Validation of ERA5 Wave Energy Flux through Sailor diagram in Spain (2005-2014)

J. Sáenz, S. Carreno-Madinabeitia, G. Ibarra-Berastegi, A. Ulazia and M. Garro

Abstract—**The aim of this study is to validate the estimations of Wave Energy Flux (WEF) from ERA5 against observational data. To that end, 0.5° x 0.5° resolution WEF values from ERA5 reanalisis and corresponding data from 15 directional REDEXT buoys of Puertos del Estado surrounding the Iberian and Canary Islands' coast covering a period of ten years (2005-2014) have been used.**

In this study, the Sailor diagram and its methodology [1] have been used to compare the skill of the wave model used in ERA5 for the WEF two-dimensional variable. The methodology on which the Sailor diagram is built proposes a diagram along with different statistical verification indices. In this particular case, WEF is a vectorial magnitude, and its zonal and meridional components (WEFu and WEFv) have been assessed at the same time by applying this methodology.

In order to compare the observed and the ERA5 data, well-known indices extended to two dimensions such as RMSE, correlation and bias have been used, as well as variances of each component. Furthermore, to analyse if the WEFu and WEFv components of the ERA5 data are rotated in relation to the observations, their relative rotation angle, eccentricity, and the congruence coefficient of their first EOF (Empirical Orthogonal Function) have also been calculated. This way, the analysis is extended to fully cover the two dimensional structure of the vector WEF data.

Index Terms—**Wave Energy Flux, ERA5, Sailor-diagram and Bi-dimensional variables.**

I. Introduction

In the field of climate and meteorology, statistical indices such as Root Mean Square Error (RMSE), correlation [2], and standard deviation are essential tools commonly used for the validation of model results against reference data, most often observations [2], [3]. This is the reason that led to the common use of the Taylor diagram [4]. It employs the cosine theorem to geometrically relate the centered RMSE, correlation, and standard deviation indices.

However, the Taylor diagram is based on the scalar relationship between these three coefficients and its objective is to compare the performance of scalar fields. In any case, since it is a very good model evaluation tool, its use has also been extended to vector quantities by using different strategies. In some cases, only the magnitudes of wind [5] or currents has been evaluated. Another option found in the bibliography consists in the use of two different diagrams in order to analyze the agreement of the different models for the zonal and meridional components of the vector magnitudes [6] being evaluated. Finally, some studies average the the Taylor diagrams corresponding to the zonal and meridional components [7] in order to define an overall agreement in both dimensions. In any case, the former three options lead to some limitations, since they do not completely diagnose the directional errors in a vector quantity. In fact, it is a common outcome of the use of individual Taylor diagrams for the zonal and meridional components, that a model which is identified as the optimal one for the zonal component is not the best one for the meridional component [8].

In order to provide a full diagnostic of vector quantities such as wind, currents or, as in this study, Wave Energy Flux (WEF), the authors have developed a new diagram in which it is possible to validate two-dimensional vector variables in a single graph, extending the idea of the Taylor diagram to a two-dimensional scenario. We called it the SailoR diagram [1]. This extension is not trivial, one of the biggest problems is that there is no single definition of the two-dimensional correlation, therefore this new diagram fully analyzes the structure of the two-dimensional mean squared error matrices that can be built by using observations and every model. The computations are reduced if this analysis is performed in the space spanned by the principal components of the dataset. For usual datasets in geophysics, the two-dimensional error matrices are of fully rank and the use of the principal components imply no truncation at all. If the two-dimensional mean squared error matrix is analysed in the space of the two principal components of the data, it can be shown that the error can be expressed as a separate contribution from the time-invariant bias and relative rotations of the principal components corresponding to every model with respect to the ones from the original observations.

The Sailor diagram consist in two parts: the diagram itself and a diagnostic table with statistical indices that are derived from the equations used to build

This paper is part of project PID2020-116153RB-I00 funded by MCIN/AEI/10.13039/501100011033 and has also received funding from the University of the Basque Country (UPV/EHU project GIU20/08). This contribution is part of the TED2021-132109B-C21 research project funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR

J. Sáenz is a member of the Department of Physics and PiE-UPV/EHU, BEGIK (e-mail: jon.saenz@ehu.eus).

S. Carreno-Madinabeitia is working at Department of Mathematics, University of the Basque Country (UPV/EHU) (e-mail: sheila.carreno@ehu.eus).

G. Ibarra-Berastegi is in the Energy Engineering Department and PiE-UPV/EHU, BEGIK (e-mail: gabriel.ibarra@ehu.eus).

A. Ulazia is in the Energy Engineering Department University of the Basque Country (UPV/EHU) (e-mail: alain.ulazia@ehu.eus).

M. Garro is in the Energy Engineering Department University of the Basque Country (UPV/EHU) (e-mail: mikel.garro@ehu.eus)

Digital Object Identifier:
https://doi.org/10.36688/ewtec-2023-305

TABLE I
SAILOR DIAGRAM INDICES FOR SYNTHETIC REFERENCE AND THE SYNTHETIC MOD1 AND MOD2 MODELS

Index	Ref	MOD1	MOD2
sdUx	3.27		
sdUy	6.06		
sdVx		3.27	5.22
sdVy		6.07	4.51
Sigmax	6.42	6.42	6.42
Sigmay	2.52	2.52	2.52
thetau	1.93		
thetav		1.93	2.56
thetavu		0.00	0.52
R2vec	2.00	2.00	2.00
biasMag		8.34	2.88
RMSE		8.34	4.47
Eccentricity	0.92	0.92	0.92
congruenceEOF1	1.00	1.00	0.87

TABLE II
GEOGRAPHICAL LOCATION OF THE 15 BUOYS OBTAINED FROM PUERTOS DEL ESTADO.

Name	Lon (°E)	Lat (°N)
Bilbao Vizcaya (BB)	-3.05	43.64
Cabo Begur (CB)	3.65	41.90
Cabo De Gata (CDG)	-2.34	36.57
Cabo de Palos (CDP)	-0.31	37.65
Cabo Penhas (CP)	-6.18	43.75
Cabo Silleiro (CS)	-9.43	42.12
Dragonera (DR)	2.09	39.56
Estaca de Bares (EB)	-7.68	44.12
Golfo de Cadiz (GDC)	-6.96	36.49
Gran Canaria (GC)	-15.80	28.20
Mahon (MH)	4.42	39.71
Tarragona (TG)	1.47	40.69
Tenerife Sur (TS)	-16.61	28.00
Valencia Copa (VC)	0.20	39.51
Villano Sisargas (VS)	-9.21	43.50

the diagram. The diagnostic table that can be built from these indices includes quantities such as the two-dimensional correlation, two-dimensional RMSE, the rotation angle of the models with respect to the reference dataset, their relative rotation angle, the eccentricity of the ellipses from the model and observations, and the congruence coefficient of the first EOF (Empirical Orthogonal Function) of every model data with respect to the one from the observations. An open software R package has been developed by the authors, and it can be easily installed in R (https://cran.r-project.org/web/packages/SailoR/index.html).

In the following Figure 1 and Table I, an example of these indices generated from synthetic data is shown. These data contain a two-dimensional dataset (Ref) and two models: MOD1, which contains the Ref data plus a constant bias of 8.3 m/s, and MOD2, which has been built by rotating the Ref data 30°.

The Sailor diagram shown in Figure 1 consists of three ellipses representing the Ref data in gray, MOD1 in red, and MOD2 in blue. Each ellipse is formed by considering the directions of the two EOFs existing in the data distribution and their axes are proportional to the standard deviations of every principal component, which allows for the representation of the data's variability and its shape. There are two different options of representing the Sailor diagram. In Figure 1 (a), all the ellipses are centered on the mean of the Ref data, with the averages of the models shown as colored points, while in the Figure 1 (b), each ellipse is centered in its respective average. Figure 2 shows two Taylor diagrams for these two-dimensional data, and it can be observed, among other things, that the Sailor diagram allows for easier analysis of the two-dimensional bias in these synthetic datasets.

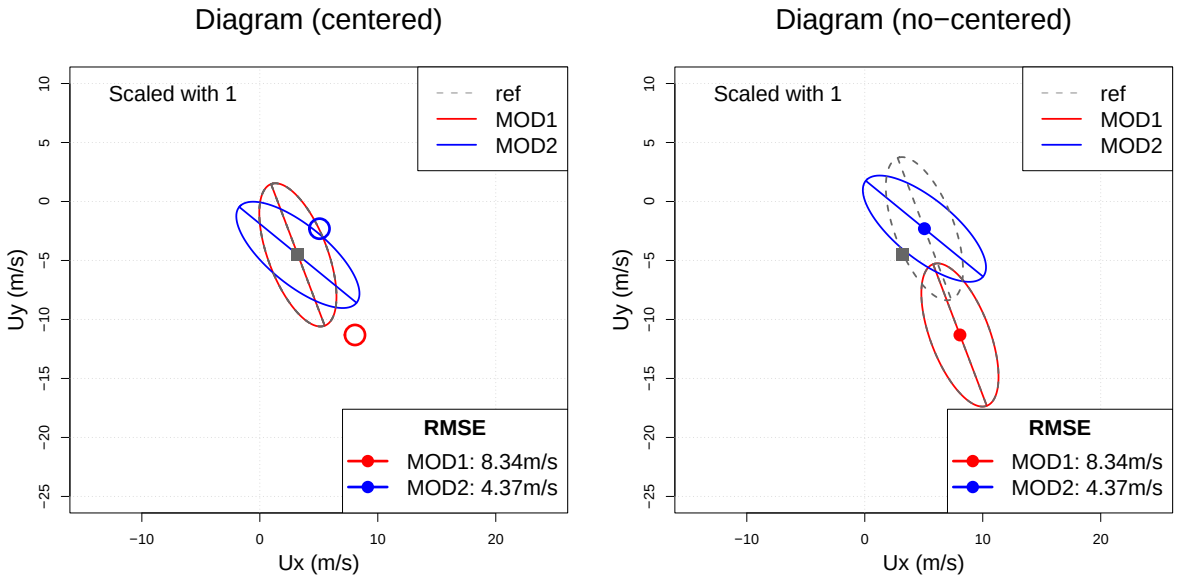
The aim of this study is to validate the two-dimensional wave energy flux variable of the ERA5 reanalysis by comparing it with buoy measurements from the Iberian Peninsula and the Canary Islands over a long period of ten years using to that end the Sailor diagram as the verification tool. This way, we extend the original aim of the SailoR diagram, which was intended for the intercomparison of different models against a unique observation dataset. By this contribution, we show that the methodology can also be applied for the assessment of a vector field from a single model against multiple observations, too.

II. METHODOLOGY

Two different data sources in the period from 2005 to 2014 have been used to carry out this research. The first source consisted of hourly observations obtained from Puertos del Estado (https://www.puertos.es/en-us). Data from 15 REDEXT buoys (Table II) have been used. The acronyms used for the different buoys in the maps are also shown in parenthesis at every buoy.

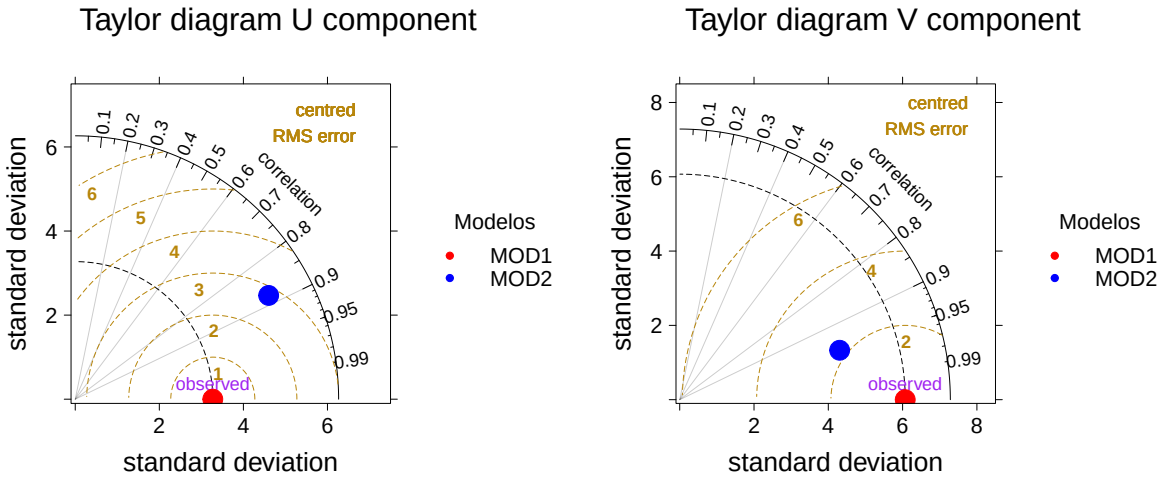
The second source of data was ERA5 reanalysis [9] which was accessed through Copernicus Climate Data Store (https://cds.climate.copernicus.eu/). This reanalysis provides hourly time data with a spatial resolution of 0.5º x 0.5º on a regular longitude and latitude grid. In both cases, the Wave Energy Flux variable (kW/m) was calculated by downloading these three variables: significant wave height, mean wave period and direction of incoming waves from Copernicus Climate Data Store.

The SailoR package, as available from CRAN, was used to obtain the following verification indices by using the time series of each ERA5 grid with the closest buoy series. The starting point of the analysis using the SailoR diagram is that we can use two datasets $\mathbf{U}$ and $\mathbf{V}$ arranged as $N \times 2$ matrices, with N the number of samples (the same for both datasets) and two columns (X and Y components of the WEF data in our case). The reference dataset ($\mathbf{U}$) in our case is taken from in-situ observations of WEF and the $\mathbf{V}$ dataset corresponds to the WEF components computed from ERA5 for every of the closest grid points surrounding the buoy used to derive $\mathbf{U}$. For every pair of $\mathbf{U}$ and $\mathbf{V}$ combinations, the (2×2) matrix representing the mean



(a) (b)

Fig. 1. Centered (a) and no-centered (b) Sailor diagrams built by using the synthetic reference and MOD1 and MOD2 models.



(a) (b)

Fig. 2. Taylor diagrams for U component (a) and V component(b) of synthetic reference and the synthetic MOD1 and MOD2 models.

squared error between them is calculated by means of the expression:

$$\mathbf{\Delta}_{uv}^2 = \frac{1}{N}\left(\mathbf{V}-\mathbf{U}\right)^T\left(\mathbf{V}-\mathbf{U}\right). \quad (1)$$

From this matrix, we derive the scalar mean squared error corresponding to each $\mathbf{U}$ and $\mathbf{V}$ pair by computing its Frobenius norm:

$$\varepsilon^2 = ||\mathbf{\Delta}_{uv}^2||_F. \quad (2)$$

As fully developed in [1], this squared error can be separated into two parts, one derived from the bias and another one derived from the covariances between both datasets, which can also be reduced to the directional analysis between the empirical orthogonal functions describing every dataset. In order to calculate the EOFs, it has to be taken into account that the covariance matrices are of full rank for any realistic dataset in N is larger than 2 samples. Following the strategy mentioned above to the closest grid point in ERA5 to every buoy, the following results, as described in [1], are calculated by using the SailoR package:

- Two-dimensional correlation of grid point time series against the closest buoy.
- RMSE of grid point data against the closest buoy.
- Magnitude of vector bias from every grid point against the closest observation.
- Relative rotation angle from reference EOFs (in-situ data) to nearest model EOFs.

- Eccentricity: his index evaluates the reliability of rotation angles due to the degeneracy of the eigenvalues, if any. This is evaluated both for grid points and in-situ observations.
- Congruence coefficient of the first EOF: This index calculates the consistency of the EOF pairs from the reference (in-situ data) and model (nearest grid-point) datasets.
- Standard deviation of the first and second principal components, both for in-situ and grid points.

The way the SailoR diagram is used in this contribution is ilustrated in Figure 3 for two points in the Bay of Biscay (panel a, Bilbao Vizcaya buoy) and in the Mediterranean (panel b, Cabo Begur). The magnitude of the distance between the colored points and the grey squared represent the bias between the estimations of WEF. The R^2 squared two-dimensional correlation coefficient evaluates the temporal agreement of changes in the underlying time series for both directions. The congruence coefficient addresses the question whether the leading EOFs from model and observations point in the same direction. It corresponds to the absolute value of the inner product of the unit vectors oriented along the semi-major axes of the ellipses. Finally, the relative rotation between the leading EOFs ilustrates the rotation of these major axes with respect to the grey dashed semi-major axis, which corresponds to observations. The overal RMSE value is shown in the legends of the Sailor diagrams in Figure 3. As shown in the figure, it can easily be seen at the first glance that in the BB (Bilbao Vizcaya) case (left panel), the bias of the red point is very small, that the semi-major blue axis is smaller than the observed gray one, and that in general the orientation of the ellipses is good. On the other hand, in the right panel (Mediterranean area, Cabo Begur), the relative rotation is larges, the bias is small for both points and the blue point overestimates the variability in the leading EOF. The information contained in these examples is the one that will be discussed in detail.

As mentioned above, in the case of in-situ data, fifteen buoys (fifteen **U** datasets above) are used (see Table II). However, the figures below show more grid points than fifteen. The values of WEF (zonal and meridional components) derived from ERA5 at every grid point (**V**) are compared with the closest in-situ buoy. The maximum distance between both points is in any case is limited to 3.5° longitude and 1.5° latitude (approximately 385 km and 165 km, respectively), so that ERA5 points beyond this distance from any buoy are not used.

III. Results

The root mean squared error including both zonal and meridional components shows smaller values over the Mediterranean areas (Figure 4a) if compared with the Bay of Biscay or the Atlantic areas. It is also smaller close to the coastal regions over the southwestern Iberian Peninsula and between the Canary Islands and the African coast. However, inferring from this indicator alone that the reanalysis data are better over those areas would be misleading, since the comparison of the reanalysis data with in-situ observations can also include the analysis of the directional fields or the correlation of both WEF components. On the other hand, the magnitude of the bias vector (Figure 4) is generally small except in open areas of the southwestern Iberian Peninsula and in the areas between the Canary Islands an the African coast. The lowest values are found in the Mediterranean areas, with low values, albeit a little bit higher over the Bay of Biscay. The values provided in kW/m in Figure 4 (left panel) can be expressed as relative errors, since the values of WEF are not uniform at every place. The relative errors represent values close to 45 % in the Atlantic facade to 15 % in the Mediterranean areas or 30 % in the Canary Islands. On the other side, the bias represents errors that range from 22 % in the Atlantic areas to 7 % in the Mediterranean or 40 % in the Canary Islands.

The northern areas of the domain in the Atlantic shore (Bay of Biscay) exhibit the highest isotropy in the two-dimensional R^2 field, as shown in Figure 5a. The values of R^2 in those areas achieve values close to two over the majority of the area. It must be noted that for the two-dimensional correlation coefficient, the optimal value of R^2 for linearly dependent signals is 2 [10]. These good results for the squared 2D correlation can be explained by the fact that swell is quite uniform for this area, which tends to produce a spatially uniform field of WEF. The values of the root-mean square error over those areas, however, are relatively high, as previously indicated, something that is consistent with the moderate-high bias inducing the relatively high values of the RMSE errors despite a good agreement of the temporal position of the anomalies and their amplitudes (high values of R^2), as shown by these results. On the other hand, over the Mediterranean, the values of R^2 are limited to reduced areas close to the buoys, and they present distinctive bull-eye like patterns, pointing to the relevance of local factors in this area. In any case, even in areas apart from the in-situ observations, the values of R^2 are still higher than one over the whole region apart from the area close to Gibraltar and areas in the Mediterranean far from the buoys. Finally, there exist strong horizontal gradients of R^2 and RMSE in areas close to the Gibraltar Strait or the Canary Islands. Over these areas the spatial representativeness of in-situ observations is relatively low. This can be explained by the shadow effect of the islands and the strait itself that modify the general WEF flow when analyzed at a very local level.

Before proceeding to the analysis of the directional errors, we have also calculated the eccentricity ε of the distribution of WEF over the grid points. This parameter describes the eccentricity of the ellipse defined by the principal components of the observations and model results (in this case, ERA5 WEF) in the space spanned by both principal components [1]. The reason for analyzing this parameter is that, if in a given area the eccentricty is close to zero, the EOFs will form a degenerated multiplet and it will be impossible to correctly identify the EOFs. In all areas, the eccentricity (Figure 5b) is clearly different from zero, which means

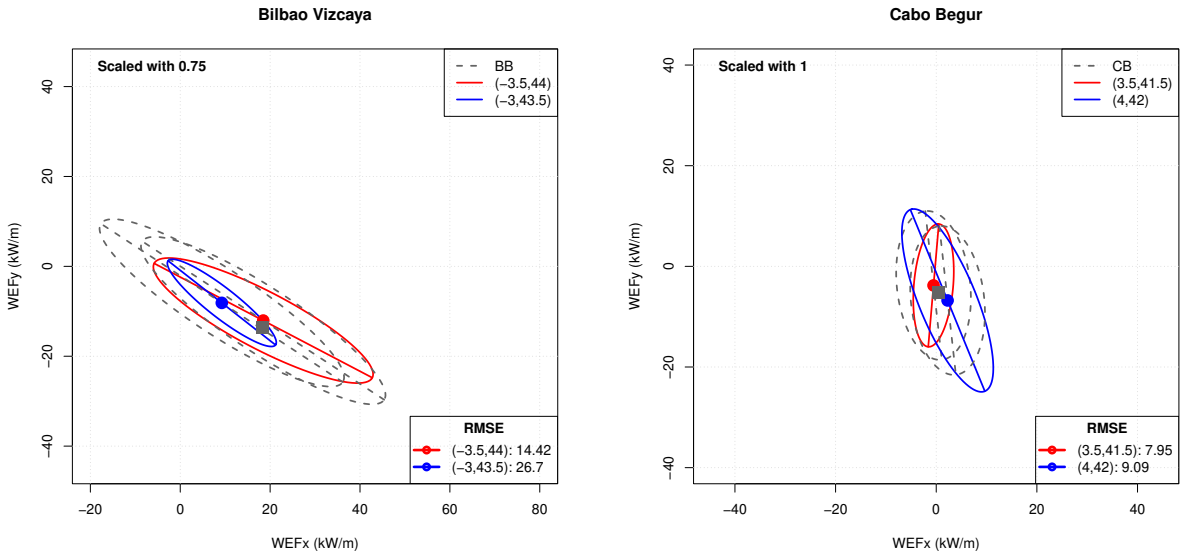


Fig. 3. Sample (non-centered) SailoR diagrams for WEF constructed by using two points close to the Bay of Biscay buoy (panel a) and Cabo Begur (CB), in the Mediterranean (panel b).

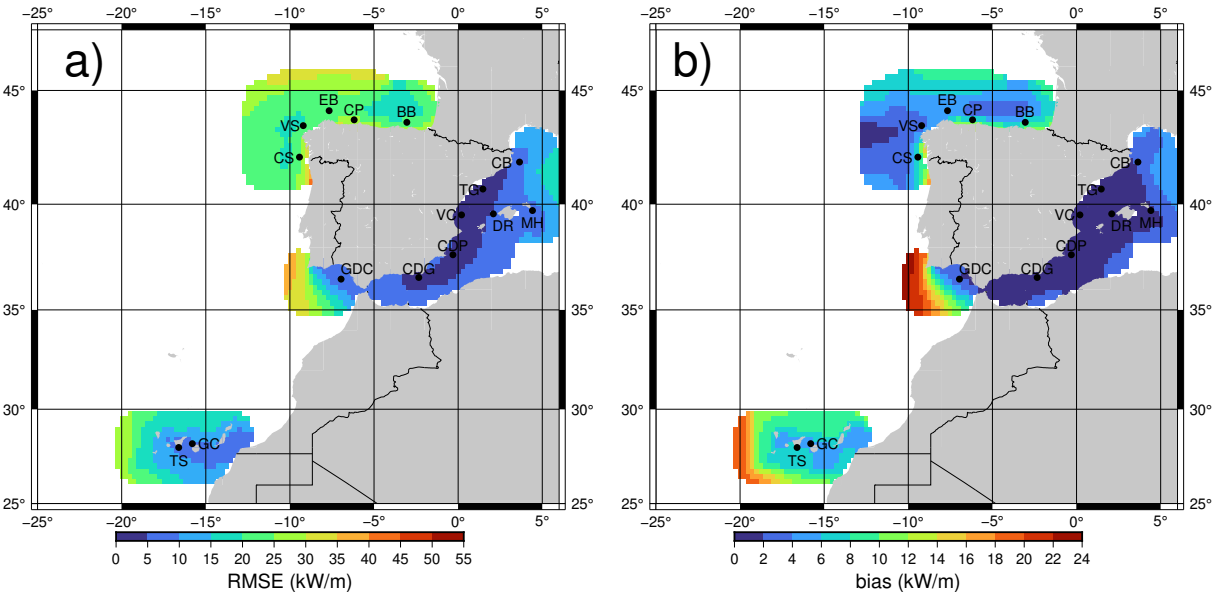


Fig. 4. RMSE corresponding to both zonal and meridional components (panel a) and magnitude of two-dimensional bias (panel b).

that we have a strong confidence that the estimation of directional errors will be robust [1]. In this case, since the eccentricity ε can be identified both in the in-situ observations and the grid points, the values from the in-situ dataset are showed as circles drawn using the same color as the underlying grid points. It can be seen that the eccentricity of the model fields matches very good the one in the observations except in Dragonera (DR), Tarragona (TG) and Tenerife Sur (TS) (Figure 5b).

Therefore, as could be expected from the fact that the directional structure of the data is quite different from a circumference (high values of the eccentricity), the robustness of the identified EOFs is high. Besides that, there is a very good match of the leading EOF of the in-situ observations of WEF and its values as simulated by the wave model in ERA5 (Figure 6a). There are only a few areas in the Mediterranean and in the Canary Islands in which the simulated values of WEF do not achieve absolute values of the congruence coefficients [1] $|g_{11}| = |\mathbf{e}_{u1} \cdot \mathbf{e}_{v1}|$ (absolute value of the inner product of the leading EOF of observations and simulations) higher than 0.8. This shows that there is in general a very good agreement of the directional distribution of WEF values. As a consequence, the directional errors, expressed as the relative rotation of the leading EOF of model with respect to the one corresponding to the closest observations [1] are in general low in Atlantic areas or the Bay of Biscay. There are higher directional errors over the Mediterranean or the Canary Islands, in areas severely affected by local factors associated to the position of the sensors with respect to the coasts and the main areas of generation of the waves.

Finally, after having analyzed the directional agreement of the in-situ and reanalysis WEF data, we diagnose the standard deviations of the corresponding

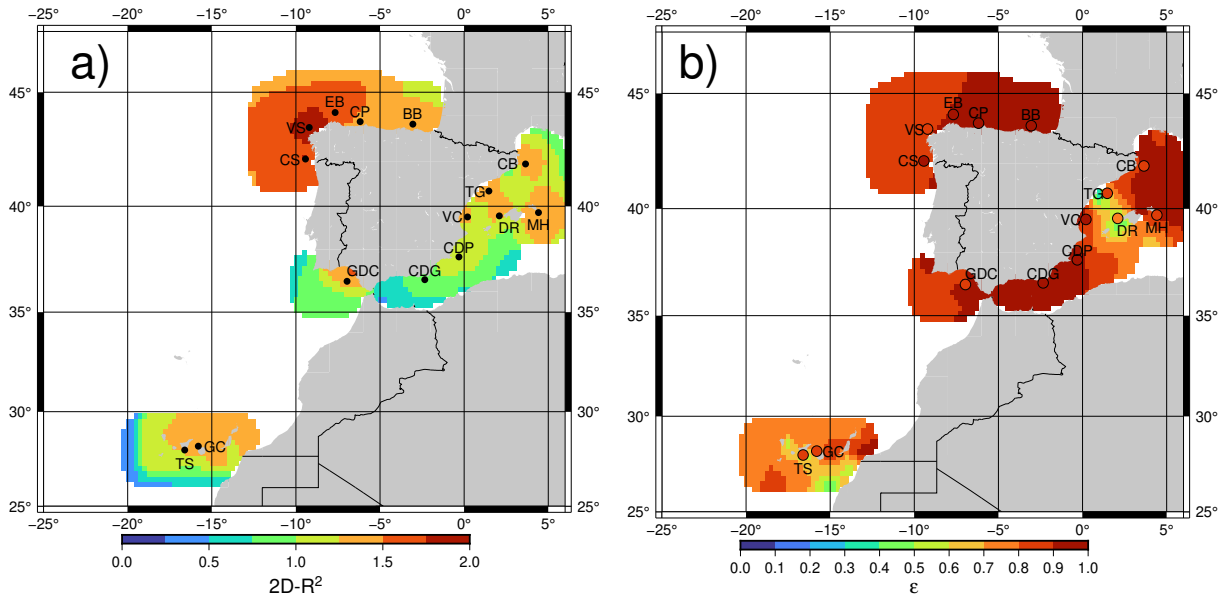


Fig. 5. Two-dimensional squared correlation R^2 (panel a) between the in-situ WEF data and ERA5 estimations of WEF at every grid point. Eccentricity ε of the ellipse spanned by both EOFs for the ERA5 data at every grid point (panel b).

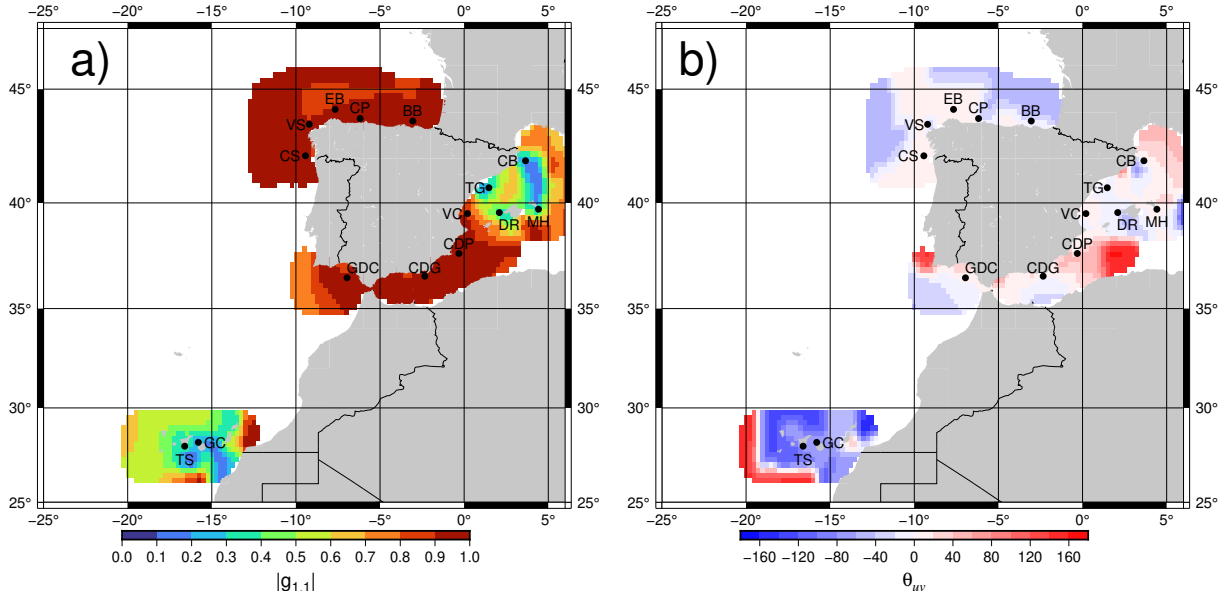


Fig. 6. Absolute value of the congruence coefficient $|g_{1,1}|$ (panel a) between the leading EOF corresponding to in-situ and ERA5 estimations of WEF at every grid point. Relative rotation angle θ_{uv} between the estimations of WEF by ERA5 relative to the observed ones (panel b).

leading and second principal components in Figure 7.

It can be seen that the model slightly underestimates the standard deviation of the WEF field over the Bay of Biscay and the Atlantic, while it represents much better the structure of this leading principal component in the Mediterranean (Figure 7). In the case of the second principal component, the agreement of the model with observations is in general better over the whole domain. The values shown in Figure 7 correspond roughly to 70 % of the average WEF in the Atlantic facade, the 10 % of the average WEF in the Mediterranean areas and 6 % of the average WEF values around the Canary Islands.

IV. Discussion

The use of the SailoR diagram in the verification of vector fields allows us to make important advances in the diagnostics of the relative merits of the simulations in terms of several characteristics.

First, it has been noted that the raw use of the RMSE error would yield a very simplistic scenario which would indicate that the simulation agrees much better with the observations in the Mediterranean.

However, the complete analysis of the squared correlation (Figure 5), congruence coefficients and the relative rotation indicates that the simulation over the Atlantic yields better results in terms of two-dimensional correlation and directional errors. This shows that the smaller RMSE in the Mediterranean despite larger directional errors and similar biases is to some extent a consequence of the fact that the variance of the WEF variability is larger in the Atlantic coast and the Bay of Biscay, so that smaller relative errors end up building a larger RMSE error.

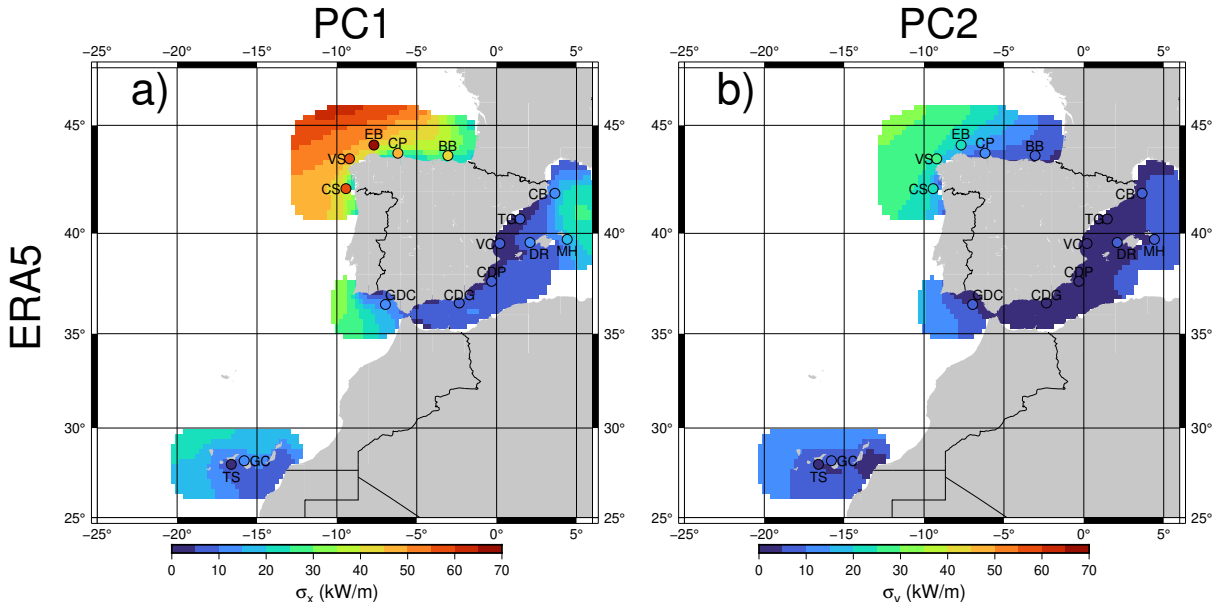


Fig. 7. Standard deviations of the leading (panel a) and second (panel b) principal component of the two-dimensional WEF field for the model (grid points) and observations (colored circles).

V. Conclusions

The subset of the verification indices provided by the SailoR package used in this paper for the analysis of the agreement between wave energy flux computed from ERA5 data and observed by buoys around the Spanish coast shows that they allow to identify areas with higher or smaller representativeness of the sampling due to the existence of local effects. This is reflected in different components of the error such as bias, root-mean square error or two-dimensional correlation. Beyond the simple analysis of the difference between model-based estimations of WEF with respect to their observational counterparts, the use of the indices calculated by means of the SailoR package allows us to make a deeper analysis of the error, identifying differences in the way the differences between model and observations affect terms such as the bias, the correlation of both components of the vector field, the orientation of the major axis of the distribution of vectors or their relative rotation. This allows the scientists to analyse vector fields to achieve a better understanding of the relative contributions of each term to the overall error. This is particularly important in variables such as WEF, in which directional errors might be important.

Acknowledgement

The authors would like to express their gratitude to the Spanish Puertos del Estado https://www.puertos.es/en-us for kindly providing with data for this study. All the calculations have been carried out in the framework of R [11], including the Sailor package, developed by authors [1]. The maps have been produced by using the Generic Mapping Tools system [12].

References

[1] J. Sáenz, S. Carreno-Madinabeitia, G. Esnaola, S. J. González-Rojí, G. Ibarra-Berastegi, and A. Ulazia, "The Sailor diagram–a new diagram for the verification of two-dimensional vector data from multiple models," *Geoscientific Model Development*, vol. 13, no. 7, pp. 3221–3240, 2020.

[2] H. Brooks, B. Brown, B. Ebert, C. Ferro, I. Jolliffe, T.-Y. Koh, P. Roebber, and D. Stephenson, "WWRP/WGNE joint working group on forecast verification research," *Collab. Aust. Weather Clim. Res. World Meteorol. Organ*, 2015.

[3] H. R. Stanski, L. J. Wilson, and W. R. Burrows, *Survey of common verification methods in meteorology*. Geneva: World Meteorological Organization, 1989.

[4] K. E. Taylor, "Summarizing multiple aspects of model performance in a single diagram," *Journal of geophysical research: atmospheres*, vol. 106, no. D7, pp. 7183–7192, 2001.

[5] A. Ulazia, J. Sáenz, G. Ibarra-Berastegui, S. J. González-Rojí, and S. Carreno-Madinabeitia, "Using 3DVAR data assimilation to measure offshore wind energy potential at different turbine heights in the West Mediterranean," *Applied energy*, vol. 208, pp. 1232–1245, 2017.

[6] P. Lorente, S. Piedracoba, J. Soto-Navarro, and E. Alvarez-Fanjul, "Evaluating the surface circulation in the Ebro delta (northeastern Spain) with quality-controlled high-frequency radar measurements," *Ocean Science*, vol. 11, no. 6, pp. 921–935, 2015.

[7] T. Lee, D. E. Waliser, J.-L. F. Li, F. W. Landerer, and M. M. Gierach, "Evaluation of CMIP3 and CMIP5 wind stress climatology using satellite measurements and atmospheric reanalysis products," *Journal of Climate*, vol. 26, no. 16, pp. 5810–5826, 2013.

[8] D. Walters, M. Best, A. Bushell, D. Copsey, J. Edwards, P. Falloon, C. Harris, A. Lock, J. Manners, C. Morcrette *et al.*, "The Met Office Unified Model global atmosphere 3.0/3.1 and JULES global land 3.0/3.1 configurations," *Geoscientific Model Development*, vol. 4, no. 4, pp. 919–941, 2011.

[9] H. Hersbach, C. Peubey, A. Simmons, P. Berrisford, P. Poli, and D. Dee, "ERA-20CM: a twentieth-century atmospheric model ensemble," *Quarterly Journal of the Royal Meteorological Society*, vol. 141, no. 691, pp. 2350–2375, 2015.

[10] D. S. Crosby, L. C. Beaker, and W. H. Gemill, "A proposed definition for vector correlation in geophysics: Theory and application," *Journal of Atmos. and Oceanic Technol.*, vol. 10, pp. 355–367, 1993.

[11] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2022. [Online]. Available: https://www.R-project.org/

[12] P. Wessel, J. F. Luis, L. Uieda, R. Scharroo, F. Wobbe, W. H. F. Smith, and D. Tian, "The Generic Mapping Tools version 6," *Geochemistry, Geophysics, Geosystems*, vol. 20, p. 5556–5564, 2019.