

On the temporal sensitivity of the reference data in bias correction techniques for improving metocean datasets

F. Callea, A. Martinez-Perurena, A. Zarketa-Astigarraga

Fluid Mechanics Department, Mondragon University, Arrasate, Gipuzkoa, Spain.

M. Penalba

Fluid Mechanics Department, Mondragon University, Arrasate, Gipuzkoa, Spain.

Ikerbasque - Basque Foundation for Science, Bilbao, Bizkaia, Spain.

G. Cervelli & G. Giorgi

Marine Offshore Renewable Energy Lab (MOREnergy Lab), Politecnico di Torino, Torino, Italy.

B. Robertson

Civil and Construction Engineering, Oregon State University, Corvallis, OR, USA

G. Iglesias

MaREI, Environmental Research Institute & School of Engineering, UCC, Cork, Ireland

University of Plymouth, School of Engineering, Computing & Mathematics, Plymouth, UK

ABSTRACT: The paper presents a preliminary study limited to the use of wave height for the data corresponding to the Gulf of Biscay, a location with a very dominant North-West wave rose. It benchmarks different bias correction (BC) techniques and evaluates their performance, as well as the sensitivity of the BC to the reference dataset employed for the identification of the BC parameters. More precisely, the amount of data (number of years) and the selected period (exact years) are analysed. Overall, the results demonstrate that the Gumbel-based BC techniques overperform the linearly-spaced BC techniques, the directional-adjusted Gumbel Quantile Mapping technique showing the lowest bias. With respect to the sensitivity of the reference data, BC seems to provide satisfactory results even when only 1 year of data are used. However, the dispersion among the different selected periods is large, resulting in large uncertainties. This dispersion reduces significantly once 3 or more years of data are used, independently of the selected period.

1 INTRODUCTION

Offshore renewable energy (ORE) systems are one of the alternatives that can assist the energy transition towards a carbon-neutral energy system. In fact, The International Energy Agency predicts that about 45% of CO₂ emissions savings by 2050 will come from technologies that are still under development (Stéphanie Bouckaert et al. 2021). 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.

The effective and reliable design of successful ORE technologies, including all the aspects mentioned above, relies on accurate metocean data. The use of incomplete or inaccurate metocean data can incorporate a higher uncertainty to a design process; an area where the uncertainty level is

already significant (Haselsteiner et al. 2021). 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 (Penalba et al. 2021). In this sense, ORE 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. Therefore, understanding and reducing the uncertainty of the metocean data across the whole operational domain including both regions, is of high relevance.

The most common sources of metocean data in the ORE sector are observation buoys and re-analysis datasets. 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. Therefore, re-analysis datasets and data from climate models (referred to as *simulated* datasets in the following) are often used. These datasets cover very long time periods (can go back to 1900 in some cases (Ulazia et al. 2017, Carreno-Madinabeitia et al. 2021)) and provide metocean data at any point in the world for virtually no cost. However, the main problem of *simulated* datasets is their limited accuracy under certain conditions and, thus, calibration techniques based on observations are used to reduce the differences (bias). To the best knowledge of the authors, the first benchmarking of different BC techniques is in the offshore context is presented in (Lemos et al. 2020), where the authors exclusively focus on the significant wave height (H_s) and use ERA5 datasets are the reference dataset for BC, while (Penalba et al. 2023) present an extension of that study with a focus on the ORE sector and using *in-situ* observations at four diverse locations for the assessment of the different BC techniques. In addition, specific statistical metrics for evaluating BC techniques in the ORE context, considering both the *Operational* and the *Survivability* regions, are presented in (Penalba et al. 2023): The *PDF – score* and the Distribution Added Value (*DAV*).

Despite the broad assessment of different BC correction methods for the ORE industry and the clear conclusion endorsing the quantile-mapping (QM) techniques, particularly that based on the Gumbel distribution, two main aspects remain unanswered:

- i. The selection of the reference datasets for BC: from the type of data (re-analysis datasets (Lemos et al. 2020) or observations (Penalba et al. 2023)), to their temporal and spatial coverage (*i.e.* the minimum amount of data).
- ii. The impact of data correction on the final design parameters of the different ORE technology components and subsystems.

A preliminary analysis of the latter is covered in (Penalba et al. 2023), so the present study aims at shedding light on the former by incorporating a directional BC technique and performing a preliminary temporal sensitivity analysis of the reference data used for BC. Hence, the remainder of the paper is as follows: Section 2 presents the considered BC techniques; Section 3 describes the methodology, with the specific statistical techniques, and the case study; Section 4 presents the most relevant results; and Section 5 draws the main conclusions.

2 BIAS CORRECTION TECHNIQUES

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, meaning that no knowledge of the underlying physics is needed (Ehret et al. 2012), and have become a common practice in climate and meteorological studies in the last two decades (Teutschbein & Seibert 2012, Maraun 2016). (Penalba et al. 2023) concludes that only techniques based on QM provide the necessary fidelity. Therefore, the two QM techniques that show the best results in (Penalba et al. 2023), *i.e.* the linearly-spaced QM (LQM) and the Gumbel-based QM (GQM) described in Sections 2.1 and 2.2, respectively, are considered here.

2.1 Linearly-spaced Quantile mapping

The LQM method computes the correction factor X^{LQM} at each quantile (q_j) of the Cumulative Density Function. To that end, firstly the *simulated* and measured datasets are distributed into different quantiles. In this study, similarly to (Penalba et al. 2023), 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,

$$X^{LQM}(q_j) = CDF_{as}^{-1}(q_j) - CDF_{obs}^{-1}(q_j). \quad (1)$$

This difference between the inverse CDFs is then fitted with an n -order polynomial function. Hence, the dataset corrected at each quantile via the LQM technique is given as,

$$y^{BC}(q_j) = y^{sim}(q_j) + f(X^{LQM}(q_j), n). \quad (2)$$

where y represents the metocean variable, and y^{BC} and y^{sim} refer to the bias corrected and *simulated* variables, respectively.

2.2 Gumbel-based quantile mapping

The GQM technique allows for adequately capturing the error in extreme events by means of a more dense characterisation of the higher quantiles, while preserving the main information around the most common quantiles. The GQM technique, uses the same quantile mapping technique as in the LQM technique but placing the quantiles following a Gumbel distribution function (Gumbel 1935):

$$F(x; \mu, \beta) = e^{-e^{-(x-\mu)/\beta}}, \quad (3)$$

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.

Once the quantiles are identified, the GQM technique is implemented following Equations (1) and (2).

2.3 Directional adjusted quantile mapping

The LQM and GQM methods use univariate metocean data. However, the characteristics of metocean data can vary significantly depending on the wave/wind direction. In fact, data can be corrected differently for each wave/wind direction (Semedo et al. 2011, Semedo et al. 2018). Therefore, an additional BC method where the calibration is adjusted to each directional sector is also included: the directional adjusted QM (DAQM) technique (Mínguez et al. 2011). The DAQM is based on a similar approach as the LQM and the GQM techniques when as for the use of quantiles, but has two main differences. On the one hand, the DAQM presented here is based on a parametric nonlinear regression problem instead of a polynomial function and, on the other hand, the correction parameters vary smoothly along the range of possible wave/wind directions by means of cubic splines:

$$y^{BC} = a^{DAQM}(\theta)(y^{sim})^{b^{DAQM}(\theta)} \quad (4)$$

where a^{DAQM} and b^{DAQM} are the correction spline parameters that depend on the wave/wind direction under analysis θ . These correction parameters are identified so that the solution of the optimisation problem returns the calibrated values of the variable y (Mínguez et al. 2011):

$$\begin{aligned} \min(a, b) &= [\sum (y_i^{obs} - y_i^{BC})^2] = \\ &= \sum [y_i^{obs} - a_i^{DAQM}(\theta_i)(y_i^{sim})^{b_i^{DAQM}(\theta_i)}]^2, \\ i &= 1, \dots, S \end{aligned} \quad (5)$$

where S is the number of sectors. The complete 360° directional range is divided into 360 moving sectors ($S = 360$). However, each of these sectors covers a range of 22.5°, rotating around the directional circumference 1° at a time. Note that, in the following, the DAQM technique is analysed based on both QM methods, the linearly-spaced quantiles (DALQM) and the Gumbel distribution (DAGQM).

3 METHODOLOGY AND CASE STUDY

The present study aims at analysing the suitability of the directional BC technique and the characteristic of the dataset used as reference data for computing BC techniques. To that end, metocean data from *in-situ* observations and ERA5 re-analysis datasets are used. In this preliminary study, the analysis is limited to the Bay of Biscay and H_s , as described in Section 3.1. The main statistical metrics for evaluating the performance of each BC technique are briefly described in Section 3.2. Finally, the definition of the different reference datasets for the sensitivity analysis is explained in Section 3.3.

3.1 Wave data

As a preliminary analysis, the case study includes only H_s data corresponding to the Bay of Biscay that represents 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) predominantly incoming from the North-West direction, as shown in Figure 1.

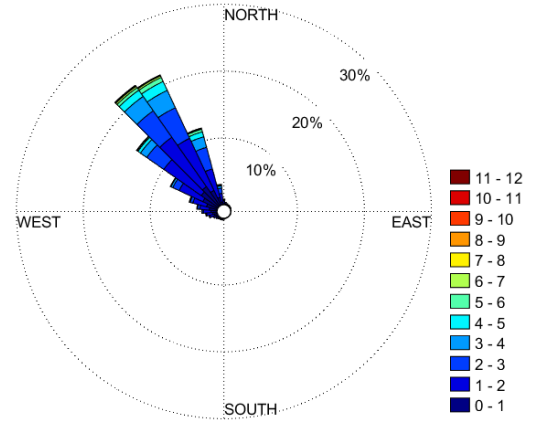


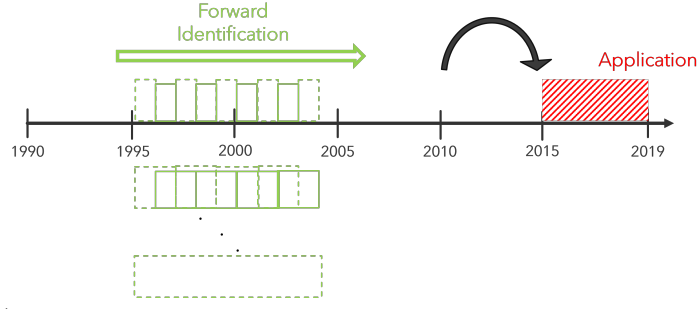
Figure 1: Wave rose of the wave data in Bay of Biscay.

3.2 Statistical metrics

The assessment of the different BC techniques and datasets used as reference data for the identification of the BC parameters is carried out based on three main statistical metrics.

The most essential metric is the bias, which is defined as the difference between the ground truth represented by y^{obs} , and either y^{sim} or y^{BC} . In order to provide a single metric for each variable and reference data scenario, the mean of the absolute bias is considered as,

$$bias(y^{obs}, y^{sim} || y^{BC}) = \text{mean}(abs(y^{obs} - y^{sim} || y^{BC})). \quad (6)$$



a) Schematic representation of the different reference data scenarios.

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
1-year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
2-year		1995-96	1996-97	1997-98	1998-99	1999-00	2000-01	2001-02	2002-03	2003-04
3-year			1995-97	1996-98	1997-99	1998-00	1999-01	2000-02	2001-2003	2002-04
4-year										
5-year										
6-year										
7-year										
8-year										
9-year									1995-03	1996-04
10-year										1995-04

b) Definition of the reference data scenarios.

Figure 2: Representation of the reference data scenarios: (a) schematic diagram and (b) definition.

In addition, two specific statistical metrics suggested in (Penalba et al. 2023) are also employed. On the one hand, the *PDF - score* is a simple but practical metric of the similarity of two PDFs, allowing the comparison along the entire distribution (Perkins et al. 2007):

$$PDFscore = \sum_{m=1}^M \min(PDF(y_m^{obs}) - PDF(y_m^{as})), \quad (7)$$

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

On the other hand, the *DAV* metric allows for a normalised comparison between two *PDF - scores* as follows (Soares and Cardoso 2018),

$$DAV = \frac{PDF - score_{BC} - PDF - score_{as}}{PDF - score_{as}} \times 100, \quad (8)$$

where $PDF - score_{BC}$ and $PDF - score_{as}$ represent the *PDF - scores* of the corrected and raw assimilated datasets, respectively.

3.3 Reference data scenarios

Due to the well known intra- and inter-annual variations of the resource, two main aspects are criti-

cal for the definition of the reference data scenarios: the amount of data (number of months/years) and the specific period in time (decade, year or season). For the evaluation of the performance of the BC technique, a validation dataset is also essential, which should be the same for all reference datasets.

In the present study, as illustrated in Figure 2 (a), reference datasets of different lengths and corresponding to different periods between 1995-2004 are analysed, using as the validation dataset the period between 2015-2019. Figure 2 (b) defines all the different scenarios.

4 RESULTS

Results are divided in two main parts. First a benchmarking of different BC techniques is carried out in Section 4.1 to analysed the performance of each techniques. Once the most suitable BC technique is identified, the sensitivity analysis of the reference dataset is conducted in Section 4.2, where the impact of considering shorter and longer datasets is assessed by means of different statistical metrics.

4.1 BC benchmarking

Figure 3 (a) depict the bias of the raw ERA5 re-analysis dataset as a function of the wave direction and the wave height, showing a bias between -0.5 and 0.25 across most of the polar diagram, except

for extreme events for which the bias increases considerably. Figures 3(b)-(e) illustrate the bias of the four BC techniques, where all of them are shown to improve the quality of the data significantly. As expected, the bias of the LQM-based BC techniques continues showing a considerable bias for the higher H_s quantiles, which is significantly reduced by GQM-based techniques. In addition, the directional-adjusted techniques demonstrate a better performance, reducing the bias ranges in about 50% in both cases. As a conclusion, the best performance is shown by the DAGQM technique, which

is consistent with (Lemos et al. 2020) and, thus, selected for the temporal sensitivity analysis.

4.2 Temporal sensitivity

The sensitivity analysis addresses the characteristics of the BC in each scenario. Figure 4 (a) illustrates the mean bias off the *simulated* dataset (-0.4 approximately) and the mean bias among the different scenarios with the same amount of data used for the identification of BC parameters (*i.e* identification years). The boxplot graph shows the

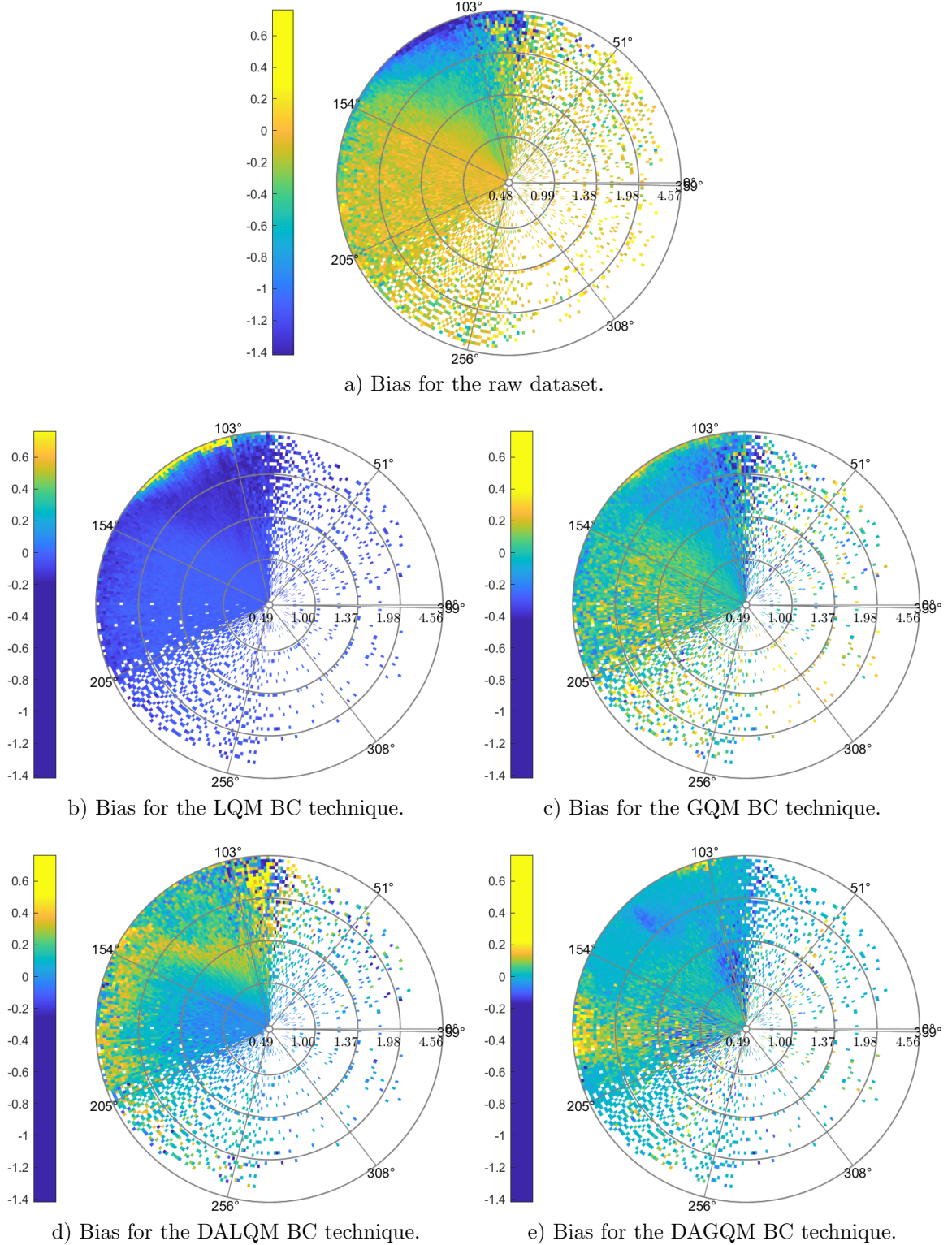


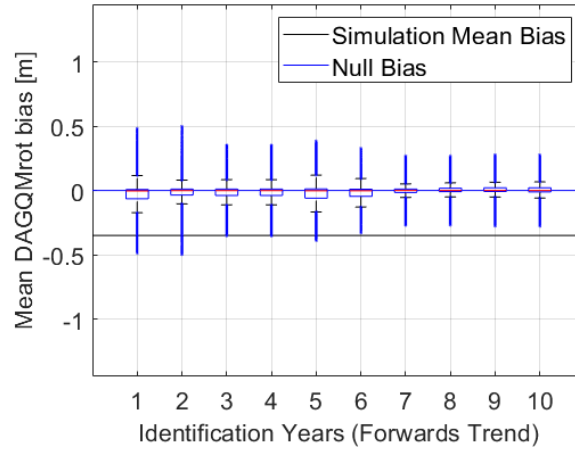
Figure 3: Bias comparison of unidirectional and directional adjusted BC techniques: Bias of the (a) raw dataset, (b) .

mean (in red), the interquartile range (IQR), and the dispersion in terms of minimum, maximum and outliers. All these metrics are shown to shrink while converging towards the 0 bias, demonstrating that the more data are considered in the reference dataset, the better the performance of the BC technique is. However, no improvement is observed once the data corresponding to seven years is included.

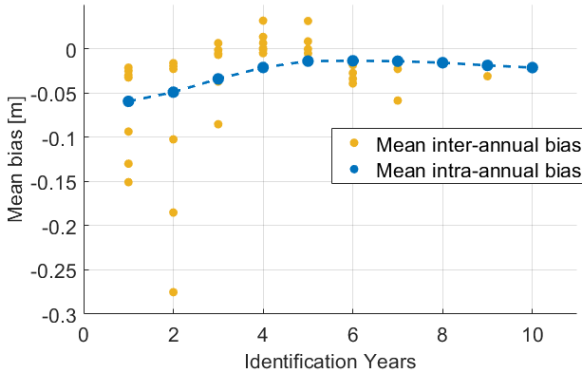
Figures 4 (b)-(e) decompose the information of the boxplot graph by showing the results of the mean bias, mean dispersion, mean PDF-score and DAV, respectively, differing between intra- (in yellow) and inter-annual (in blue) results for the dif-

ferent identification years. Intra-annual results refer to the mean computed over each scenario, while the inter-annual results refer to the mean of the results computed at each set of identification years. Hence, mean bias and dispersion are shown to converge once three or more years of data are considered in the reference dataset, while the PDF-score and DAV show the optimal results when including 5 years of data. However, the PDF-score and DAV show poorer results when more than 5 years are included, although the variation is insignificant.

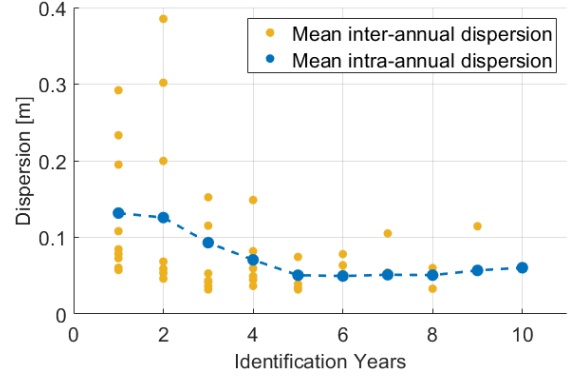
Although the mean inter-annual results are shown to converge very quickly, the dispersion among the different mean intra-annual results is



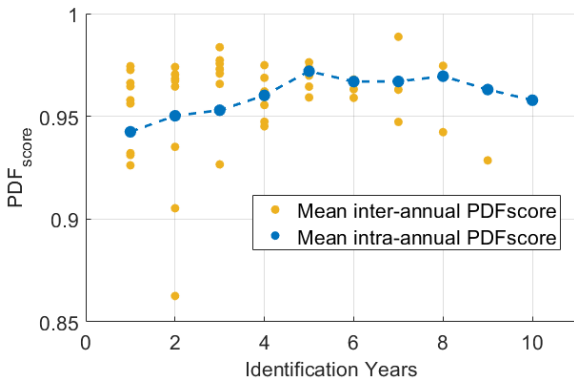
a) Boxplot diagram.



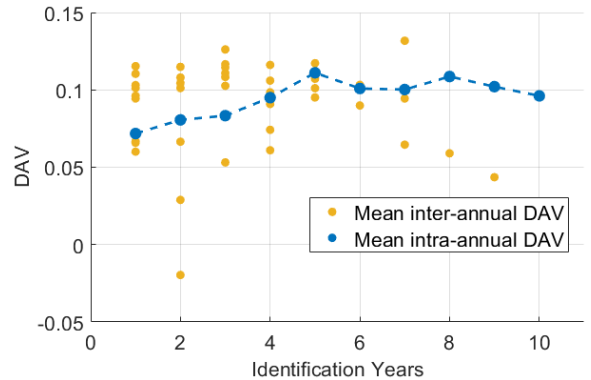
b) mean bias.



c) mean dispersion.



d) mean PDF-score.



e) mean DAV.

Figure 4: Temporal sensitivity of BC techniques: (a) boxplot representation (b) mean bias, (c) mean dispersion, (d) mean PDF-score and (e) mean DAV.

significant when only one or two years of data is considered for the identification of BC parameters. This means that when a limited amount of data is used, the quality of the correction is sensitive to the selected period and, thus, the uncertainty of the BC is large. In contrast, when more than 3 years of data are used, that uncertainty is reduced significantly.

5 CONCLUSIONS AND FUTURE WORK

The present paper studies the impact of wave direction and selection of reference data for the calibration of metocean data via bias correction (BC) techniques. The linearly-spaced (LQM) and Gumbel based quantile mapping (GQM) techniques are first implemented, extending these two techniques with the directional-adjusted adaptations, DALQM and DAGQM, respectively. The paper shows a preliminary study limited to the use of wave height for the data corresponding to the Gulf of Biscay, a location with a very dominant North-West wave rose.

On the one hand, the Gumbel-based techniques are shown to overperform the linearly-spaced techniques, being the DAGQM the technique with the lower bias. It is to be expected that, for a location with a broader wave rose, the directional-adjusted methods will show an even better performance, although this needs to be verified by extending the analysis to other regions and datasets.

On the other hand, the sensitivity to the amount of data (number of years) and the period (exact years) considered as the reference dataset is evaluated via the mean bias, dispersion, PDF-score and distribution added value (DAV). All parameters are shown to reach satisfactory values when including 3 or more years of data, independently of the selected years. Although reduced datasets (1 or 2 years of data) also show relatively good results, the dispersion among the different years is high, which results in a large uncertainty of the calibration.

ACKNOWLEDGEMENTS

This publication is part of the research project PID2021-124245OA-I00 funded by MCIN/AEI/10.13039/501100011033 and by ERDF A way of making Europe, and the research project funded by the Basque Government's ELKARTEK 2024 program under the grant No. KK-2024/00086. Partial funding is related to 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). 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. 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.

REFERENCES

- Carreno-Madinabeitia, S., G. Ibarra-Berastegi, J. Sáenz, & A. Ulazia (2021). Long-term changes in offshore wind power density and wind turbine capacity factor in the Iberian Peninsula (1900–2010). *Energy* 226.
- Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, & J. Liebert (2012). Hess opinions "should we apply bias correction to global and regional climate model data?". *Hydrology and Earth System Sciences* 16(9), 3391–3404.
- Gumbel, E. (1935). Les valeurs extrêmes des distributions statistiques. *Annales de l'Institut Henri Poincaré* 5, 115–158.
- Haselsteiner, A. F., E. Mackay, & K.-d. Thoben (2021). Reducing conservatism in highest density environmental contours. *Applied Ocean Research* 117(November), 102936.
- Lemos, G., M. Menendez, A. Semedo, P. Camus, M. Hemer, M. Dobrynin, & P. M. Miranda (2020). On the need of bias correction methods for wave climate projections. *Global and Planetary Change* 186(December 2019), 103109.
- Maraun, D. (2016, dec). Bias Correcting Climate Change Simulations - a Critical Review. *Current Climate Change Reports* 2(4), 211–220.
- Mínguez, R., A. Espejo, A. Tomás, F. Méndez, & I. Losada (2011). Directional calibration of wave reanalysis databases using instrumental data. *Journal of Atmospheric and Oceanic Technology* 28(11), 1466–1485.
- Penalba, M., J. I. Aizpurua, & A. Martinez-perurena (2021). On the definition of a risk index based on long-term metocean data to assist in the design of Marine Renewable Energy systems. *Ocean Engineering* 242(March), 110080.
- Penalba, M., C. Guo, A. Zarketa-Astigarraga, G. Cervelli, G. Giorgi, & B. Robertson (2023). Bias correction techniques for uncertainty reduction of long-term metocean data for ocean renewable energy systems. *Renewable Energy* 219, 119404.
- Penalba, M., A. Zarketa-astigarraga, P. Branson, & B. Robertson (2023). Impact of resource uncertainties on the design of wave energy converters. pp. 529.
- Perkins, S. E., A. J. Pitman, N. J. Holbrook, & J. McAneney (2007). Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *Jour-*

- nal of Climate* 20(17), 4356–4376.
- Semedo, A., M. Dobrynin, G. Lemos, A. Behrens, J. Staneva, H. De Vries, A. Sterl, J.-R. Bidlot, P. M. Miranda, & J. Murawski (2018). C mip5-derived single-forcing, single-model, and single-scenario wind-wave climate ensemble: Configuration and performance evaluation. *Journal of Marine Science and Engineering* 6(3), 90.
- Semedo, A., K. Sušelj, A. Rutgersson, & A. Sterl (2011). A global view on the wind sea and swell climate and variability from era-40. *Journal of Climate* 24(5), 1461–1479.
- Soares, P. M. & R. M. Cardoso (2018). A simple method to assess the added value using high-resolution climate distributions: Application to the EURO-CORDEX daily precipitation. *International Journal of Climatology* 38(3), 1484–1498.
- Stéphanie Bouckaert, A. F. Pales, C. McGlade, U. Remme, & B. Wanner (2021). Net Zero by 2050: A Roadmap for the Global Energy Sector. Technical report, International Energy Agency, Paris.
- Teutschbein, C. & J. Seibert (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of hydrology* 456, 12–29.
- Ulazia, A., M. Penalba, G. Ibarra-Berastegui, J. Ringwood, & J. Saénz (2017). Wave energy trends over the Bay of Biscay and the consequences for wave energy converters. *Energy* 141.