The *Delta* technique was the first suggested approach, but more complex and effective techniques have been suggested in different climate studies, mainly for temperature and precipitation. All these techniques can be classified into four main groups [24]:

1. The linear-scaling technique [27] enables correcting monthly mean values based on the difference between the assimilated and observed data, as in the Delta technique, but incorporates an additional factor based on the long-term monthly means.
2. The variance-scaling technique corrects both the variance and the mean in two steps: in the first step, only the mean is corrected, using the linear-scaling technique; the variance is corrected in the second step by means of a factor based on the ratio of the standard deviation (SD).
3. Power transformation techniques use a nonlinear correction function in the form of an exponential, which enables adjusting the variance statistics, as opposed to the linear-scaling technique.
4. The distribution mapping techniques aim at correcting the distribution function of the assimilated data to agree with the observations by means of a transfer function created based on the differences between the two datasets.

One of the most thorough reviews of different BC techniques is presented in [24] applied to hydrological studies, where the authors conclude that all BC techniques improve the quality of the simulated data, but only the *higher-skill* techniques, such as power transformation and distribution mapping, have the sufficient complexity to correct different statistical properties beyond the mean. In fact, the distribution mapping techniques have shown to be the most popular BC approaches in the last decade, regardless of the fields and variables. These techniques have been given different names in the different studies, such as 'probability mapping' [28] 'histogram equalization' [29], 'statistical downscaling' [30] or 'quantile–quantile mapping' [31].

1.2. Metocean data in the ORE literature

In contrast to the previous subsection, although the use of climate models and re-analysis datasets is relatively common in the ORE literature, the use of BC techniques is remarkably scarce. The

vast majority of the ORE studies use raw simulated or assimilated datasets regardless of the final application of the metocean data, such as production assessment [32], classification of the global resource [33], resource exploitability analysis [34], O&M assessment [35] and identification of extreme design conditions [8]. A few exceptions employ BC techniques [36–38] or analogous approaches integrating satellite, re-analysis and in-situ measurements [39]. However, these studies consider one approach and apply it without providing any comparative analysis or demonstration of its suitability against other similar approaches. To the best knowledge of the authors, the only study partially related to the ORE sector that presents a thorough benchmarking of different BC techniques is [40], where the authors exclusively focus on the significant wave height and use a multi-forcing 8-member global dynamic ensemble of wave climate projections in a global scale.

The re-analysis datasets commonly used for the analysis of ORE systems are generated by the European Centre for Medium-Range Weather Forecasts (ECMWF), which provides a set of different datasets [41]. The first atmospheric re-analysis covering the whole 20th century, from 1900–2010, is ERA20, which includes observations An updated version, although covering only the latest part of the 20th century, is the ERA-Interim dataset, which improves the data assimilation method and the wave model including shallow-water physics. Ulazia et al. compare both datasets over different areas around the world, *e.g.* the Bay of Biscay [22] and the Chilean coast [42], concluding that ERA-Interim provides significantly better results. However, the most recent re-analysis dataset by the ECMWF is the ERA5 dataset, *i.e.* the 5th generation ECMWF atmospheric reanalysis of the global climate, which provides a higher resolution (0.25° and 0.5° for the atmosphere and ocean, respectively) and assimilates more observations than any other previous re-analysis dataset. ERA5 is the most widely used and the most accurate dataset, outperforming both ERA20 and ERA-Interim re-analysis datasets. In fact, some studies use ERA5 as the reference dataset [43]. Therefore, the present study uses the ERA5 dataset as the source of metocean data.

This paper presents a comprehensive benchmarking of four different BC techniques applied to the most relevant metocean variables. The combination of the significant wave height (H_s), the peak period (T_p) and the wind speed (U_w) enables the assessment of the different crucial aspects used in the design process of ORE systems: power production, fatigue and extreme loading, system degradation, and O&M accessibility. Therefore, in the present study, H_s, T_p and U_w are obtained directly from the ERA5 re-analysis for the different analysed locations. The four different BC techniques include the *Delta* technique and three different *distribution mapping* techniques, which are assessed by using a novel metric suggested in the present paper. This novel metric enables simultaneously considering *PP* and *Surv* conditions when determining the most appropriate BC technique for the ORE sector. Also, the effectiveness of the BC techniques is evaluated in four different locations around the Iberian Peninsula, for which *in-situ* observations are also available. All the four locations have different metocean characteristics in terms of wave and wind resource, which generates the ideal pool for the evaluation of the different BC techniques.

The remainder of the paper is as follows: Section 2 describes the BC techniques and the systematic multi-level assessment methodology metrics to identify the most effective BC technique(s); Section 3 defines the geographical locations and characterises the ORE resource data upon which the different BC techniques are evaluated, including the statistical techniques for that evaluation, Section 4 presents the performance of the different BC techniques; and Section 5 draws the most relevant conclusions.

2. Bias correction techniques

In this study, the implementation of the BC techniques targets correcting the assimilated metocean data in order to improve the agreement with observations (referred to as *Obs* in the following), which are considered to be the *ground truth*. To that end, four BC techniques are applied to the different datasets.

2.1. Delta-change

As mentioned earlier in this paper, the *Delta*-change technique, first suggested in [44], is the simplest BC technique. It consists in adjusting the distribution of the assimilated dataset (y^{as}) by adding a constant correction factor (the *Delta* factor) calculated using the measured (y^{obs}) and assimilated datasets. More specifically, this constant correction factor is computed as the subtraction between the average values of the assimilated ($y^{\hat{as}}$) and the measured datasets ($y^{\hat{obs}}$). Hence, the dataset corrected via the *Delta* BC method is given as follows:

$$y_i^{BC} = y_i^{as} + (\hat{y}^{obs} - \hat{y}^{as}), \tag{1}$$

where $i = 1, \dots, N$, N being the number of points considered from the datasest. This number can vary depending on the available information of the measuring buoy at each location and for each analysed variable. It should be noted that, in the present study, y can represent three different variables, *i.e* H_s, T_p and U_w.

2.2. Full distribution mapping

The full distribution mapping (FDM) technique is the most elementary method within the distribution mapping group. In distribution mapping techniques, the correction factor is defined as a statistical relationship between the assimilated cumulative density function (CDF_{as}) and measured CDFs (CDF_{obs}). In both cases, and in all the BC techniques presented in this manuscript, the CDF is computed following the empirical cumulative distribution function. The full distribution mapping uses the whole CDF all at once to identify the statistical relationship between assimilated and observed datasets. In this case, the statistical relationship (X^{FDM}) is calculated as the difference between the inverse CDF of the assimilated (CDF_{as}^{-1}) and observed datasets (CDF_{obs}^{-1}):

$$X^{FDM} = CDF_{as}^{-1} - CDF_{obs}^{-1}, \tag{2}$$

This difference between the inverse CDFs is then fitted with an n-order polynomial function in order to transform the correction factor X^{FDM} into time-domain correction factors to be applied to the raw assimilated dataset. Hence, the dataset corrected via the FDM technique is given as,

$$y_i^{BC} = y_i^{as} + f(X^{FDM}, n), \tag{3}$$

where f denotes the n-order polynomial function and n the order of this polynomial function ($n = 3$ in the present study).

The *FDM* technique usually obtains a better improvement of the assimilated data than the *Delta* technique, but both are expected to provide a poor representation of the extreme conditions due to the predominance of the milder conditions in the definition of the correction factors.

2.3. Quantile mapping

The quantile mapping (QM) method, also known as the empirical quantile mapping, is also classified within the distribution mapping group and is almost identical to the FDM method. The only difference is that the correction factor X^{QM} is identified at each quantile (q_j) of the CDF, as illustrated in Fig. 1.

Firstly the assimilated and measured datasets need to be distributed into different quantiles. In this study, 50 linearly-spaced quantiles are defined between the 1st and the 99th quantiles, both included ($q_j = 1, \dots, 99$). Hence, the correction factor for each quantiles is computed as,

$$X^{QM}(q_j) = CDF_{as}^{-1}(q_j) - CDF_{obs}^{-1}(q_j), \tag{4}$$

which is incorporated in the BC method as in Eq. (3), except that the correction factor is applied at each quantile:

$$y^{BC}(q_j) = y^{as}(q_j) + f(X^{QM}(q_j), n). \tag{5}$$

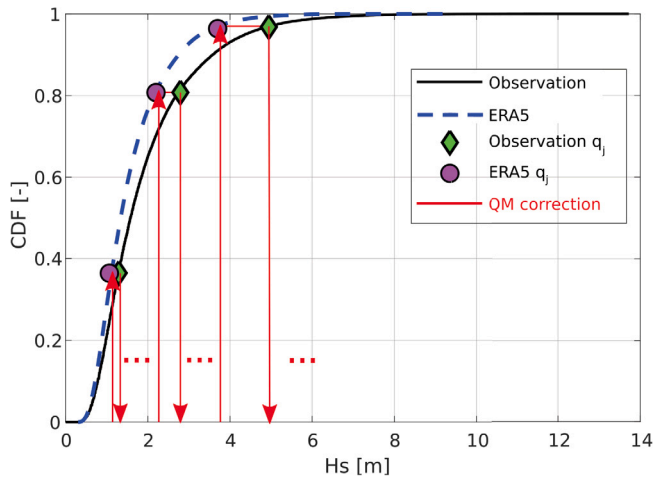


Fig. 1. Schematic representation of the QM technique, where q_j represents each quantile considered in the calibration process.

This technique is expected to outperform the FDM technique, since the bias characteristics at different quantiles are considered. However, similarly to the *Delta* and *FDM* techniques, linearly spacing the quantiles is still expected to lead to underrepresented extreme events.

2.4. Gumbel quantile mapping

The Gumbel quantile mapping (GQM) technique allows for adequately capturing the error in extreme events while preserving the main information around the most common quantiles. The GQM technique, also known as the empirical Gumbel quantile mapping, uses the same quantile mapping technique as in the QM technique but placing the quantiles following a Gumbel distribution function (GDF), which is a particular case of the generalised extreme value distribution, defined as [45],

$$F(x;\mu,\beta) = e^{-e^{-(x-\mu)/\beta}}, \tag{6}$$

where μ and β are, respectively, the location and scale parameters. That way, the GDF provides a better representation of the upper tail of the distribution function, since over 50% of the quantiles are commonly placed above the 99th quantile. Fig. 2 illustrates the differences in the quantile distribution between the QM and GQM BC techniques.

Once the quantiles are identified, the GQM technique is implemented as in Section 2.3, computing the correction term based upon the inverse CDFs at each quantile, as in Eq. (4), and applying that correction term to the assimilated data via a polynomial function, as in Eq. (5). Due to the large number of quantiles placed above the 99th quantile, the correction of the extreme values, where the higher biases are usually found, is expected to improve significantly with respect to the rest of the BC techniques [40].

3. Methodology and case study

The BC techniques introduced in Section 2 are evaluated in order to identify the most effective techniques for the correction of the raw ERA5 data. The characteristics of the resource within *PP* and *Surv* are expected to be significantly different and, thus, the measurement of the effectiveness of BC techniques requires to be adjusted to each mode. To that end, a systematic multi-level methodology is suggested, which provides significantly greater insight to understand the implications of using the different BC techniques. The multi-level assessment consists on three stages, enabling the elimination of ineffective BC techniques at each level. First, traditional statistical metrics are used for an overall evaluation, as shown in Section 3.2.1, which enables the quantification of the bias for the raw ERA5 signal (y^{ERA5}) and the elimination of the simplest and most inefficient techniques. However, these overall metrics mask the effectiveness of the BC techniques for each range of metocean conditions (*e.g.* mild, intermediate, harsh conditions). To that end, as described in Section 3.2.2, metocean conditions are partitioned into different parts, analysing the effect of BC techniques in each quantile and also over the full probability density function (PDF). A third layer is incorporated in order to synthesise the outcomes extracted from the full distribution analysis, using specific statistical metrics described in Section 3.2.3, so that the effectiveness of the different BC techniques can be quantified by means of a single metric. That way, effective BC techniques are selected as the techniques that show a good performance through all the three levels. Finally, the impact of the selected BC techniques on the prediction of the ORE potential is evaluated, including the mean and the variability of the resource. The flowchart depicted in Fig. 3 illustrates the systematic multi-level assessment method that incorporates different statistical metrics and tools. Red arrows illustrate the BC techniques being eliminated in the different stages of the process due to their poor performance, while green lines depict the BC techniques that succeed.

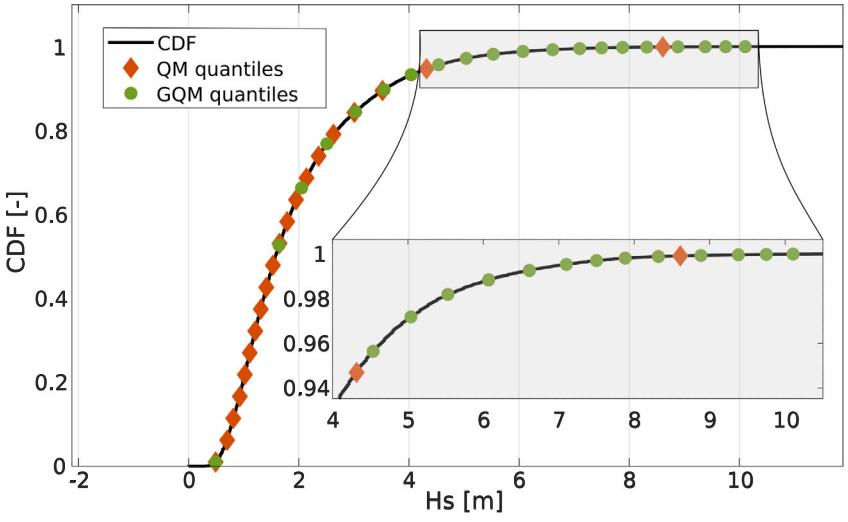


Fig. 2. Comparison of the quantile distribution in the QM and GQM BC techniques.

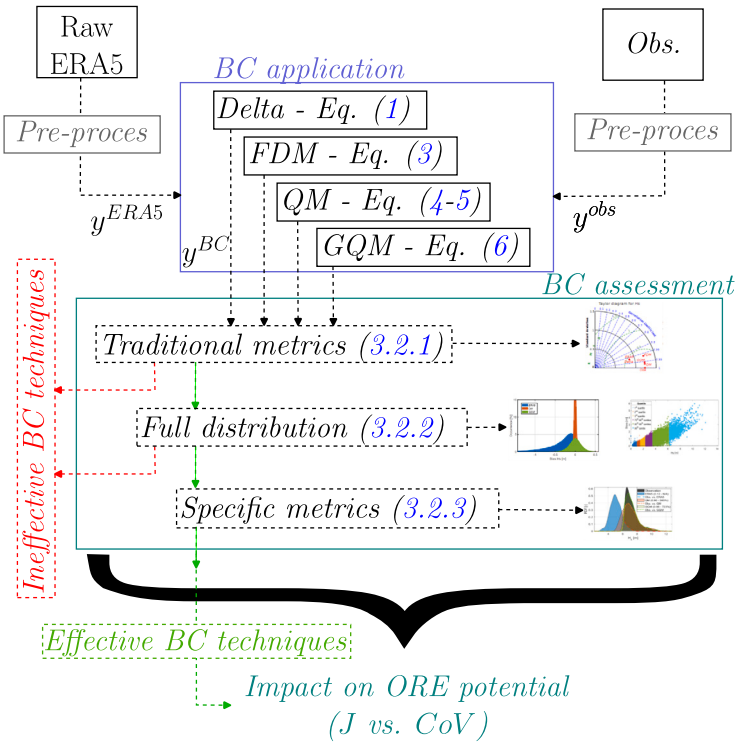


Fig. 3. Flowchart of the systematic multi-level assessment method for BC techniques.

3.1. Data pre-processing

All the re-analysis data is obtained from the last-generation ERA5 product provided by the ECMWF, while the observations data is provided by the Spanish Port Authorities coordination agency *Puertos del Estado*, that owns a large amount of measurement buoys along the Spanish coast. Hence, all the data is openly available for the public.

However, it is important to note that a pre-processing exercise of the datasets is necessary, as illustrated in Fig. 3, so that re-analysis datasets and observations become comparable. Since the correction factors for the different BC techniques are computed based on the difference between the observation and raw assimilation datasets, it is crucial that both datasets have the same length and discretisation. However, data from measurement buoys are commonly irregular and include temporal gaps due to different issues, such as faults in the measuring sensors, communication problems, mooring system failures, or energy shortage [46]. Therefore, the pre-processing of the data consists on two main actions depending on the characteristics of the missing data:

i. If the gap consists on a single sample, the observation data is interpolated using the information of the preceding and subsequent samples;
ii. If the gap extends along several samples (*e.g.* more than one sample), the data corresponding to that gap is neglected.

It should be noted that, in any case, the missing data always corresponds to less than a 20% of the data, ensuring that the particular metocean characteristics of the different locations are preserved in the pre-processed dataset.

3.2. BC assessment techniques

3.2.1. Traditional statistical metrics

The characterisation of the metocean conditions is usually determined by means of standard statistical metrics [14,22]. In this study four traditional metrics are employed:

(I.) The bias is the difference between the ground truth represented by y^{obs} and the different versions of y^{as}, either raw assimilated data or corrected data. In order to provide a single metric for each variable and location analysed in this study, the mean of the absolute bias is considered as,

$$bias(y^{obs}, y^{as}) = mean(abs(y^{obs} - y^{as})). \tag{7}$$

(II.) The Root Mean Square Deviation (RMSD) also represents the deviation between the observations and the different versions of the assimilated datasets (either raw or corrected), but computed as the square root of the average of squared errors:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{y^{obs}}_i - y_i^{as}\right)^2}{N}}, \tag{8}$$

(III.) The standard deviation (σ_y) measures the variability of the dataset, which is calculated for each dataset:

$$\sigma_y = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \mu_y\right)^2}, \tag{9}$$

where y can be either the measured or assimilated dataset, N is the length of the analysed dataset and μ_y represents the mean of all the samples.

(IV.) The correlation, given by the Pearson Correlation (PC) coefficient in this case, indicates the statistical relationship between two datasets, regardless whether that relationship is causal or not:

$$PC = \frac{cov(y^{obs}, y^{as})}{\sigma_{y^{obs}}\sigma_{y^{as}}}, \tag{10}$$

where $cov(y^{obs}, y^{as})$ is the covariance between the measured and assimilated datasets.

In the case of the bias and the RMSD, lower values of the metrics denote a better suitability of the assimilated datasets (0 being the perfect agreement). In contrast, PC shows a better agreement as it is

closer to 1. Finally, the better agreement of the BC technique is shown by σ_y values that are close to the observation σ_y. These four metrics are commonly used when characterising meteorological and metocean conditions. The combination of these metrics enables quantifying the capacity of the BC techniques to correct specific characteristics of the raw assimilated dataset and, thus, provides the relative skill of the BC techniques. For example, PC coefficient quantifies the similarity in pattern, RMSD the level of deviation and σ_y the similarity of the variability.

In addition to the above mentioned metrics, graphical methods are also considered in this study in order to visually depict the effectiveness of the different BC techniques. Taylor diagrams illustrate the RMSD, σ_y and PC of each dataset at once, enabling the comparative assessment of different BC techniques based on these three metrics in a single graph. Based on these traditional metrics, the first evaluation is carried out, excluding those BC techniques that show a poor performance.

3.2.2. Full distribution

However, the comparative assessment provided by Taylor diagrams is based upon average metrics, hindering a more comprehensive analysis comparing the correction of the BC techniques along different quantiles. Therefore, a second level of evaluation is defined for the BC techniques that succeeded the evaluation at the first level, in order to enable a more thorough assessment. Quantile–quantile (Q–Q) plots are commonly used in the comparison of BC techniques, which, in this case, are complemented with PDF comparisons. The latter is used both for the overall comparison of the BC techniques across the whole space and a dissected analysis by partitioning the space into statistically relevant sections. That way, the features of each BC technique, including their weaknesses, can be analysed. The datasets are divided into 6 partitions:

- 1st quartile: low-energetic conditions, around the cut-in limit for power production of ORE technologies.
- 2nd quartile: mild conditions, low power production.
- 3rd quartile: medium-energetic conditions, the conditions for which the ORE technologies are usually designed.
- 75th–90th centiles: medium-high-energetic conditions, highest metocean conditions for power production.
- 90th–99th centiles: high-energetic conditions, around cut-off conditions for power production of ORE technologies.
- +99th centile: extreme conditions, the harshest conditions for survivability.

The dissected analysis is illustrated by means of PDFs for each partition and conditional scatter diagrams.

3.2.3. Specific statistical metrics

In the third level of the assessment approach, the potential improvement of the different BC techniques is evaluated by means of two more specific metrics: The $PDF-score$ and the Distribution Added Value (DAV). The $PDF-score$ is a simple but practical metric of the similarity of two PDFs, allowing the comparison along the entire distribution [47]. Effectively, the $PDF-score$ calculates the common area between two PDFs, providing the minimum cumulative value:

$$PDFscore = \sum_{m=1}^{M} min(PDF(y_m^{obs}) - PDF(y_m^{as})), \tag{11}$$

where M is the number of bins used to represent the PDF.

The DAV metric allows for a normalised comparison between two $PDF-score$s as follows [48],

$$DAV = \frac{PDF-score_{BC} - PDF-score_{as}}{PDF-score_{as}} \times 100, \tag{12}$$

where $PDF-score_{BC}$ and $PDF-score_{as}$ represent the $PDF-score$s of the corrected and raw assimilated datasets, respectively. The DAV is particularly useful for the present study, for it directly represents the improvement achieved by the implementation of each BC technique.

However, the original DAV metric suggested in [48] considers the whole PDF, which, in the present study, prevents discerning the metocean conditions of the *PP* region from the extreme events of the *Surv*. As a consequence, since the extreme events correspond to the $\geq 99th$ quantile of the dataset [40], the original DAV hinders the effectiveness of the BC techniques to represent extreme events. Therefore, an adapted metric is suggested here, where the relevance of the *PP* and *Surv* modes is balanced via evenly distributed weights:

$$DAV_{ORE} = \frac{DAV_{PP} + DAV_S}{2}, \tag{13}$$

where DAV_{PP} and DAV_S are the DAV metrics corresponding to the *PP* and *Surv* modes, respectively. As a first suggestion, the same weight is provided to both modes. In any case, it should be noted that this is a preliminary assumption for this novel metric. Hence, further investigation, outside the scope of this work, could suggest that different weights may be more appropriate, eventually to be tuned according to the specific requirements of each given study.

Finally, the effectiveness of BC techniques selected through the process is evaluated when predicting the resource potential. To that end, the mean wave and wind power density, and the variability of the resource are assessed, analysing the impact on the ORE resource prediction.

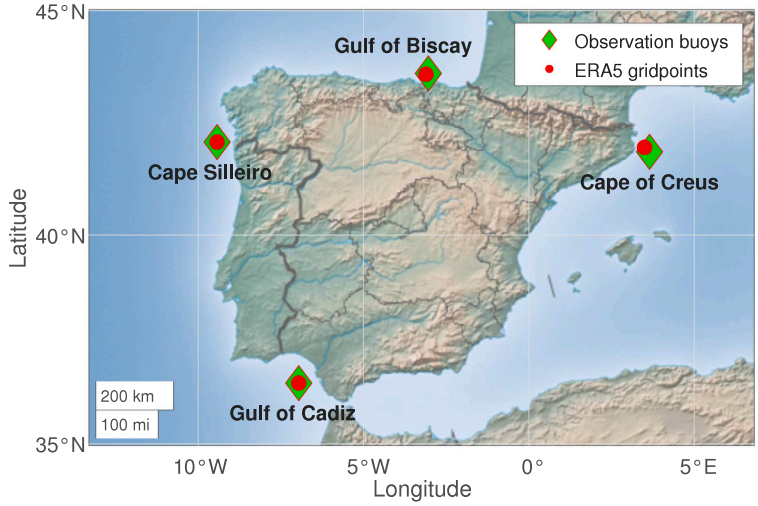
3.3. Metocean data

The performance of the different BC techniques may be influenced by the characteristics of the metocean data used in the analysis. Therefore, in order to evaluate the robustness of the BC techniques, the present study uses four different locations around the Iberian peninsula, as illustrated in Fig. 4(a), where the metocean characteristics are significantly different. The ERA5 dataset for each location is selected from the closest gridpoint to the *in-situ* measurement buoy. The main characteristics of these four locations obtained from ERA5 are depicted in Fig. 4(b), represented via non-parametric PDFs of the three main metocean variables (H_s, T_p and U_w). Each location is described here below:

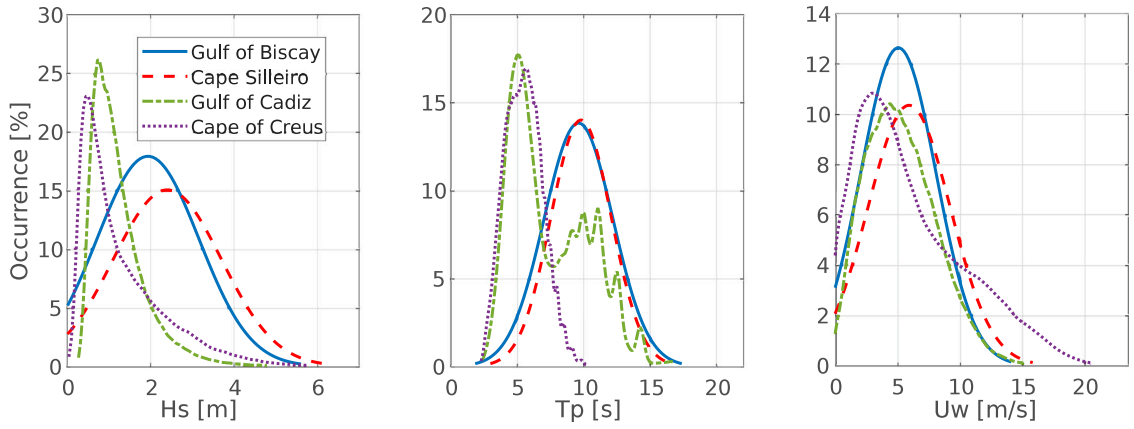
a. *Gulf of Biscay* to the north, a relatively sheltered area in the North-East Atlantic Ocean, clearly dominated by swell waves ($\hat{H}_s$ = 1.9 m and $\hat{T}_p$ = 9.6 s) and consistent wind speeds ($\hat{U_w}$ = 5 m/s) with a relatively low variability.
b. *Cape Silleiro* to the west, an open area facing the open North-East Atlantic Ocean, also dominated by even larger swell waves ($\hat{H}_s$ = 2.4 m and $\hat{T}_p$ = 9.8 s) and consistent wind speeds ($\hat{U_w}$ = 6 m/s) with a relatively low spread.
c. *Gulf of Cadiz* to the south, where the Atlantic Ocean connects to the Mediterranean Sea, resulting in a combination of swells and wind-seas, the latter being predominant ($\hat{H}_s$ = 1.2 m, $\hat{T}_p^{swell}$ = 10 s and $\hat{T}_p^{wind}$ = 5 s approximately). Similarly to the two previous cases, the wind resource in *Gulf of Cadiz* is consistent ($\hat{U_w}$ = 5.3 m/s) and shows a relatively low variability.
d. *Cape of Creus* to the east, in the West of the Mediterranean Sea, clearly dominated by wind-seas ($\hat{H}_s$ = 1.3 m and $\hat{T}_p$ = 5.5 s) and less consistent wind speeds, showing a significant variability with the tail of the PDF extending up to 20 m/s that result in a higher mean wind speed ($\hat{U_w}$ = 6.1 m/s).

4. Performance evaluation

Following the main goal of the present study, the evaluation of the different BC techniques is comprehensively assessed for the *Gulf of Biscay* ia the multi-level assessment approach presented in Section 3. Combining all the different locations, metocean variables and BC techniques would result in an unfeasible large number of combinations for a meaningful and clear presentation in this paper. Therefore, the results at the other three locations are obtained following the same



a) Geographical location of the four locations analysed in this study, including the location of the measurement buoys (large green diamonds) and the ERA5 re-analysis gridpoints (small red circles).



b) PDFs for the four analysed locations: H_s (left), T_p (middle) and U_w (right).

Fig. 4. Definition of the four locations analysed in the present study.

process, but only the final metrics are shown in Appendix of the present paper for the sake of clarity. However, the reader can access the more comprehensive results for all the locations in the supplementary data provided with the paper.

That way, it is ensured that the correction factors are adequately calculated comparing the data for the exact same instants.

4.1. Bias correction at the Gulf of Biscay

The development of ORE technologies requires the understanding of the different metocean conditions. Hence, following the systematic multi-level assessment approach, the evaluation of the different BC techniques is divided into three parts.

4.1.1. Traditional statistical properties

The main statistical metrics defined in Section 3.2.1 provide the overall information on the characteristics of the resource, which simplifies the preliminary comparison of the different BC techniques.

Table 1 shows all the metrics for the different BC techniques, with green highlighting showing the best performance. Darker green denotes the best values, while the lighter green indicates metrics that also

Table 1
Statistical properties of the measured, raw assimilated and the different BC techniques at the Gulf of Biscay for H_s, T_p and U_w. Best performance BC techniques are highlighted in shades of green for each metric.

Dataset	H_s [m]				T_p [s]				U_w [m/s]			
	Bias [m]	$RMSD$ [m]	σ_y [m]	PC [−]	Bias [s]	$RMSD$ [s]	σ_y [m]	PC [−]	Bias [m/s]	$RMSD$ [m/s]	σ_y [m]	PC [−]
Obs.	–	–	1.22	1	–	–	2.59	1	–	–	3.03	1
ERA5	0.38	0.57	0.87	0.96	1.50	2.07	2.91	0.82	1.55	2.06	3.45	0.84
Delta	0.32	0.45	0.87	0.96	1.03	1.70	2.91	0.82	1.39	1.86	3.45	0.84
FDM	0.22	0.32	1.22	0.97	0.91	1.57	2.59	0.82	1.28	1.70	3.03	0.84
QM	0.04	0.11	1.22	0.99	0.10	0.25	2.59	0.99	0.07	0.19	3.03	0.99
GQM	0.12	0.15	1.22	0.99	0.48	0.95	2.59	0.93	0.65	0.95	3.03	0.95

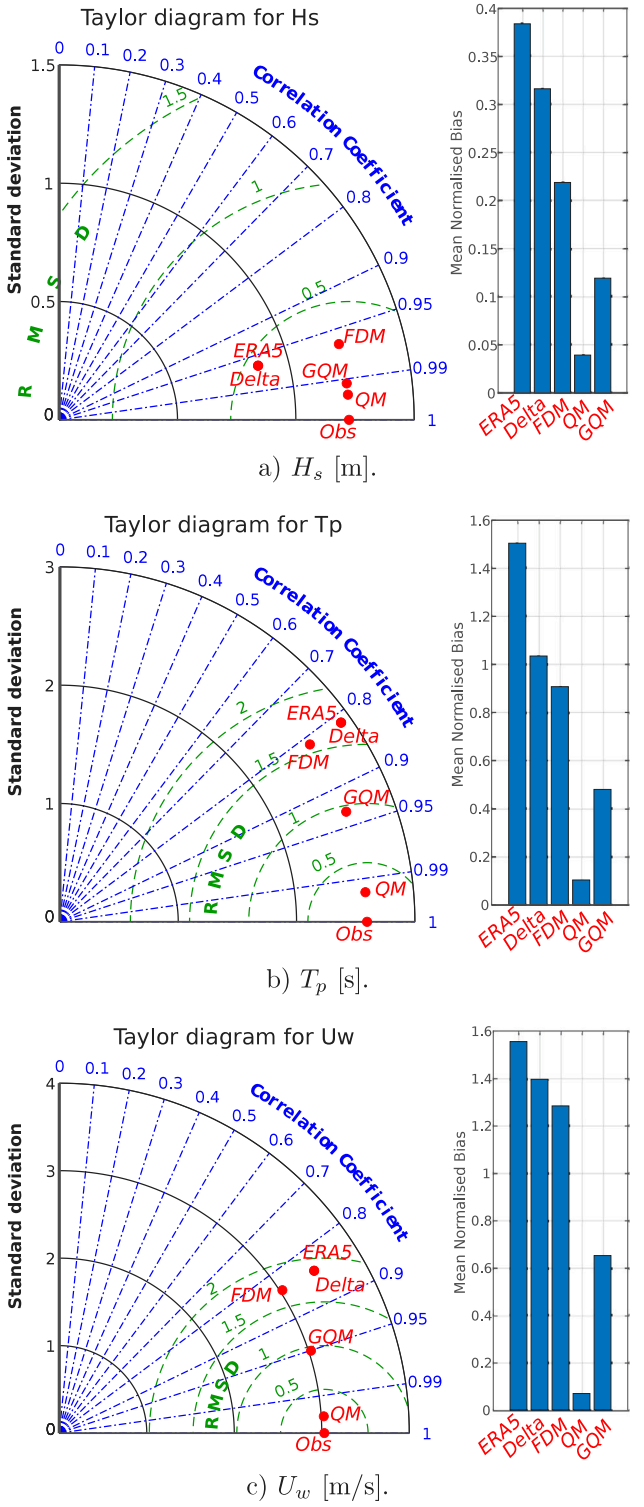


Fig. 5. Taylor diagrams and bias of the different BC techniques: (a) H_s, (b) T_p and (c) U_w.

show a good agreement. The first outcome from the table is that the raw ERA5 dataset provides a poor representation of the real resource, regardless of the metocean variable, as particularly demonstrated with the *Bias* and *RMSD* metrics. Another clear aspect that can be extracted from Table 1 is that the *Delta* method does not provide any significant improvement, while all the distribution mapping methods demonstrate to provide better results. Note that the same PC metric is obtained with the three distribution mapping methods, achieving the very same value as for the observations. Among the distribution mapping methods, it is clear that the QM method is the most effective for all metocean variables, the GQM achieving similar performance metrics in the case of H_s, but significantly poorer with T_p and U_w. The comparison between the different BC techniques and the superiority of the *QM* method is clearly shown in Fig. 5 by means of the Taylor diagrams and the bar plot representation of the bias.

Furthermore, the results presented in Table 1 and Fig. 5 show that all BC techniques perform significantly better for H_s than T_p or U_w. The main reason for that is believed to be the lower variability of H_s, as shown by the σ_y metric, which is twice as high for T_p and almost three times higher for U_w. In addition, T_p is defined as a discrete statistical parameter estimated from the sea-state spectrum, meaning that small variations of the spectrum shape can lead to a different bin of the T_p vector. Since all BC techniques are pure statistical techniques and the same number of quantiles is used for all the variables, the higher variability of T_p and U_w expands the bounds of the correction space, reducing the effectiveness of the correction factors that are computed at each quantile. However, the *QM* method shows a remarkable performance, even with T_p and U_w, reducing the bias of the raw assimilated dataset by an order of magnitude. Overall, it can be concluded that the performance of the *Delta* and *FDM* methods is unsatisfactory, although both improve the quality of the raw dataset, as concluded in [24]; while the *GQM* technique looks promising, especially because it is designed to focus on the extreme values. Therefore, the *Delta* and *FDM* techniques are discarded for the following stages.

4.1.2. Analysis of the full distribution

The analysis via the baseline statistical metrics reduces the set of potential BC techniques to two: *QM* and *GQM* techniques. Both are assessed through the overall and dissected analysis across the six partitions.

Overall analysis. The simplest and most consistent full distribution analysis is the visual comparison of the CDFs, as illustrated in Fig. 6. Overall, apart from the deviation of the raw *ERA5* dataset, *QM* and *GQM* methods seem to have captured the characteristics of the observation dataset perfectly. Zooming in the upper quantiles (96th–100th), the agreement between the observation, and the *QM* and *GQM* techniques is still remarkable. It is only in the very extreme conditions (99.995th–100th quantiles) when the *QM* method seems to diverge slightly, while the *GQM* method still overlays the observation dataset, as expected. Similarly, the CDF results for T_p and U_w also show almost imperceptible differences.

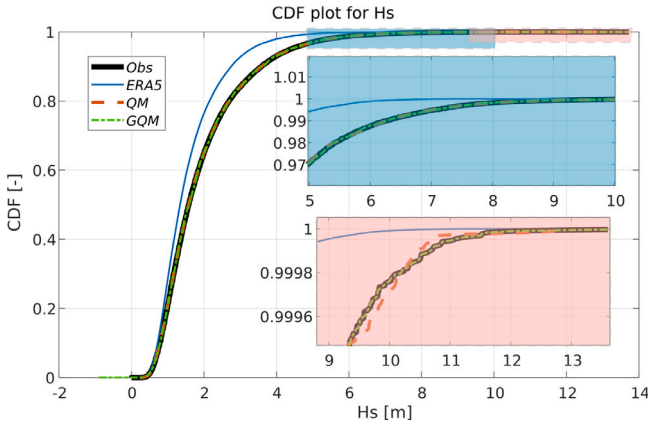


Fig. 6. CDF for H_s: Observations (black), raw *ERA5* (blue), *QM* (red) and *GQM* (green) techniques. Blue shaded area zooms in the last 4 percentiles, while the red shaded area zooms further up to the 99.995 percentile.

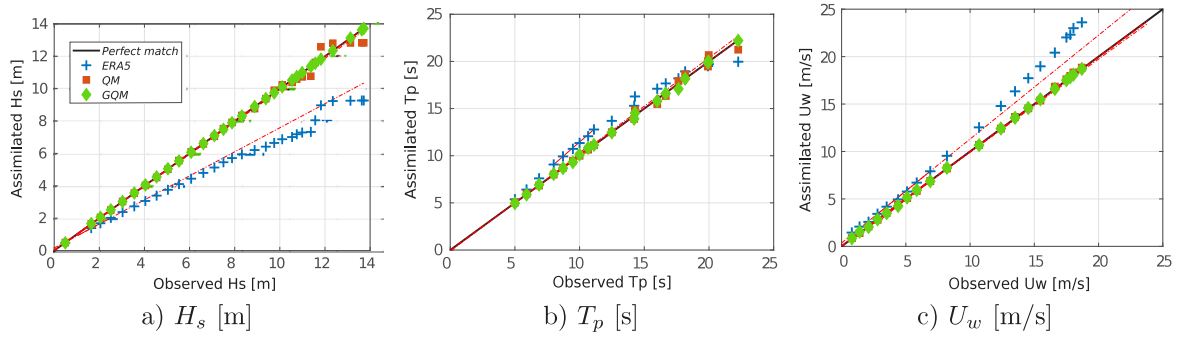


Fig. 7. Q–Q plots of the different BC techniques: (a) H_s, (b) T_p and (c) U_w.

In order to represent the differences more clearly, a comparative analysis is carried out in different quantiles, for which Q–Q plots are commonly used, comparing the corrected datasets to the observations. In this case, the comparison is carried out for the three metocean variables H_s, T_p and U_w, as illustrated in Fig. 7(a), (b) and (c), respectively. Yet, similar results to those obtained in Fig. 6 are shown in Fig. 7, where differences between the *QM* and *GQM* BC techniques seem negligible.

However, these results seem to contradict the results shown in Table 1 and Fig. 5, where the *QM* technique is shown to be the technique that provides the most accurate data. The main reason for such a contradiction is that, in Figs. 6 and 7, the performance of the BC techniques at each quantile is given by a single averaged number, which hinders the dispersion within each quantile. The only manner to analyse this dispersion is going beyond the average values and analysing the probability distributions. Fig. 8(a), (b) and (c) illustrate, respectively, the overall PDFs for H_s, T_p and U_w.

The results shown in Fig. 8 are in line with the metrics shown in Table 1. The outstanding performance of the *QM* BC technique is clearly shown, with the vast majority of the bias distribution being concentrated very close to 0 for the three metocean variables. Likewise in Table 1, the bias distribution of the *GQM* method also shows a significant improvement with respect to the raw *ERA5* dataset, especially for H_s.

Analysis of the partitions. Nevertheless, the distributions depicted in Fig. 8 correspond to the whole dataset, where mild metocean conditions (*i.e.* $H_s \leq 2$ m, $T_p \leq 12$ s and $U_w \leq 20$ m/s) are largely predominant. However, due to the characteristics of the PDFs assumed in each BC technique, the performance of these BC techniques is expected to vary significantly across the different conditions.

Similarly to Fig. 8, bias distributions for each BC technique and partition are illustrated in Fig. 9. Here, the analysis is restricted to H_s in order to avoid repetitiveness, but the same conclusions can be applied to T_p and U_w. To some extent, these graphs are the equivalents of the Q–Q plots shown in Fig. 7, but provide greater insight into the effectiveness of each BC technique. Between low-energetic and medium-high energetic conditions, which corresponds to the 90% of the data, the *QM* method is clearly shown to outperform the *GQM* method. The latter exhibits difficulties to adequately correct low-energetic and mild conditions, which is to be expected given that only 1 quantile is placed in each of these partitions (see Fig. 2). However, it should be noted that, since the raw values in these partitioning are low, the bias is also modest within these partitions. Yet, the effectiveness of the *GQM* technique increases substantially as it approaches to the high-energetic conditions, as is to be expected. In fact, the *GQM* technique demonstrates to outperform the *QM* method, especially in the extreme partition, where the *QM* struggles to accurately represent the realistic conditions.

The conditional scatter diagrams in Fig. 10 show the trends of the biases with respect to the metocean conditions for each assimilated dataset, where the colour code illustrates the different partitions. In the case of the raw *ERA5* dataset, the trend is close to linear, with the bias increasing as H_s increases. This trend is overturned when implementing the *GQM* technique, showing particularly low bias values at high-energetic and extreme H_s conditions. In contrast, the *QM* method is able to vanish the bias between low- and medium-high-energetic conditions, but increases drastically at high- and extreme conditions (*i.e.* $H_s \geq 5$ m). Note that the conditional scatter diagrams for both corrected datasets show a sawtooth pattern. The lower bias values of these sawtooth pattern correspond to the H_s values where the quantiles are placed when identifying the correction factor. The bias slightly increases for H_s values between the quantiles, which is particularly evident in the case of the *GQM* method (7 quantiles are placed for $H_s \leq 4$ m, see Fig. 2, and 7 troughs can be identified in that range).

Similar trends can also be identified for T_p and U_w, as illustrated in Fig. 11, which demonstrates the applicability of these methods to

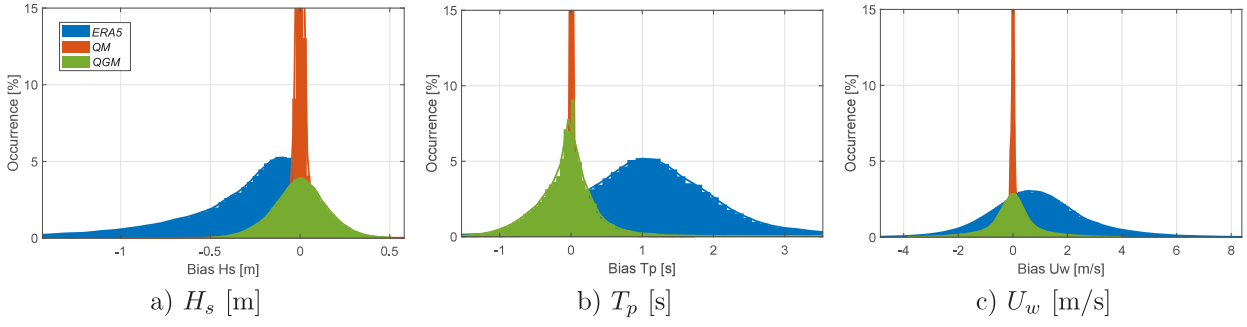


Fig. 8. PDFs of the different BC techniques: (a) H_s, (b) T_p and (c) U_w.

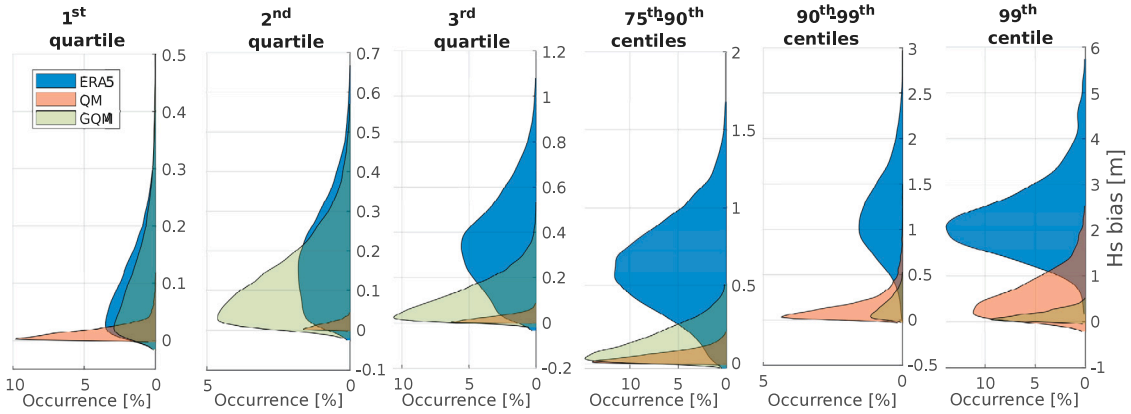


Fig. 9. PDF across the different partitions, comparing the bias of the raw assimilation, and the *QM* and *GQM* BC techniques.

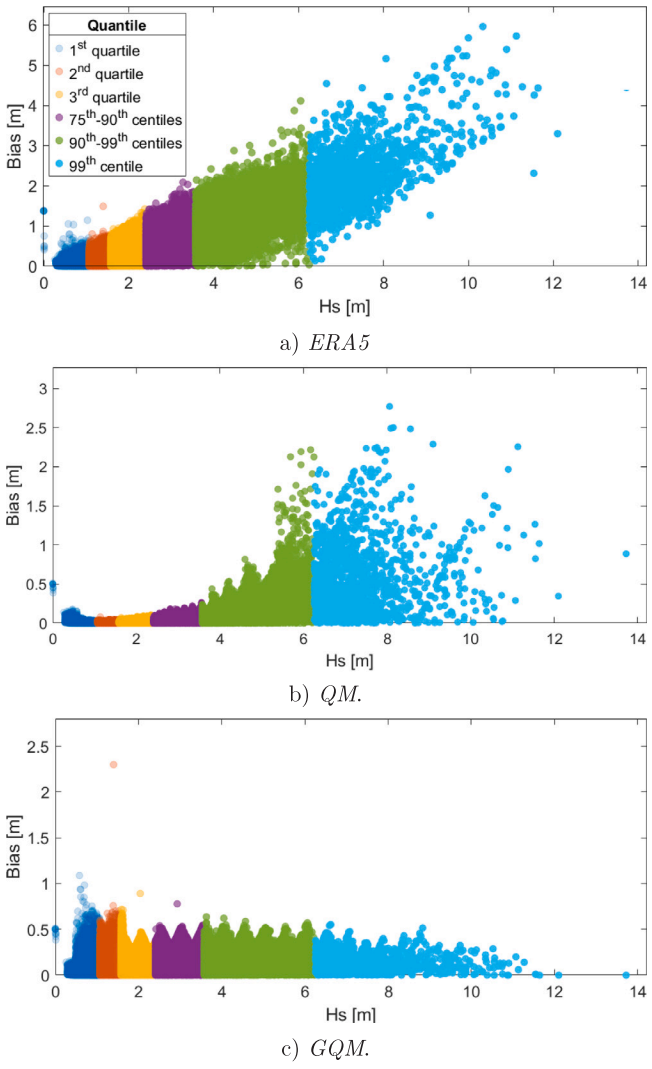


Fig. 10. Conditional scatter diagrams for each BC technique applied to H_s: (a) *ERA5*, (b) *QM* and (c) *GQM*.

Table 2
Evaluation of the BC techniques: Overall assessment for H_s and U_w.

H_s	Overall		Extremes		Mean
	$PDF\ score$	DAV_{PP}	$PDF\ score$	DAV_S	DAV_{ORE}
ERA5	0.87	–	0.12	–	–
QM	0.99	13%	0.8	585%	299%
GQM	0.97	11%	0.95	723%	367%
U_w	$PDF\ score$	DAV	$PDF\ score$	DAV	DAV_{ORE}
ERA5	0.91	–	0.47	–	–
QM	0.99	9%	0.79	69%	38.5%
GQM	0.96	5%	0.95	102%	53.5%

any metocean variable. However, one can observe that the absolute biases corresponding to T_p and U_w are considerably larger compared to H_s. In any case, both the *QM* and the *GQM* techniques demonstrate their efficacy in substantially reducing the bias, showing that the *QM* is particularly adequate for the *PP* region, while the *GQM* demonstrates a better performance in the *Surv* region. Therefore, both methods are selected to be evaluated in the third level by means of the specific metrics.

4.1.3. Specific bias correction metrics

The more specific $PDF-score$ and DAV metrics also yield to similar conclusions. Table 2 shows the improvement of both the QM and GQM BC techniques, reaching $PDF-scores$ of 0.99 in the case of the QM and above 0.96 for GQM. However, the improvement with respect to the raw *ERA5* data is above 10% for H_s and slightly lower (between 5 and 9%) for U_w.

In contrast, the same metrics computed for extreme conditions result in a significantly different scenario. First, the raw *ERA5* dataset is again shown to provide a very poor representation of the extremes (with a $PDF-score$ of 0.12 and 0.47 for H_s and U_w, respectively), which confirms the results also shown in Figs. 9–11. Secondly, although both BC techniques show a remarkable improvement of the raw *ERA5* dataset, the GQM clearly shows to overperform QM, achieving an improvement of over 720% for H_s and 100% for U_w.

Fig. 12 illustrates the poor performance of the raw *ERA5* dataset in extreme conditions, showing substantially shifted PDFs with respect to the observations both for H_s and U_w. Although the QM technique shows an important correction, especially in the case of the U_w, the GQM technique almost overlaps the observation PDF, demonstrating again its suitability for extreme conditions.

Differences between the QM and GQM BC techniques are clearly demonstrated by the different sets of results presented in this section. Connecting with the two operational regions mentioned in Section 1,

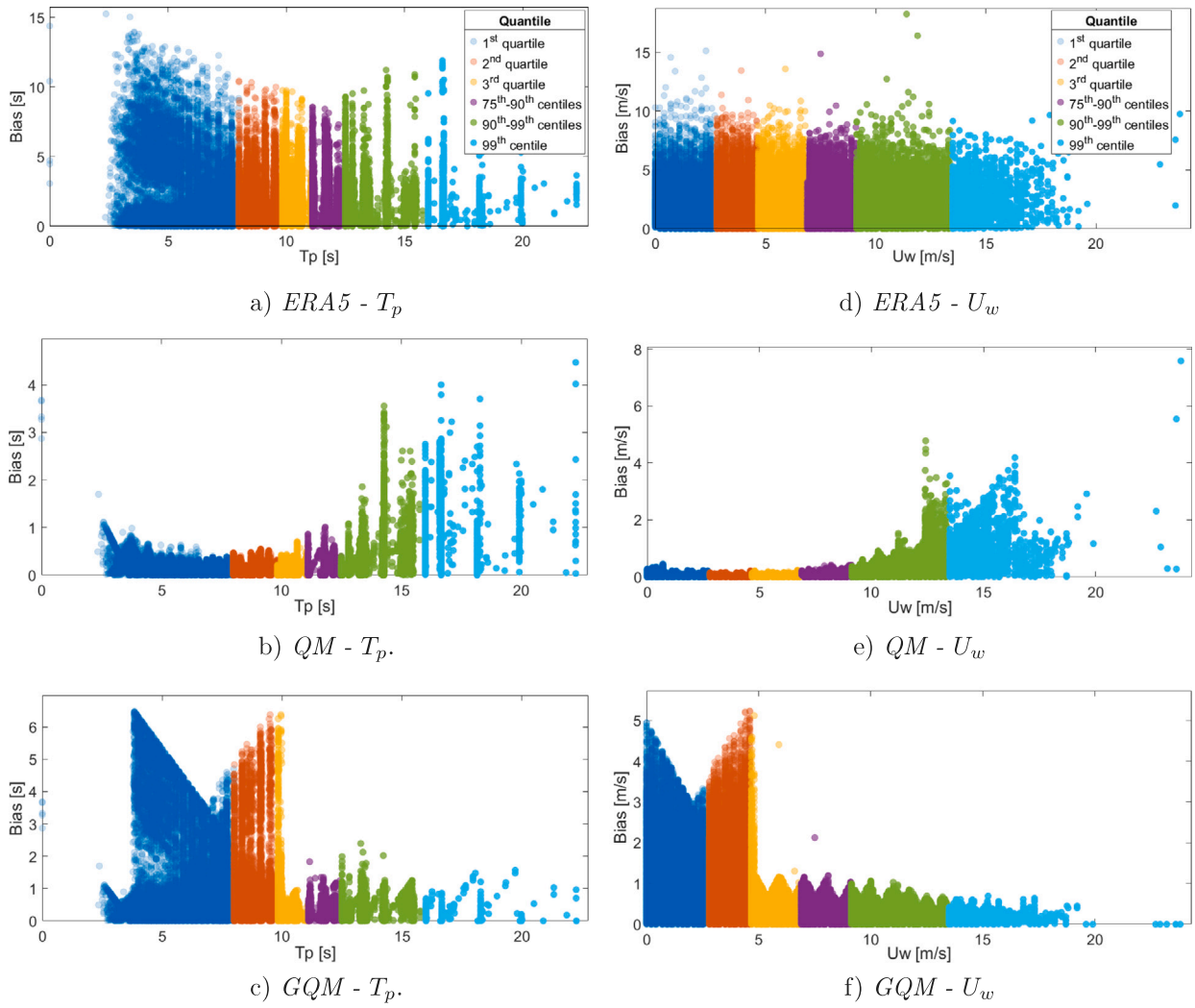


Fig. 11. Conditional bias for each partitioning and BC technique: T_p (a–c) and U_w (d–f).

the QM technique is shown to be the most appropriate for the studies focused on the power production regions, while the GQM technique is demonstrated to be more suitable for survivability studies. Since the number of extreme events is scarce, they appear underrepresented when considering the whole dataset, leading to an arguably better suitability of the QM technique for overall studies. Yet, survivability is a key aspect on the design of any ORE technology, meaning that extreme events are highly relevant despite their lower frequency. Therefore, statistical metrics for the evaluation of the most suitable BC technique may need to be revised.

The revision suggested in this study is presented in Eq. (13), where the same relevance is given to the correction of the ORE resource data for *PP* and *Surv* studies. Hence, Table 2 presents the results obtained with the revised metric, being the GQM metric the most suitable, both for H_s and U_w.

4.2. Impact on ORE potential analysis

One of the main objectives of analysing the ORE resource is the assessment of the power density or potential of the different ORE technologies. In this sense, two main variables are commonly studied: the wave and wind power density (J_{wave} and J_{wind}, respectively) and the Coefficient of Variation (CoV), which illustrate the average potential and variability of the resource in a given location, respectively. The three ORE potential variables are defined as follows:

$$J_{wave} = 0.49\alpha H_s^2 T_p, \tag{14}$$

where $\alpha = 0.9$ is [22],

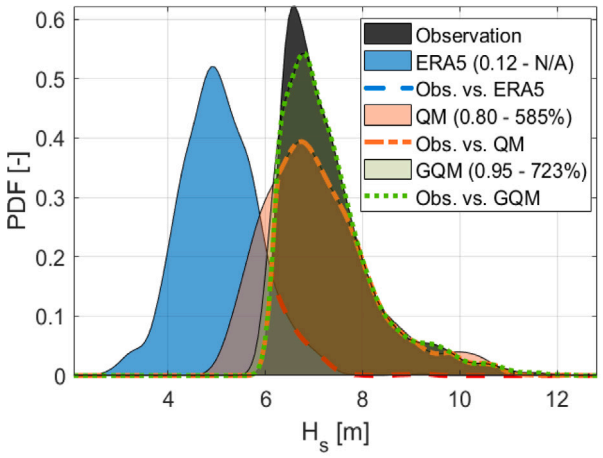
$$J_{wind} = 0.5\rho_{air} U_w^3, \tag{15}$$

ρ_{air} is the air density, and

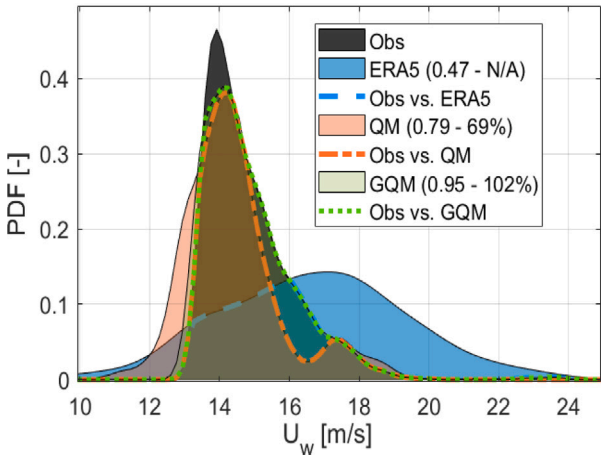
$$CoV = \frac{\sigma(J)}{\hat{J}}, \tag{16}$$

$\sigma(J)$ and $\hat{J}$ representing the standard deviation and the average power density, respectively.

Although the different BC techniques are only applied to the main ORE resource variables, the impact of these techniques on the ORE potential evaluation is crucial. Fig. 13 (left) and (right) illustrate the ORE potential evaluated by means of different BC techniques for wave and wind energy, respectively, where two main outcomes can be extracted. On the one hand, the raw $ERA5$ dataset is once more demonstrated to be inadequate also for resource potential evaluation, misrepresenting the real wave and wind energy resource by over 30% and 50%, respectively. On the other hand, no substantial difference is observed between QM and GQM for the estimation of the average potential, but the GQM approach seems to be slightly more accurate for the correction of the resource variability.



a) H_s - Extreme distribution



b) U_w - Extreme distribution

Fig. 12. PDF-score and DAV computation for extreme conditions (a) H_s and (b) U_w.

In a sense, these results support the idea suggested with the revised DAV_{ORE} metric and shown in Table 2, claiming that the relevance of the extreme events turns out to be hindered by the overall analyses via traditional statistical metrics. However, it should be noted that this conclusion is not yet definitive and requires a more thorough study including the most critical ORE technology design aspects beyond pure resource datasets, before it can be confirmed.

4.3. Sensitivity of BC techniques' performance to resource characteristics

Finally, the sensitivity of the performance of these BC techniques is studied in this section, including the rest of the locations described in Section 3 and characterised in Fig. 4. It is crucial for a BC technique to be consistent across different resource conditions, since the resource assessment is commonly carried out across a wide area with significantly different resource characteristics.

For the sake of clarity, all the results for the additional three locations are shown in Appendix. Table A.1 presents all the traditional metrics for the four locations analysed, for which identical patterns are identified, as highlighted by the green cells. In all the locations the QM technique appears to be the most suitable BC technique, with the GQM approach showing a very good agreement with the observation datasets. It is only in the case of the Gulf of Cadiz and H_s variable, that the $Delta$ and FDM seem also to perform well. However, the raw $ERA5$ dataset at the Gulf of Cadiz shows a relatively low bias and the variability of the resource (represented by σ_y) is also significantly lower than at the rest of locations.

Similarly, the variability σ_y is shown the lowest in the case of H_s, with U_w showing the largest. Accordingly, the calibration of the resource variables via BC techniques is particularly good for H_s. However, the QM is shown to perform remarkably well with all the different variables and locations.

The performance of the QM and the GQM techniques for different partitions is also illustrated in Figs. A.1–A.3 for Cabo Silleiro, Gulf of Cadiz and Cape of Creus, respectively. The biases obtained from the different BC techniques in the different locations show identical patterns to those shown in Figs. 10 and 11 for the Gulf of Biscay, including the sawtooth structure. Hence, it is demonstrated that the sensitivity of the BC techniques to the resource characteristics of specific locations is negligible.

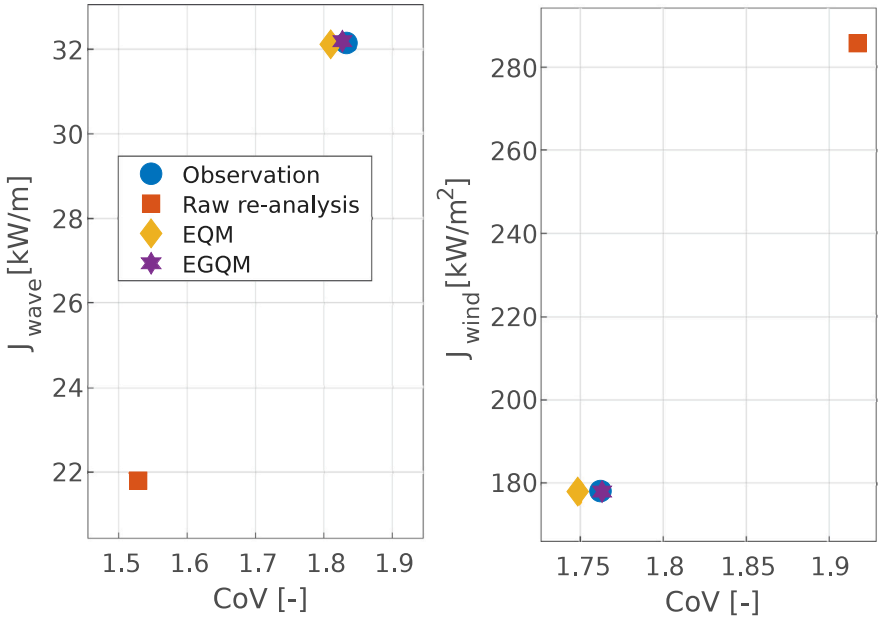


Fig. 13. Evaluation of the ORE potential via J and CoV: (left) wave and (right) offshore wind energy.

In order to facilitate reproducibility and data reuse, all the datasets, including observations, raw ERA5 and corrected datasets using the different BC techniques, are available as research data in the following repository: https://github.com/MGEP-Fluidos/BC4ORE/tree/main/RENEpaperData.

5. Conclusions

Th paper presents a thorough evaluation of different bias correction (BC) techniques applied to the reduction of uncertainty on long-term metocean datasets required for the design of any Offshore Renewable Energy (ORE) technology. To that end, four different BC techniques are analysed by means of a systematic multi-level assessment method: *i.e.* the *Delta* approach and three different distribution-mapping techniques based on the full distribution mapping FDM, quantile-mapping QM and Gumble quantile-mapping GQM. The assessment methodology includes a set of traditional statistical metrics at the first stage, where the simplest and ineffective methods are excluded. In the following stage of the methodology, a full-distribution analysis is carried out, evaluating the performance of the BC techniques across the whole space. In the third and last stage, additional metrics specifically adapted for the ORE sector are employed in order to select the most suitable BC techniques. This methodology considers the particularity of ORE technologies, which are designed to efficiently generate energy under the power production conditions and survive the extreme survivability conditions.

In addition, the methodology is applied at four locations with significantly different metocean characteristics around the Iberian Peninsula in order to verify the consistency of the BC techniques: the Gulf of Biscay, Cape Silleiro, Gulf of Cadiz and Cape of Creus. The metocean data for all the locations is retrieved from the ERA5 re-analysis dataset provided by the European Centre for Medium-Range Weather Forecasts, considering the main variables that characterise the ORE resource, *i.e.* the significant wave height H_s, the peak period T_p and the wind speed U_w are considered. Additionally, *in-situ* measurements provided by the Spanish Port Authorities coordination agency *Puertos del Estado* are used as the *ground truth* to determine the correction factors of the different BC techniques.

Regardless of the location, the raw ERA5 datasets show a relatively poor agreement with the observations, being particularly poor for the extreme events appeared under survivability conditions. Therefore, as a first conclusion, the need for calibration of the raw ERA5 datasets is confirmed. In this sense, all the BC techniques suggested in this paper demonstrate to improve the quality of the raw dataset. However, the improvement of the *Delta* and the FDM techniques is shown to be insufficient for a significant uncertainty reduction. Therefore, these two BC techniques are discarded in the first stage of the methodology, focusing on the more effective QM and GQM methods through the following two stages. The more traditional QM technique is demonstrated to be the most effective technique for the datasets focused on the power production conditions. In contrast, its performance decreases considerably beyond the power production region, increasing the bias substantially when applied to extreme events. The GQM technique largely outperforms the QM under extreme condition within the survivability region. However, the statistical metrics evaluating the overall performance of the BC techniques are biased by the predominance of the conditions within the power production region, hindering the great relevance of the extreme events despite their lower occurrence.

In order to overcome this limitation of the traditional metrics, alternative metrics adapted to better represent the needs of ORE systems are considered in the third stage. These alternative metrics provide the same relevance to the calibration of the datasets within the power production and survivability regions. Hence, it is shown that the GQM technique may be more suitable for overall assessments that consider metocean data from both regions simultaneously. This conclusion is also supported by the analysis of the wave and wind energy potential in the different locations. The GQM and QM show similar capabilities for the estimation of the average power density (J), but the GQM is shown to better capture the variability, providing a better agreement for the coefficient of variation (CoV) when comparing to observation datasets. However, although out of the scope of the current study, it should be noted that the impact of corrected ORE resource variables should be evaluated on key design parameters, not only pure resource parameters as J and CoV.

Accordingly, potential future research lines should incorporate:

i Evaluating different probability density functions
ii Other more complex bias correction techniques that allow the consideration of wave/wind direction, *e.g.* the directional adjusted Gumbel quantile mapping method,
iii Assessing the impact of bias correction on the ORE technologies' design process using specific ORE technology design parameters,
iv Assessing the spatial distribution of the impact of bias correction, identifying the locations where bias correction is crucial.

CRediT authorship contribution statement

Markel Penalba: Conceptualization, Methodology, Software, Visualization, Data curation, Writing – original draft. **Chao Guo:** Software, Visualization. **Ander Zarketa-Astigarraga:** Methodology, Data curation, Writing – review & editing. **Giulia Cervelli:** Visualization, Writing – review & editing. **Giuseppe Giorgi:** Conceptualization, Writing – review & editing. **Bryson Robertson:** Conceptualization, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Markel Penalba reports financial support was provided by Mondragon University Higher Polytechnic School. Markel Penalba reports a relationship with Basque Foundation for Science that includes: funding grants. Markel Penalba reports a relationship with Spanish Scientific research Council that includes: funding grants. Markel Penalba reports a relationship with Basque Government Department of Environment and Territorial Policy that includes: funding grants.

Acknowledgements

This publication is part of the research project PID2021-124245OA-I00 funded by MCIN/AEI/10.13039/501100011033 and by ERDF Away of making Europe (Spain), and the research project funded by the Basque Government's ELKARTEK 2022 program under the grant No. KK–2022/00090 (Spain). In addition, the present work was partially carried out during Dr. Penalba's research visit at Oregon State University under Dr. Robertson's supervision, which was funded by the Basque Government's IKERMUGIKORTASUNA research mobility program under the grant No. MV-2022-1-0017 (Spain). Finally, the authors from the Fluid Mechanics research group at Mondragon University are also supported by the Basque Government's Research Group Program under the grant No. IT1505-22 (Spain).

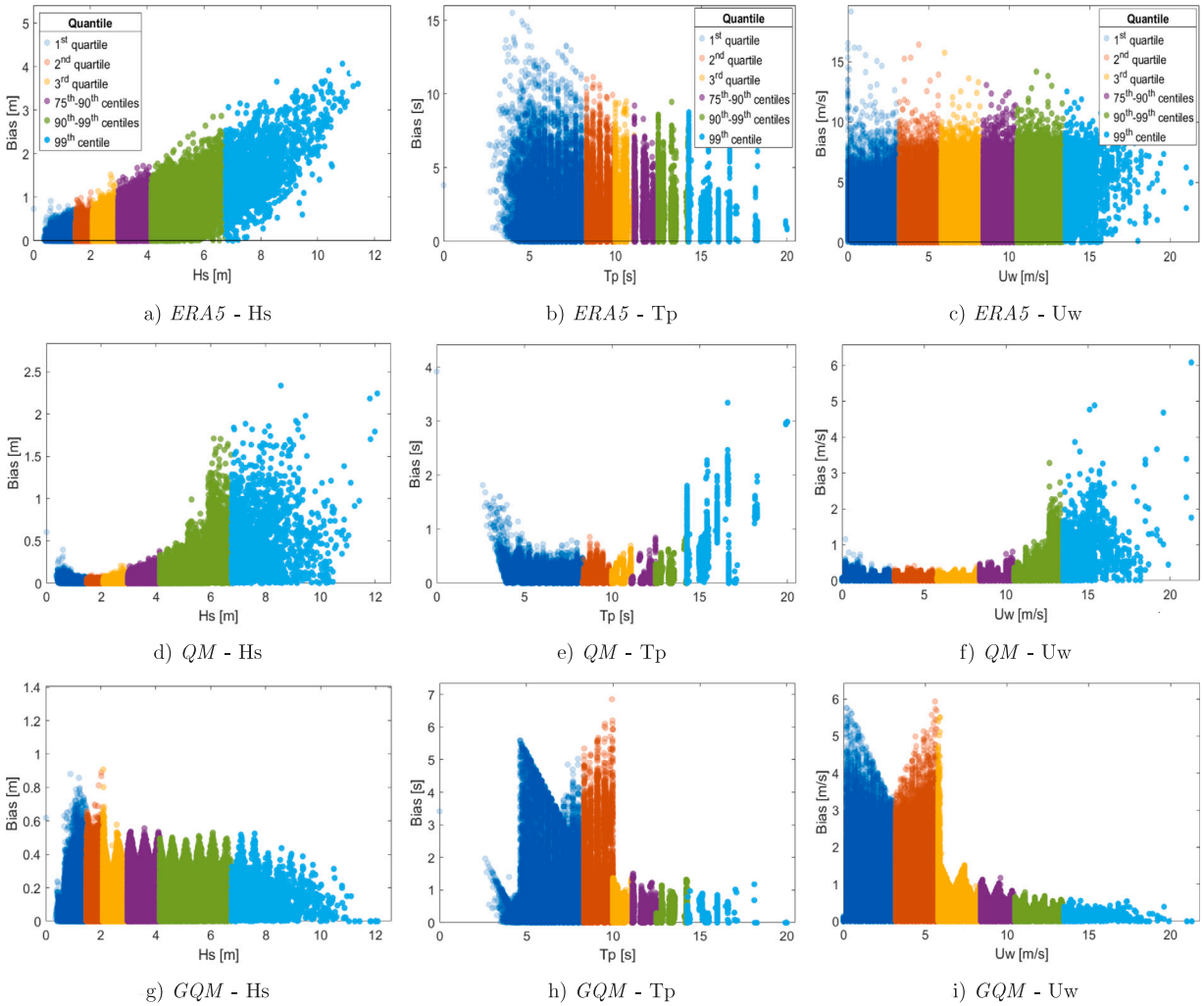
Appendix. Bias correction results in additional locations

Statistical properties of the different BC techniques in different locations are shown in Table A.1, comparing the results obtained at the Gulf of Biscay with the results computed at Cape Silleiro, Gulf of Cadiz and Cape of Creus.

In addition, partitioned scatter diagrams for Cape Silleiro, Gulf of Cadiz and Cape of Creus are illustrated in Figs. A.1–A.3, respectively.

Table A.1
Statistical properties of the measured, raw assimilated and the different BC techniques for H_s, T_p and U_w in all the four analysed locations.

Dataset		H_s [m]				T_p [s]				U_w [m/s]			
		Bias [m]	$RMSD$ [m]	σ_y [m]	PC [–]	Bias [s]	$RMSD$ [s]	σ_y [m]	PC [–]	Bias [m/s]	$RMSD$ [m/s]	σ_y [m]	PC [–]
Gulf of Biscay	*Obs.*	–	–	1.22	1	–	–	2.60	1	–	–	3.03	1
	ERA5	0.38	0.57	0.87	0.96	1.50	2.07	2.91	0.82	1.55	2.06	3.45	0.84
	Delta	0.32	0.45	0.87	0.96	1.04	1.70	2.91	0.82	1.40	1.86	3.45	0.84
	FDM	0.22	0.32	1.22	0.97	0.91	1.57	2.59	0.82	1.28	1.70	3.03	0.84
	QM	0.04	0.11	1.22	1.00	0.10	0.25	2.59	1.00	0.07	0.19	3.03	0.99
	GQM	0.12	0.15	1.22	0.99	0.48	0.95	2.59	0.93	0.65	0.95	3.03	0.95
Cape Silleiro	*Obs.*	–	–	1.30	1	–	–	2.28	1	–	–	3.31	1
	ERA5	0.29	0.43	1.04	0.97	1.47	1.90	2.69	0.83	2.59	3.16	4..9	0.89
	Delta	0.26	0.43	1.04	0.97	0.96	1.50	2.69	0.83	1.65	2.09	4.39	0.89
	FDM	0.21	0.29	1.30	0.97	0.82	1.34	2.27	0.83	1.20	1.56	3.31	0.89
	QM	0.04	0.10	1.30	1.00	0.07	0.14	2.28	1.00	0.08	0.16	3.31	1.00
	GQM	0.13	0.16	1.30	0.99	0.42	0.79	2.28	0.94	0.65	0.95	3.31	0.96
Gulf of Cadiz	*Obs.*	–	–	0.67	1	–	–	2.96	1	–	–	2.82	1
	ERA5	0.17	0.23	0.63	0.94	2.40	3.50	3.40	0.64	2.19	2.85	3.68	0.82
	Delta	0.17	0.23	0.63	0.94	2.07	2.72	3.40	0.64	1.61	2.12	3.68	0.82
	FDM	0.17	0.23	0.67	0.94	1.86	2.50	2.94	0.64	1.28	1.70	2.82	0.82
	QM	0.02	0.12	0.67	1.00	0.09	0.17	2.96	1.00	0.07	0.17	2.82	1.00
	GQM	0.09	0.12	0.67	0.98	0.68	0.94	2.96	0.95	0.65	0.93	2.82	0.95
Cape of Creus	*Obs.*	–	–	1.02	1	–	–	1.42	1	–	–	4.40	1
	ERA5	0.32	0.43	0.82	0.95	0.81	1.21	1.47	0.71	2.81	3.59	5.05	0.83
	Delta	0.26	0.34	0.82	0.95	0.74	1.10	1.47	0.71	2.19	2.82	5.05	0.83
	FDM	0.22	0.30	1.02	0.96	0.74	1.09	1.41	0.71	1.93	2.51	4.40	0.84
	QM	0.03	0.15	1.02	1.00	0.06	0.13	1.41	1.00	0.11	0.22	4.40	1.00
	GQM	0.12	0.15	1.02	0.99	0.32	0.51	1.41	0.94	0.92	1.27	4.40	0.96



a) *ERA5* - Hs b) *ERA5* - Tp c) *ERA5* - Uw

d) *QM* - Hs e) *QM* - Tp f) *QM* - Uw

g) *GQM* - Hs h) *GQM* - Tp i) *GQM* - Uw

Fig. A.1. Conditional bias for each partitioning and BC technique including H_s, T_p and U_w for ERA5 (a–c), *QM* (d–f) and *GQM* (g–i) at CS.

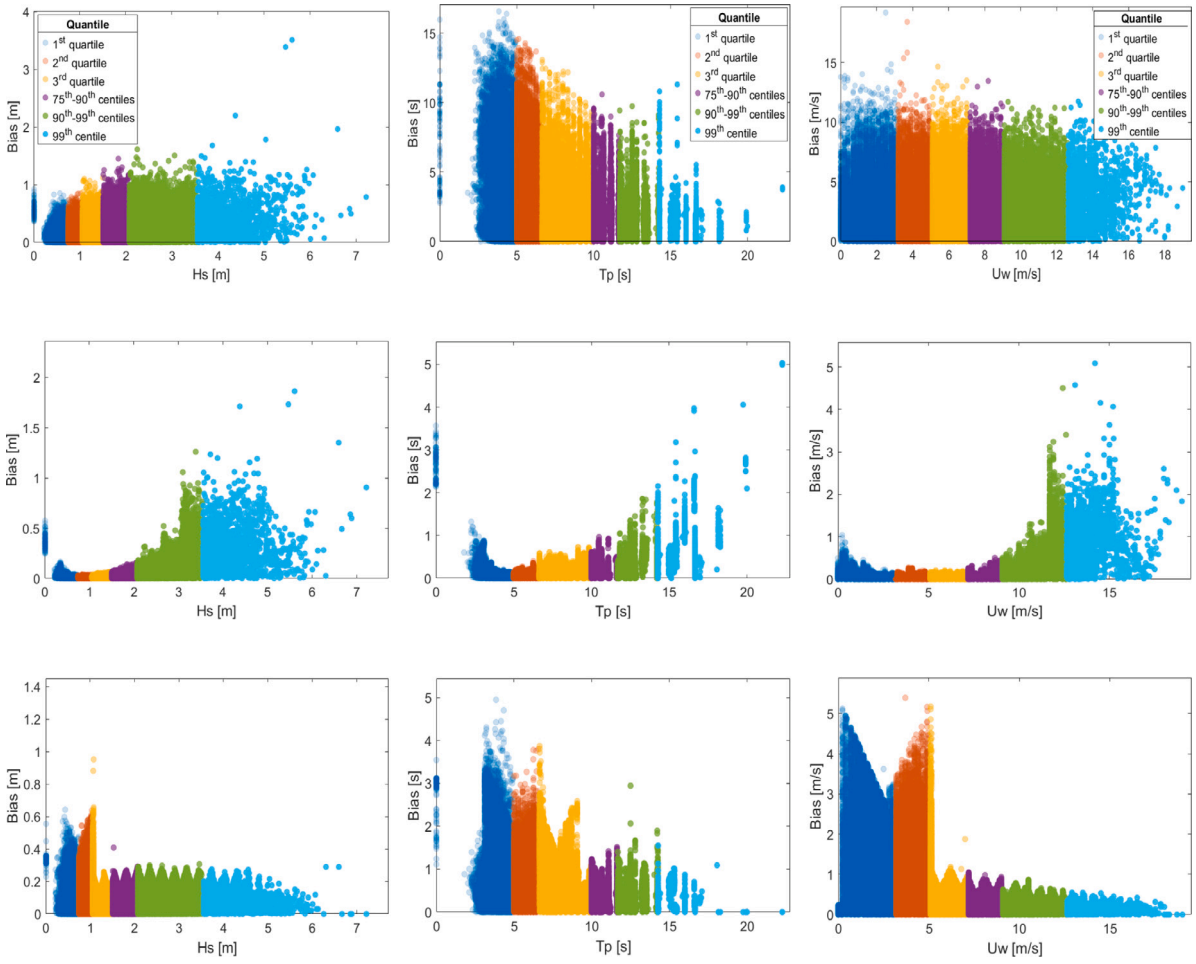
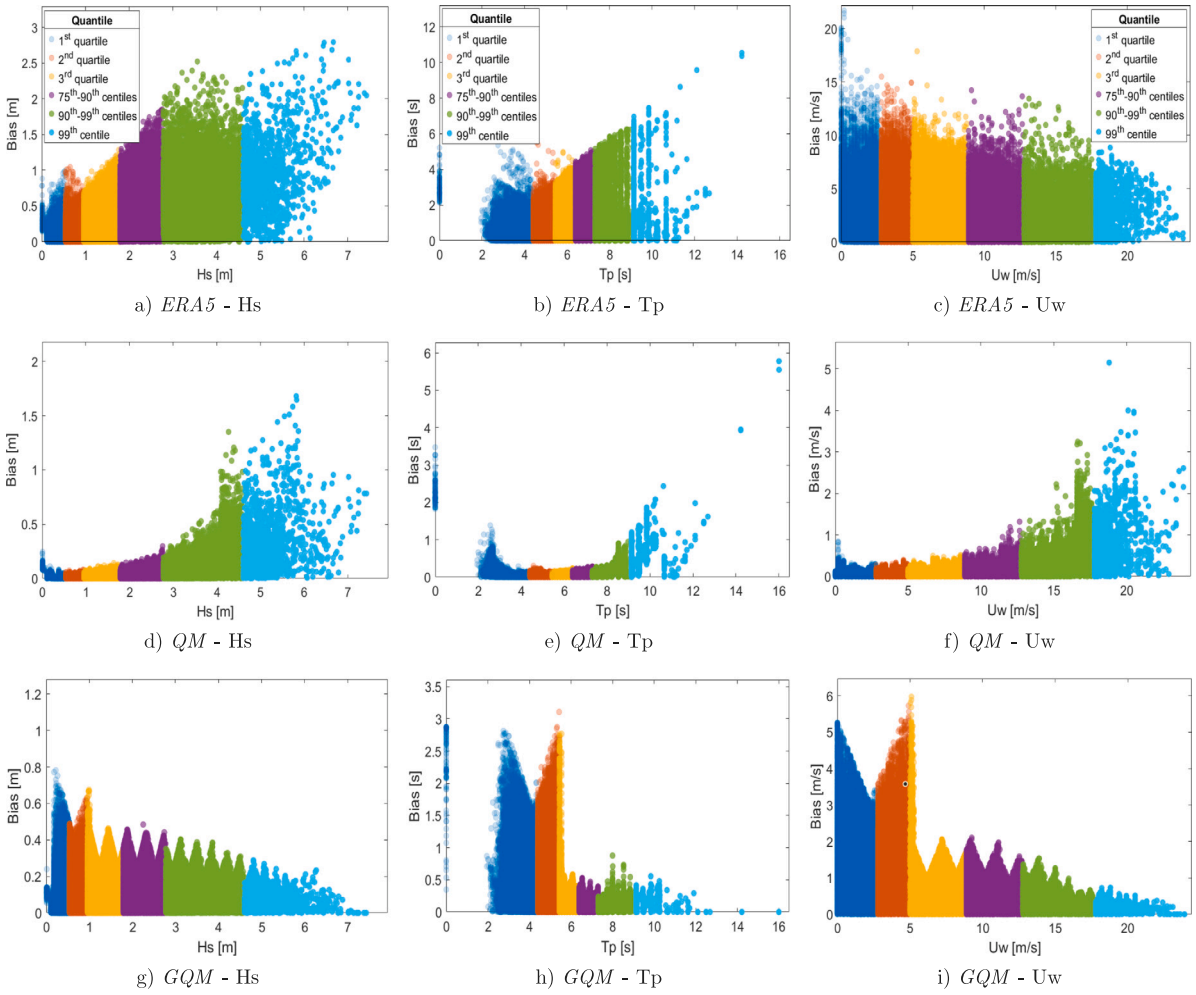


Fig. A.2. Conditional bias for each partitioning and BC technique including H_s, T_p and U_w for ERA5 (a–c), *QM* (d–f) and *GQM* (g–i) at GC.



a) *ERA5* - Hs b) *ERA5* - Tp c) *ERA5* - Uw

d) *QM* - Hs e) *QM* - Tp f) *QM* - Uw

g) *GQM* - Hs h) *GQM* - Tp i) *GQM* - Uw

Fig. A.3. Conditional bias for each partitioning and BC technique including H_s, T_p and U_w for ERA5 (a–c), *QM* (d–f) and *GQM* (g–i) at CC.

References

[1] UN, Adoption of the paris agreement, 2015, https://www.iea.org/publications/freepublications/publication/ElectricityInformation2017Overview.pdf.

[2] IPCC, Synthesis Report of the IPCC Sixth Assessment Report (AR6), Tech. Rep., Intergovernmental Panel on Climate Change (IPCC), ISBN: 978-92-9169-151-7, 2023, URL https://report.ipcc.ch/ar6syr/pdf/IPCC_AR6_SYR_SPM.pdf.

[3] IPCC, Global Warming of 1.5°C, Tech. Rep., Intergovernmental Panel on Climate Change (IPCC), ISBN: 978-92-9169-151-7, 2018.

[4] IRENA, Future of Wind: Deployment, investment, grid integration and socio-economic aspects (A Global Energy Transformation paper), Tech. Rep., International Renewable Energy Agency, Abu Dhabi, ISBN: 978-92-9260-121-8, 2019, Available in https://www.irena.org/publications/2019/Apr/Global-energy-transformation-A-roadmap-to-2050-2019Edition.

[5] Stéphanie Bouckaert, A.F. Pales, C. McGlade, U. Remme, B. Wanner, Net Zero by 2050: A Roadmap for the Global Energy Sector, Tech. Rep., International Energy Agency, Paris, 2021, p. 224, URL https://www.iea.org/reports/net-zero-by-2050.

[6] Ocean Energy Europe, 2030 Ocean Energy Vision, Tech. Rep., 2020, URL https://www.oceanenergy-europe.eu/wp-content/uploads/2020/10/OEE_2030_Ocean_Energy_Vision.pdf.

[7] NREL, Marine Energy in the United States : An Overview of Opportunities, Tech. Rep., (February) 2021, URL https://www.nrel.gov/docs/fy21osti/78773.pdf.

[8] M. Penalba, J.I. Aizpurua, A. Martinez-perurena, On the definition of a risk index based on long-term metocean data to assist in the design of Marine Renewable Energy systems, Ocean Eng. 242 (March) (2021) 110080, http://dx.doi.org/10.1016/j.oceaneng.2021.110080.

[9] A.F. Haselsteiner, E. Mackay, K.-d. Thoben, Reducing conservatism in highest density environmental contours, Appl. Ocean Res. 117 (November) (2021) 102936, http://dx.doi.org/10.1016/j.apor.2021.102936.

[10] S. Rose, J. Apt, Quantifying sources of uncertainty in reanalysis derived wind speed, Renew. Energy 94 (2016) 157–165, http://dx.doi.org/10.1016/j.renene.2016.03.028.

[11] B. Robertson, H. Bailey, D. Clancy, J. Ortiz, B. Buckham, Influence of wave resource assessment methodology on wave energy production estimates, Renew. Energy 86 (2016) 1145–1160, http://dx.doi.org/10.1016/j.renene.2015.09.020.

[12] I. Fairley, H.C. Smith, B. Robertson, M. Abusara, I. Masters, Spatio-temporal variation in wave power and implications for electricity supply, Renew. Energy 114 (2017) 154–165, http://dx.doi.org/10.1016/j.renene.2017.03.075.

[13] B. Reguero, I. Losada, F. Méndez, A recent increase in global wave power as a consequence of oceanic warming, Nature Commun. 10 (205) (2019) http://dx.doi.org/10.1038/s41467-018-08066-0.

[14] Markel Penalba, J.I. Aizpurua, A. Martinez-Perurena, G. Iglesias, A data-driven long-term metocean data forecasting approach for the design of marine renewable energy systems, Renew. Sustain. Energy Rev. 167 (2022) 112751, http://dx.doi.org/10.1016/j.rser.2022.112751.

[15] ISO, ISO 19901-1:2015 Petroleum and natural gas industries - specific requirements for offshore structures - Part 1: Metocean design and operating considerations., Tech. Rep., International Organization for Standardizatio, ISO/TC 67/SC 7 Offshore structures, 2015, p. 206.

[16] IMAREST, Metocean procedures guide for offshore renewables, Tech. Rep., Institute of Marine Engineering, Science & Technology (Offshore Renewables Special Interest Group, 2018.

[17] D. Christie, S.P. Neill, P. Arnold, Characterising the wave energy resource of lanzarote, canary islands, Renew. Energy 206 (2023) 1198–1211, http://dx.doi.org/10.1016/j.renene.2023.02.126, URL https://www.sciencedirect.com/science/article/pii/S0960148123002768.

[18] L. Rusu, C. Guedes Soares, Local data assimilation scheme for wave predictions close to the portuguese ports, J. Oper. Oceanogr. 7 (2) (2014) 45–57.

[19] M.A. Hoque, W. Perrie, S.M. Solomon, Application of SWAN model for storm generated wave simulation in the Canadian beaufort sea, J. Ocean Eng. Sci. 5 (1) (2020) 19–34.

[20] K. Amarouche, A. Akpınar, M.B. Soran, S. Myslenkov, A.G. Majidi, M. Kankal, V. Arkhipkin, Spatial calibration of an unstructured SWAN model forced with CFSR and ERA5 winds for the black and azov seas, Appl. Ocean Res. 117 (2021) 102962.

[21] F. Islek, Y. Yuksel, Inter-comparison of long-term wave power potential in the black sea based on the SWAN wave model forced with two different wind fields, Dyn. Atmos. Oceans 93 (2021) 101192.

[22] A. Ulazia, M. Penalba, G. Ibarra-Berastegui, J. Ringwood, J. Saénz, Wave energy trends over the Bay of Biscay and the consequences for wave energy converters, Energy 141 (2017) http://dx.doi.org/10.1016/j.energy.2017.09.099.

[23] S. Carreno-Madinabeitia, G. Ibarra-Berastegi, J. Sáenz, A. Ulazia, Long-term changes in offshore wind power density and wind turbine capacity factor in the Iberian Peninsula (1900–2010), Energy 226 (2021) http://dx.doi.org/10.1016/j.energy.2021.120364.

[24] C. Teutschbein, J. Seibert, Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, J. Hydrol. 456 (2012) 12–29.

[25] D. Maraun, Bias Correcting Climate Change Simulations - a Critical Review, Curr. Clim. Chang. Rep. 2 (4) (2016) 211–220, http://dx.doi.org/10.1007/S40641-016-0050-X.

[26] U. Ehret, E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, J. Liebert, HESS opinions "should we apply bias correction to global and regional climate model data?", Hydrol. Earth Syst. Sci. 16 (9) (2012) 3391–3404, http://dx.doi.org/10.5194/hess-16-3391-2012, URL https://hess.copernicus.org/articles/16/3391/2012/.

[27] G. Lenderink, A. Buishand, W. Van Deursen, Estimates of future discharges of the river rhine using two scenario methodologies: direct versus delta approach, Hydrol. Earth Syst. Sci. 11 (3) (2007) 1145–1159.

[28] P.J. Block, F.A. Souza Filho, L. Sun, H.-H. Kwon, A streamflow forecasting framework using multiple climate and hydrological models 1, JAWRA J. Am. Water Resour. Assoc. 45 (4) (2009) 828–843.

[29] R. Rojas, L. Feyen, A. Dosio, D. Bavera, Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations, Hydrol. Earth Syst. Sci. 15 (8) (2011) 2599–2620.

[30] C. Piani, J. Haerter, E. Coppola, Statistical bias correction for daily precipitation in regional climate models over europe, Theor. Appl. Climatol. 99 (1) (2010) 187–192.

[31] F. Sun, M.L. Roderick, W.H. Lim, G.D. Farquhar, Hydroclimatic projections for the murray-darling basin based on an ensemble derived from intergovernmental panel on climate change AR4 climate models, Water Resour. Res. 47 (12) (2011).

[32] I.J.L. Michele Martini, Raúl Guanche, José A. Armesto, C. Vidal, Met-ocean conditions influence on floating offshore wind farms power production, Wind Energy 19 (2016) 339—-420, http://dx.doi.org/10.1002/we.

[33] I. Fairley, M. Lewis, B. Robertson, M. Hemer, I. Masters, J. Horrillo-Caraballo, H. Karunarathna, D.E. Reeve, A classification system for global wave energy resources based on multivariate clustering, Appl. Energy 262 (October 2019) (2020) 114515, http://dx.doi.org/10.1016/j.apenergy.2020.114515.

[34] A. Martinez, G. Iglesias, Wave exploitability index and wave resource classification Significant height of combined wind waves and swell, Renew. Sustain. Energy Rev. 134 (March) (2020) 110393, http://dx.doi.org/10.1016/j.rser.2020.110393.

[35] G. Rinaldi, A. Garcia-Teruel, H. Jeffrey, P.R. Thies, L. Johanning, Incorporating stochastic O&M models into the techno-economic analysis of floating offshore wind farms, Appl. Energy 301 (July) (2021) 117420, http://dx.doi.org/10.1016/j.apenergy.2021.117420.

[36] M. Penalba, A. Ulazia, G. Ibarra-Berastegui, J. Ringwood, J. Sáenz, Wave energy resource variation off the west coast of Ireland and its impact on realistic wave energy converters' power absorption, Appl. Energy 224 (2018) 205–219, http://dx.doi.org/10.1016/j.apenergy.2018.04.121, URL https://www.sciencedirect.com/science/article/pii/S0306261918306895.

[37] M. Badger, T. Ahsbahs, P. Maule, I. Karagali, Inter-calibration of SAR data series for offshore wind resource assessment, Remote Sens. Environ. 232 (June) (2019) 111316, http://dx.doi.org/10.1016/j.rse.2019.111316.

[38] C.E. Hiles, B. Robertson, B.J. Buckham, Extreme wave statistical methods and implications for coastal analyses, Estuar. Coast. Shelf Sci. 223 (April) (2019) 50–60, http://dx.doi.org/10.1016/j.ecss.2019.04.010.

[39] M.M. Nezhad, M. Neshat, A. Heydari, A. Razmjoo, G. Piras, D.A. Garcia, A new methodology for offshore wind speed assessment integrating Sentinel-1, ERA-Interim and in-situ measurement, Renew. Energy 172 (2021) 1301–1313, http://dx.doi.org/10.1016/j.renene.2021.03.026.

[40] G. Lemos, M. Menendez, A. Semedo, P. Camus, M. Hemer, M. Dobrynin, P.M. Miranda, On the need of bias correction methods for wave climate projections, Glob. Planet. Change 186 (December 2019) (2020) 103109, http://dx.doi.org/10.1016/j.gloplacha.2019.103109.

[41] ECMWF, Browse reanalysis datasets, 2023, URL https://www.ecmwf.int/en/forecasts/datasets/browse-reanalysis-datasets, Last accessed: 29/08/2023.

[42] A. Ulazia, M. Penalba, A. Rabanal, G. Ibarra-Berastegi, J. Ringwood, J. Sáenz, Historical Evolution of the Wave Resource and Energy Production off the Chilean Coast over the 20th Century, Energies 11 (9) (2018) 2289, http://dx.doi.org/10.3390/en11092289, URL http://www.mdpi.com/1996-1073/11/9/2289.

[43] S. Carreno-Madinabeitia, G. Ibarra-Berastegi, J. Sáenz, A. Ulazia, Long-term changes in offshore wind power density and wind turbine capacity factor in the iberian peninsula (1900–2010), Energy 226 (2021) 120364, http://dx.doi.org/10.1016/j.energy.2021.120364, URL https://www.sciencedirect.com/science/article/pii/S0360544221006137.

[44] L.E. Hay, R.L. Wilby, G.H. Leavesley2, A Comparison of Delta Change and Downscaled Gcm Scenarios for Three Mounfainous Basins in the United States', Tech. Rep., 36, (2) 2000, URL http://climate.usu.edu/Free/.

[45] E. Gumbel, Les valeurs extrêmes des distributions statistiques, Ann. l'Inst. Henri Poinc. 5 (1935) 115–158.

[46] G. Cervelli, L. Parrinello, C. Moscoloni, G. Giorgi, Comparison of the ERA5 wave forecasting dataset against buoy record., Instrum. Mesures Métrol. 21 (3) (2022).

[47] S.E. Perkins, A.J. Pitman, N.J. Holbrook, J. McAneney, Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions, J. Clim. 20 (17) (2007) 4356–4376, http://dx.doi.org/10.1175/JCLI4253.1.

[48] P.M. Soares, R.M. Cardoso, A simple method to assess the added value using high-resolution climate distributions: Application to the EURO-CORDEX daily precipitation, Int. J. Climatol. 38 (3) (2018) 1484–1498, http://dx.doi.org/10.1002/joc.5261.